INT. J. PROD. RES., 2000, VOL. 38, NO. 4, 765–791

Validating the Malcolm Baldrige National Quality Award Framework through structural equation modelling

SIME CURKOVIC†*, STEVE MELNYK‡, ROGER CALANTONE‡ and ROBERT HANDFIELD‡

Since its introduction in 1987, the Malcolm Baldrige National Quality Award (MBNQA) has come to have a major influence in the assessment of quality efforts and systems. Many public and private institutions, for example, use the measures derived from this award programme to carry out self-assessments. Many programmes at the local, state and international levels are based on the MBNQA. Implicit in this evidence is the assumption that the MBNQA adequately captures the major dimensions of Total Quality Management. This study empirically assesses this assumption. Specifically, it assesses the extent of fit between the factors of the MBNQA and their measures. It also evaluates the extent to which these factors really do capture this important higher-level construct known as TQM. The results reported in this study are based on a field survey consisting of responses gathered from 526 plant managers within the US automotive industry. The assessment is carried out using confirmatory factory analysis and structural equation models.

1. Introduction

Since its introduction in 1987, the Malcolm Baldrige National Quality Award (MBNQA) has come to have a major impact on how quality is viewed and, more importantly, assessed. This impact is most clearly manifested through the various criteria and measurement instruments developed for and used by this programme. These criteria and measures are used to assess the performance of the firm on the various dimensions of quality. They are also used to evaluate how effectively and efficiently the firm has been able to use quality both as a strategic vehicle and as a means of satisfying its customers' needs. These criteria and measures form the means by which progress on quality and customer satisfaction are measured and evaluated.

It can be argued that one of the major contributions of the MBNQA has been the establishment of these criteria and measures. There is significant evidence that these elements have achieved widespread acceptance. For example, there is the Baldrige 'effect', whereby firms use these criteria and measures to perform critical self-assessments (Steeples 1993, Ettorre 1996). A wide variety of public and private organizations, including corporations, governments, hotels, schools, religious affiliations, and prisons, are now applying the MBNQA criteria. In addition, the MBNQA criteria and measurements have been adopted at the international, national, state, and local

Revision received July 1999.
† Western Michigan University, Haworth College of Business, Management Department, Kalamazoo, MI 49008, USA.
‡ Michigan State University, The Eli Broad Graduate School of Management, Department of Marketing & Supply Chain Management, East Lansing, MI 48824, USA.
* To whom correspondence should be addressed. e-mail: sime.curkovic@wmich.edu

International Journal of Production Research ISSN 0020–7543 print/ISSN 1366–588X online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journals/tf/00207543.html

levels. For example, the Egyptian National Quality Award and the Greater Memphis (Tennessee) Award for Quality are based in large part on the MBNQA. It has been estimated that over 40 such programmes can trace their roots to the MBNQA. Even Japan has been looking at the possibility of introducing it own version of the MBNQA (Ettorre 1996).

This activity would suggest that the MBNQA and its criteria have comprehensively captured the major dimensions of a Total Quality Management (TQM) system, as envisioned by such authors as Deming (1982, 1986), Juran and Gryna (1980), and Garvin (1988). Currently, the support for this premise is based on case and anecdotal evidence drawn from many sources, including the winners of this award and their experiences. To date, however, there has been no statistical study that has attempted to validate the extent of fit between the underlying factors associated with the MBNQA and their measures and the extent to which these factors really do capture this important but higher-level factor known as TQM. These goals form the major objectives of this study.

Specifically this paper will focus on the following objectives.

- To provide an overall empirical assessment of the Total Quality Management construct.
- To compare the MBNQA framework with other competing frameworks drawn from research to determine which framework conceptually best fits the needs of a TQM system.
- Empirically to assess the MBNQA, through both confirmatory factor analysis and structural equation modelling, in terms of its ability to capture both this unobserved but critical latent variable known as Total Quality Management and the TQM process.

This study begins by reviewing the TQM literature. From this review, the elements of a TQM system are identified. This section concludes with a comparative assessment of the MBNQA with seven other TQM frameworks to determine which framework best fits the conceptual requirements of a TQM system. In the second section, a detailed discussion of the MBNQA framework is presented and then operationalized based on the award criteria. In the third section, a test of the MBNQA framework according to a structural equation modelling approach is performed. Results of the assessment of the psychometric properties of the model obtained from the test are then presented. The final section concludes with implications for managerial decision-making and recommendations for future research.

2. Total Quality Management—defining the concept and system

Based on the pioneering work of Deming, the term 'Total Quality Management' (TQM) emerged over a decade ago in the US and embodied a broad scope of activities within the framework of world class manufacturing (Deming 1981, 1982, 1986). TQM itself is an integrated management philosophy and set of practices that establishes an organization-wide focus on quality, merging the development of a quality-oriented corporate culture with intensive use of management and statistical tools aimed at designing and delivering quality products to customers (Melnyk and Denzler 1996). TQM stresses three major principles: customer satisfaction, employee involvement, and continuous improvements in quality. TQM also involves benchmarking, product and service design, process design, long-range thinking, and problem-solving tools.

Logothetis (1992) describes TQM as 'a culture; and inherent in this culture is a total commitment to quality and attitude expressed by everybody's involvement in the process of continuous improvement of products and services, through the use of innovative scientific methods.' Perhaps one of the most salient definitions of TQM was provided by the *Report of the Total Quality Leadership Steering Committee and Working Councils* (Evans 1992). This council consists of a number of CEOs from major corporations, as well as a number of academic representatives from distinguished schools across the country. This council developed the following definition of TQM:

> a people-focused management system that aims at the continual increase of customer satisfaction at continually lower real cost. Total Quality is a total system approach (not a separate area or program), and an integral part of high-level strategy; it works horizontally across functions and departments, involves all employees, top to bottom, and extends backwards and forwards to include the supply chain and customer chain. (Evans 1992)

2.1. *Categories of frameworks*

Three different categories of frameworks emerge from the literature: (1) anecdotal; (2) empirically-based research; and (3) formal assessment processes. The anecdotal category is based on the personal experiences of quality gurus such as Deming, Juran and Crosby. The empirically-based research category looks at formal research studies with increased attention given to studies whose thrust was the development and validation of TQM measurement instruments (e.g. Saraph *et al.* 1989, Flynn *et al.* 1994). The third and final category examines popular formal international and national assessment processes such as ISO 9000 and the Malcolm Baldrige National Quality Award. These three categories of frameworks are used to identify the various views of the constructs associated with TQM and the items that represent manifestations of these constructs.

2.2. *An operational framework of TQM*

The beginning of this section offered several definitions of TQM (e.g. Evans 1992, Logothetis 1992, Melnyk and Denzler 1996). A series of associated traits were identified from these definitions of TQM: (1) continuous improvement; (2) meeting customers' requirements; (3) long-range planning; (4) increased employee involvement; (5) process management; (6) competitive benchmarking; (7) team-based problem-solving; (8) constant measurement of results; (9) closer relationships with customers; and (10) management commitment. Using these traits, the various frameworks and their associated constructs were assessed to determine which framework best fits the definition of TQM (see table 1).

From this table, it can be seen that the MBNQA framework best fits the definition of TQM. As such, the MBNQA framework will be used as the operational framework of TQM for the purposes of this study. Many other researchers (e.g. Dean and Bowen 1994, Black and Porter 1996, Capon *et al.* 1994) have also adopted the MBNQA framework as the basic operational model of TQM. Given its importance, we now examine the MBNQA framework in detail.

Associated Traits of TQM	Juran	Deming	Crosby	Saraph *et al.* (1989)	Flynn *et al.* (1994)	Powell (1995)	ISO 9000	MBNQA
Continuous improvement	×	×	×	×	×			×
Meeting customer's requirements	×	×			×	×	×	×
Long-range planning	×			×				×
Increased employee improvement	×	×	×	×		×		×
Process management		×		×	×	×	×	×
Competitive benchmarking				×		×		×
Team-based problem-solving	×		×	×		×		×
Constant measurement of results	×		×	×	×	×	×	×
Closer relationships with customers	×	×			×	×		×
Management commitment	×	×	×	×	×	×	×	×
10 Traits total:	8	6	5	8	6	8	4	10

Table 1. The best definitions of TQM.

2.3. *The Malcolm Baldrige National Quality Award (MBNQA)*

In 1987, former President Ronald Reagan signed the Malcolm Baldrige National Quality Improvement Act. The act established an annual national award to recognize quality improvement among manufacturing, service, and small businesses. The MBNQA was established for three primary reasons. One reason was to raise the consciousness of US business leaders regarding the issue of quality. Another reason was to provide a comprehensive framework for measuring the quality efforts of US businesses. Lastly, it was established to provide US businesses with a template for a thorough TQM system (Hockman 1992).

The MBNQA encourages organizations to address quality on a broad range of issues. Companies that wish to compete for the award must produce evidence of leadership and long-term planning, initiate verifiable quality control procedures, address the happiness and well-being of the workforce and, above all, work toward customer satisfaction. The criteria argue strongly for customer-driven organizations, high levels of employee involvement, and information-based management. For companies not competing for the award, the application provides a framework for implementing a quality programme and establishes the benchmarks for measuring future progress.

The Baldrige Award is composed of seven separate, weighted categories: (1) leadership; (2) strategic planning; (3) customer and market focus; (4) information and analysis; (5) human resource development and management; (6) process management; and (7) business results. These categories are described in detail in table 2.

1.0 **Leadership** (110 points): The *Leadership* category examines senior leaders' personal leadership and involvement in creating and sustaining values, company directions, performance expectations, customer focus, and a leadership system that promotes performance excellence. Also examined is how the values and expectations are integrated into the company's leadership system, including how the company continuously learns and improves, and addresses its societal responsibilities and community involvement.

1.1 Leadership System (80 points)
1.2 Company Responsibility and Citizenship (30 points)

2.0 **Strategic Planning** (80 points): The *Strategic Planning* category examines how the company sets strategic directions, and how it determines key action plans. Also examined is how the plans are translated into an effective performance management system.

2.1 Strategy Development Process (40 points)
2.2 Company Strategy (40 points)

3.0 **Customer and Market Focus** (80 points): The *Customer and Market Focus* category examines how the company determines requirements and expectations of customers and markets. Also examined is how the company enhances relationships with customers and determines their satisfaction.

3.1 Customer and Market Knowledge (40 points)
3.2 Customer Satisfaction and Relationship Enhancement (40 points)

4.0 **Information and Analysis** (80 points): The *Information and Analysis* category examines the management and effectiveness of the use of data and information to support key company processes and the company's performance management system.

4.1 Selection and Use of Information and Data (25 points)
4.2 Selection and Use of Comparative Information and Data (15 points)
4.3 Analysis and Review of Company Performance (40 points)

5.0 **Human Resource Development and Management** (100 points): The *Human Resource Development and Management* category examines how the work force is enabled to develop and utilize its full potential, aligned with the company's objectives. Also examined are the company's efforts to build and maintain an environment conducive to performance excellence, full participation, and personal and organizational growth.

5.1 Work Systems (40 points)
5.2 Employee Education, Training, and Development (30 points)
5.3 Employee Well-Being and Satisfaction (30 points)

6.0 **Process Management** (100 points): The *Process Management* category examines the key aspects of process management, including customer-focused design, product and service delivery processes, support processes, and supplier and partnering processes involving all work units. The category examines how key processes are designed, effectively managed, and improved to achieve better performance.

6.1 Management of Product and Service Processes (60 points)
6.2 Management of Support Processes (20 points)
6.3 Management of Supplier and Partnering Processes (20 points)

7.0 **Business Results** (450 points): The *Business Results* category examines the company's performance and improvement in key business areas - customer satisfaction, financial and marketplace performance, human resource, supplier and partner performance, and operational performance. Also examined are performance levels relative to competitors.

7.1 Customer Satisfaction Results (130 points)
7.2 Financial and Market Results (130 points)
7.3 Human Resource Results (35 points)
7.4 Supplier and Partner Results (25 points)
7.5 Company-Specific Results (130 points)

Source: Malcolm Baldrige National Quality Award 1997 Award Criteria. Milwaukee, WI: The Malcolm Baldrige National Quality Award Consortium, Inc., 1997.

Table 2. The Malcolm Baldrige National Quality Award—1000 Points Possible.

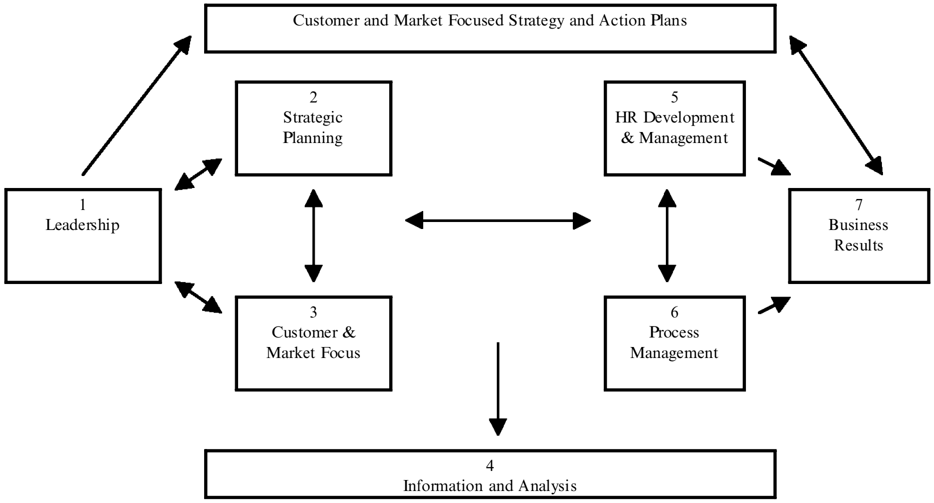


Figure 1. The Malcolm Baldrige National Quality Award Criteria: a systems perspective (1997).

Every two years, the MBNQA criteria are reviewed for potentially major revisions. Suggestions for revisions and improvements are solicited each year from examiners, judges, winners and applicants. Figure 1 describes the relationship between the MBNQA categories.

3. The operationalization of TQM

The TQM construct will be conceptualized in terms of the four basic factors described by the MBNQA framework. The 1997 MBNQA framework is described as three related subsystems (Evans 1997): (1) the 'strategic' categories of leadership, strategic planning, and customer/market focus; (2) the 'operational' categories of human resource development and process management (which lead to 'results'); and (3) the 'information' category that serves as a foundation for the other two subsystems. In summary, TQM is hypothesized to consist of the following four factors: (1) TQM Strategic Systems; (2) TQM Operational Systems; (3) TQM Information Systems; and (4) TQM Results. These factors and their proposed measures span the entire range of activities deemed critical by the MBNQA framework (see Appendix A for definitions of these factors and the selection of items).

To this point, we have established the fit between the MBNQA and TQM from a theoretical/conceptual perspective. We have yet to assess this fit from an empirical foundation. That task is left to the next section.

4. Research design

4.1. *Data collection*

In past research studies of TQM, case studies and anecdotal examples have been the predominant research methodology (Ahire *et al.* 1995). For the purposes of this study, a large scale survey was used. More specifically, an investigation of the underlying constructs was undertaken using a two-phase approach: (1) preliminary scale development was conducted using interviews from managers to provide assistance in the identification and prima facie validation of the constructs and variables in the study; and (2) implementation of a large-scale survey designed to validate scales for measuring the underlying constructs associated with TQM. This combination allowed for the exploitation of the strengths of both case studies and surveys while reducing the problems associated with both.

The primary objective of the first phase was to provide an indication of content validity, rather than build new theory. Interviews with managers in six North American manufacturing facilities were used to provide assistance in the identification and prima facie validation of the constructs and variables in the study. The scales had to be pretested for content validity before any refinement or validation was undertaken. Since items corresponding to the various constructs of the measurement instrument were derived from a comprehensive analysis of the literature, content validity was more adequately assured. However, the research questionnaire was also validated for comprehensiveness and completeness in advance through interviews with industry managers. Each manager completed the questionnaire and provided feedback regarding the wording of items, their understandability, and the overall organization of the instrument. The measurement instrument was adjusted accordingly based on their feedback.

Similar to much of the research in operations strategy, a single industry was chosen for the study (Swamidass and Newell 1987, Vickery *et al.* 1993, Whybark and Vastag 1993, Ahire *et al.* 1996). This restriction permitted the control of several potential confounding variables that often differ between industries, including the scope and complexity of quality issues. The automotive industry was selected because it has been a leader in implementing progressive quality management strategies in the United States (Cole 1990). More specifically, the sample was targeted across a 4-digit SIC code within the automotive industry—Motor Vehicle Parts & Accessories (SIC 3714); establishments primarily engaged in manufacturing motor vehicle parts and accessories, but not engaged in manufacturing complete motor vehicles or passenger car bodies (i.e. air brakes, axle housings, brake drums, bumpers, camshafts, engines, exhaust systems, fuel pumps, manifolds, mufflers, etc). An updated and comprehensive database of 2945 manufacturing facilities within this SIC code was obtained from Elm International (East Lansing, MI). The industry has already been the focus of many empirical studies which address quality management (Womack *et al.* 1990, Ahire *et al.* 1996, Curkovic *et al.* 1997).

The unit of analysis for empirical validation is the individual plant. The plant is the level of implementation for most quality management programmes, and has been used in numerous other empirical studies related to quality (Schonberger 1983, Griffin 1988, Ebrahimpour and Withers 1992, Flynn *et al.* 1994, Ahire *et al.* 1996). Furthermore, studies have shown (e.g. Saraph *et al.* 1989) that quality investments vary between plants within the same firm, indicating that a more aggregated unit of analysis, such as the parent firm level, would likely obscure important differences.

Only single respondents (plant managers) were targeted for the study. The pretest revealed that plant managers would be qualified to answer the questions objectively, while responses from quality specialists would have been biased. Plant managers have also been used as key respondents in other studies for TQM (e.g. Schonberger 1983, Griffin 1988, Ahire, Golhar *et al.* 1996). Research suggests that greater attention to informant selection can help to overcome the common method variance problem when practical considerations require single respondents. Ideally, information should be gathered from multiple respondents at each site to minimize the potential for bias from a single respondent. However, the cost and time

	Mean	Standard deviation	Median	Minimum	Maximum
Respondent's experience in current position (years)[a]:	5.898	5.186	4	0.7	30
Number of Employees[b]:	309.344	297.800	221	15	1975
Plant Size (square feet)[c]:	171 464.559	184 711.979	120 000	10 000	1 500,000
1995 sales volume ($)[d]:	53 997 726.25	66 497 170.36	35 000 000	1 000 000	640 000 000
1996 sales volume ($)[e]:	61 764 063.31	76 658 230.06	40 000 000	1 000 000	690 000 000
Average age of production euipment (years)[f]:	10.070	7.200	8	0.5	50

Held title of plant manager:	113
Held other title (e.g. V.P., President, C.E.O, G.M., etc.):	152
No responses to title:	4

Union representation:	93
Non-union representation:	176

Number of plants by region: Michigan (67); Ohio (20); Illinois (10); Kentucky (10); Indiana (9); Tennessee (9); Virginia (7); Arkansas (6); North Carolina (5); Georgia (4); Missouri (4); Wisconsin (4); Connecticut (3); New York (3); Pennsylvania (3); Texas (3); California (2); Iowa (2); New Hampshire (2); Florida (1); Louisiana (1); Minnesota (1); Mississippi (1); Nebraska (1); Oklahoma (1); South Carolina (1); South Dakota (1)

Parent firm‡:

Publicly traded	110
Foreign-owned subsidiary/transplant	35
Privately owned	121
Joint venture	10

[a] $n = 263$, [b] $n = 183$, [c] $n = 177$, [d] $n = 232$, [e] $n = 232$, [f] $n = 266$

† n varies because data elements were unavailable for some observations

‡ Note that more than one type of ownership might apply to a parent firm.

Table 3. Group 1 respondents ($n = 269$).

associated with obtaining access to individuals from large numbers of large-sized plants in a specific SIC code would be beyond those available for this study. Such a strategy was also not adopted because the response rate would likely be depressed to a critical level.

4.2. *Data collection procedure described*

A two-wave mailing, with reminder postcards sent between each wave, employed many of the techniques developed by Dillman (1978), and resulted in the return of

269 and 257 usable surveys from the first (Group 1) and second (Group 2) wave of responses, respectively. Each survey was mailed directly to the plant manager. The first mailed research questionnaire was accompanied by an explanatory letter, and then a reminder postcard was sent to those plants not yet responding one week following the first mailing. Finally, two weeks after the original packet was mailed, another survey was mailed to the non-respondents. This yielded a combined and overall response rate of 17.86%. Descriptive statistics of the survey respondents for Groups 1 and 2 are provided in tables 3 and 4, respectively.

	Mean	Standard deviation	Median	Minimum	Maximum
Respondent's experience in current position (years)[a]:	6.047	5.708	4	0.5	40
Number of Employees[b]:	362.725	456.281	235	24	3500
Plant Size (square feet)[c]:	184 578.4	240 612.387	115 000	13 000	2 000,000
1995 sales volume ($)[d]:	60 511 771.69	101 559 088.4	30 000 000	1 000 000	1 076 593 000
1996 sales volume ($)[e]:	67 409 102.4	104 760 507.6	32 000 000	1 800 000	998 789 000
Average age of production euipment (years)[f]:	11.034	8.429	10	1	50

Held title of plant manager:	96
Held other title (e.g. V.P., President, C.E.O, G.M., etc.):	160
No responses to title:	1
Union representation:	82
Non-union representation:	175

Number of plants by region: Michigan (36); Ohio (18); Illinois (17); Indiana (14); Tennessee (11); Kentucky (10); Pennsylvania (5); Virginia (5); Wisconsin (4); Georgia (3); Massachusetts (3); North Carolina (3); California (2); Florida (2); Iowa (2); Minnesota (2); Missouri (2); New York (2); Texas (2); Arkansas (1); Connecticut (1); New Hampshire (1); Oklahoma (1); South Carolina (1); South Dakota (1); Utah (1)

Parent firm‡:

Publicly traded	97
Foreign-owned subsidiary/transplant	53
Privately owned	121
Joint venture	14

[a] $n = 251$, [b] $n = 149$, [c] $n = 145$, [d] $n = 219$, [e] $n = 223$ [f] $n = 249$

† n varies because data elements were unavailable for some observations.

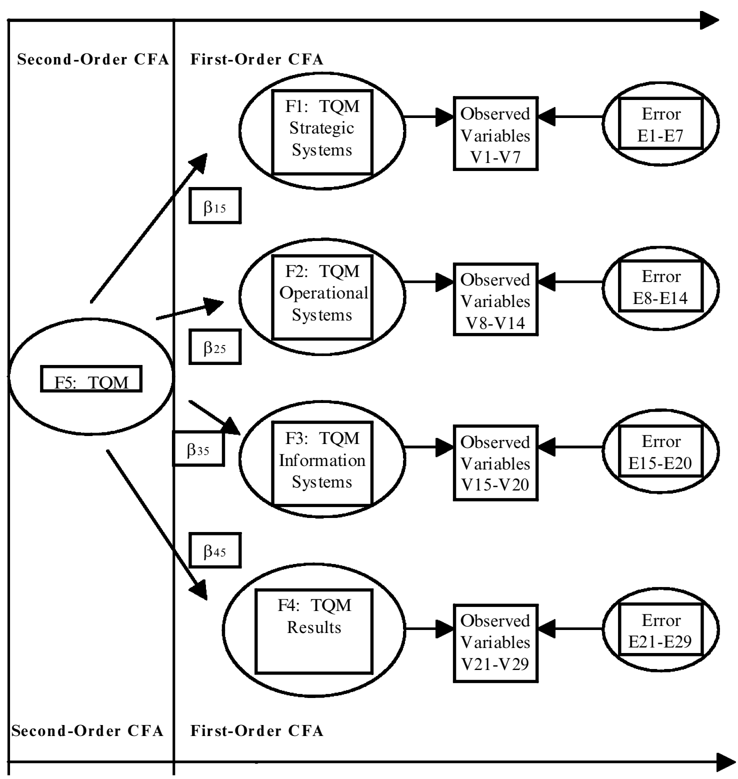
‡ Note that more than one type of ownership might apply to a parent firm.

Table 4. Group 2 respondents ($n = 2579$).

5. Data analysis

The data generated in the preceding stage were analysed using confirmatory factor analysis and structural equation modelling (SEM). SEM was selected because it was most appropriate for this study, given its objectives. In assessing empirically the extent to which the MBNQA framework really captures TQM, it is important to recognize that we are evaluating a series of relationships between manifest variables (as represented by the criteria and measures used by the MBNQA) and latent variables. This mapping also extends to relationships between latent variables and a higher-order factor referred to as TQM. The analysis of this structure is best accomplished by SEM. More important, in studying this structure, we must realize that we are studying the process by which TQM is achieved. Again, the process of studying this process is best handled by SEM because it allows us to examine multiple relationships simultaneously while also incorporating measurement error into the estimation process.

The hypothesized overall model is portrayed in figure 2 in terms of EQS notation. It represents a typical covariance structure model. In figure 2, there is a second-order TQM confirmatory factor analysis (CFA) measurement model. The single-headed arrows leading from the second-order factor of TQM (F5) to each of its underlying first-order factors (F1,F5; F2,F5; F3,F5; F4,F5) are regression paths that indicate the prediction of TQM Strategic Systems (F1), TQM Operational Systems (F2),



* Each first-order factor has a disturbance term associated with it.

Figure 2. TQM measurement model.

TQM Information Systems (F3), and TQM Results (F4) from a higher-order TQM factor. They also represent second-order factor loadings. Finally, there is a residual disturbance term associated with each first-order factor (e.g. D1, D2, D3, D4). These represent residual errors in the prediction of the first-order factors from the higher-order factor of TQM and are not shown in figure 2. However, it is assumed that any variation occurring due to factors not included in figure 2 is due to these disturbance terms.

Expressed more formally, the CFA model portrayed in figure 2 hypothesizes *a priori* that: (1) TQM can be conceptualized in terms of four factors each; (2) each observed variable will have a non-zero loading on the factor it was designed to measure and zero loadings for all other factors; (3) error terms associated with each observed variable will be uncorrelated; (4) the four first-order factors will be correlated; and (5) covariation among the four first-order factors will be explained fully by their regression onto the second-order factor. One important omission in figure 2 is the presence of double-headed arrows among the first-order factors thereby indicating their intercorrelation. This is because the development and validation of the TQM measurement models progress from the first- to second-order factors, which will indicate that all covariation among the first-order factors is explained by the second-order factor.

The development and validation of the measurement model was conducted using responses from only the first wave of surveys received. Once the measurement model was developed and validated at the first- and second-order levels using responses from the first wave, a test for invariance across both the first and second waves of responses was performed. This test determined whether the factorial structure of the measurement model replicates across independent samples of the same population, which addresses the issue of non-response bias and, more importantly, cross-validation.

5.1. *Assessment of measurement model fit*

From an operational perspective, the following minimal subset was considered important for assessing the measurement properties of a construct (Peter 1981, Venkatraman 1989, Klassen 1995): unidimensionality and convergent validity, discriminant validity, criterion-related validity, nomological validity and reliability. The measurement properties of TQM were first assessed by testing the initially hypothesized full first-order TQM measurement model using confirmatory factor analysis (CFA). A strong *a priori* basis for the hypothesized four-factor TQM measurement model warranted the use of CFA rather than exploratory factor analysis. Based on theory, past research, and exploratory factor analyses, a CFA was performed since CFA is a more rigorous method for assessing unidimensionality rather than coefficient alpha, exploratory factor analysis, and item-total correlations (Gerbing and Anderson 1988, Calantone *et al.* 1996). The purpose was to ensure unidimensionality of the multiple-item constructs and to eliminate unreliable items from them.

The four factors were combined into a structural equation model for CFA with explicit estimation of the correlation between factors. The estimation of parameters in the model was determined using maximum likelihood (ML) estimation (Bollen 1989, Bentler 1992a, Joreskog and Sorbom 1993). The applications of structural equation modelling were executed using the EQS/Windows program (Bentler 1989, 1992a).

Four criteria for assessing overall fit were used: χ^2, χ^2/df, normed fit index (NFI; Bentler and Bonett 1980), non-normed fit index (NNFI), and the comparative fit index (CFI; Bentler 1990). The statistics beyond the χ^2 were introduced to reduce or eliminate the sample size dependency noted in the χ^2. In large samples, the χ^2 statistic will almost always be significant, since χ^2 is a direct function of sample size (Hartwick and Barki 1994). The χ^2 measure is especially sensitive to sample size in cases where the sample size exceeds 200 respondents (Hair *et al.* 1996). As a result, a number of researchers (Wheaton *et al.* 1977, Byrne 1994, Hartwick and Barki 1994, Sethi and King 1994, Chau 1997) use a related measure, χ^2 divided by its degrees of freedom, which should be less than 3. A small value of χ^2 relative to the degrees of freedom signifies that the observed and estimated matrices do not differ considerably.

Several researchers support the use of these additional measures of overall model goodness of fit (Segars and Grover 1993, Hartwick and Barki 1994, Chau 1997). Although these three additional indices of fit are provided in the EQS output, Bentler (1992b) recommends the NFI and CFI to be the indices of choice. Each provides a measure of complete covariation in the data, with a value greater than 0.90 indicating an acceptable fit to the data.

Although it is unlikely that the ML estimates would be affected, non-normality could lead to downwardly biased standard errors that would result in an inflated number of statistically significant parameters (Muthen and Kaplan 1985, Byrne 1994). Since raw data were used as input, EQS automatically provides univariate as well as several multivariate sample statistics. The univariate statistics represent the mean, standard deviation, skewness and kurtosis. As expected, the items were not found to be severely kurtotic. No indications of departures from normality existed (e.g. skewness $\geq$ 2, kurtosis $\geq$ 7). The multivariate statistics reported by EQS represents variants of Mardia's (1970) coefficients of multivariate kurtosis. Two reported values bear on normal theory and two on elliptical theory. Both values failed to indicate significance. Furthermore, no significant outliers were detected in the sample of 526 firms.

5.2. *Testing the hypothesized measurement model*

A summary of selected fit indices for the EQS analysis is provided in table 5. Presented with findings of $\chi^2_{(371)} = 995.468$ and CFI $= 0.823$ for the first-order TQM CFA model, further modification was needed to improve model fit to acceptable levels. The goodness-of-fit indices were much too low for a well-fitting model. When

n	269 (Group 1: First wave of responses)
Number of latent variables	4
Total number of observed variables	9
Degrees of freedom (df)	371
χ^2 statistic	995.468
p-value	0.001
χ^2/df	2.68
Bentler–Bonett normed fit index	0.747
Bentler–Bonett non-normed fit index	0.806
Comparative fit index	0.823

Table 5. Goodness-of-fit indices for the initially hypothesized first-order TQM CFA model.

a hypothesized model is tested and the fit found to be inadequate, it is customary to proceed with post-hoc model fitting to identify misspecified parameters in the model (Bollen 1989, Byrne 1994).

After eliminating items that had low-item-construct loadings or were loaded on multiple constructs, the NFI, NNFI and CFI were iteratively used to determine whether the CFA model fitted the data well. First, to make certain that a given item represented the construct underlying each factor, a loading of 0.50 was used as the minimum cut-off. Second, to avoid problems with cross-loadings, the Lagrangian Multiplier (LM) test was used to identify significant cross-loadings (i.e. a loading on more than one factor). As recommended, only one parameter was changed at every step (Joreskog and Sorbom 1993).

EQS takes a multivariate approach based on the LM test. The objective of the test was to determine if the models that better represent the data would result with certain parameters specified as free, rather than fixed, in subsequent runs. Model modifications were continued until all parameter estimates and overall fit measures were judged to be statistically and substantively satisfactory. The revised and final full first-order TQM CFA model, consisting of 13 measures was re-estimated (see table 6). The fit of the model is satisfactory based on all the criteria of χ^2, χ^2/df, NFI, NNFI and CFI.

The revised model surpasses the hypothesized model on all fit criteria, which confirms that the modifications were meaningful. There were no examples of parameters exhibiting unreasonable estimates (e.g. correlations greater than 1.0, negative variances). Furthermore, the sign and significance of the item loadings, along with an assessment of reliability indices for each factor using Cronbach's alpha also support the satisfactory fit of the models to the data (table 7).

This final model is tenable from a content and theoretical standpoint. Furthermore, the final first-order TQM CFA model satisfied all of the measurement criteria. Cronbach's coefficient α should be 0.70 or higher for narrow constructs, and 0.55 or higher for moderately broad constructs such as those defined here (Van de Ven and Ferry 1979). All α values were higher than the minimum requirements. In terms of convergent validity, all factor loadings for each individual indicator to its respective construct were positive, greater than 0.50, and highly significant ($p < 0.001$). Also, none of the LM χ^2 values were statistically significant. This means that there were no significant cross-loadings, demonstrating discriminant

n	269 (Group 1: first wave of responses)
Number of latent variables	4
Total number of observed variables	13
Degrees of freedom (df)	59
χ^2 statistic	155.684
p-value	0.001
χ^2/df	2.64
Bentler–Bonett normed fit index	0.907
Bentler–Bonett non-normed fit index	0.920
Comparative fit index	0.939

† All of the standardized residuals were below 0.218.
‡ Distribution of standardized residuals was symmetric and centred on zero.

Table 6. Goodness-of-fit indices for the final first-order TQM CFA model.

	Measurement equation	Standard error	Test statistic
TQM strategic systems (F1):			
Cronbach's $\alpha = 0.643$	V4 = 1.294†F1 + 1.000 E4	0.121	10.705
	V6 = 1.381†F1 + 1.000 E6	0.118	11.660
Standardized solution:	V4 = 0.659†F1 + 0.752 E4		
	V6 = 0.719†F1 + 0.695 E6		
Items dropped: V1, V2, V3, V5, V7			
TQM operational systems (F2):			
Cronbach's $\alpha = 0.865$	V9 = 1.748†F2 + 1.000 E9	0.108	16.167
	V10 = 1.481†F2 + 1.000 E10	0.090	16.516
Standardized solution:	V9 = 0.873†F2 + 0.487 E9		
	V10 = 0.888†F2 + 0.460 E10		
Items dropped: V8, V11, V12, V13, V14			
TQM Information Systems (F3):			
Cronbach's $\alpha = 0.728$	V19 = 1.047†F3 + 1.000 E19	0.094	11.115
	V20 = 1.529†F3 + 1.000 E20	0.117	13.026
Standardized solution:	V19 = 0.702†F3 + 0.712 E19		
	V20 = 0.834†F3 + 0.553 E20		
Items dropped: V15, V16, V17, V18			
TQM results (F4):			
Cronbach's $\alpha = 0.872$	V21 = 2.090†F4 + 1.000 E21	0.127	16.493
	V22 = 2.411†F4 + 1.000 E22	0.126	19.096
	V23 = 1.917†F4 + 1.000 E23	0.115	16.700
	V24 = 0.858†F4 + 1.000 E24	0.091	9.419
	V26 = 0.918†F4 + 1.000 E26	0.091	10.146
	V29 = 0.800†F4 + 1.000 E29	0.088	9.100
Standardized solution:	V21 = 0.836†F4 + 0.548 E21		
	V22 = 0.917†F4 + 0.399 E22		
	V23 = 0.843†F4 + 0.537 E23		
	V24 = 0.549†F4 + 0.836 E24		
	V25 = 0.557†F4 + 0.831 E25		
	V26 = 0.584†F4 + 0.811 E26		
	V29 = 0.534†F4 + 0.846 E29		
Items dropped: V27, V28			

† All factor loadings are significant at $p < 0.001$.

Table 7. The final first-order TQM CFA model.

validity. All of the scales have statistically significant and positive correlations with the primary outcome factor of TQM Results (Factor 4). Thus, criterion-related validity is supported for all the scales. Finally, all of the inter-factor correlations were positive and significant.

5.3. *Relationships among the first-order factors*

Nomological validity was assessed from the final measurement model using the inter-factor correlations (Bagozzi 1981). All correlations were statistically significant and positive, and some of the correlations were very large (see table 8). The large correlations among some of the factors was not surprising since it was hypothesized

Factor 1: TQM strategic systems Factor 2: TQM operational systems	0.794[a]	Factor 1: TQM strategic systems Factor 3: TQM Information systems	0.766[a]
Factor 1: TQM strategic systems Factor 4: TQM results	0.134[c]	Factor 2: TQM operational systems Factor 3: TQM information systems	0.527[a]
Factor 2: TQM operational systems Factor 4: TQM results	0.103[c]	Factor 3: TQM information systems Factor 4: TQM results	0.157[b]

[a] Correlations are statistically significant ($p < 0.001$).
[b] Correlations are statistically significant ($p < 0.05$).
[c] Correlations are statistically significant ($p < 0.10$).

Table 8. Correlations among TQM constructs.

a priori that these four underlying first-order factors are associated with a higher-order factor. The lack of any negative correlations among the factors indicates that a high value on one factor does not preclude a high value on another factor. In other words, the factors complement one another.

5.4. *Second-order CFA model*

In the previous TQM factor analytic model, there were four factors that operated as independent variables. Each could be considered to be one level or one unidirectional arrow away from the observed variables. These were subsequently termed first-order factors. However, theory argues for a higher level factor that is considered accountable for the lower-order factors. Let us examine the representation of this model in figure 2.

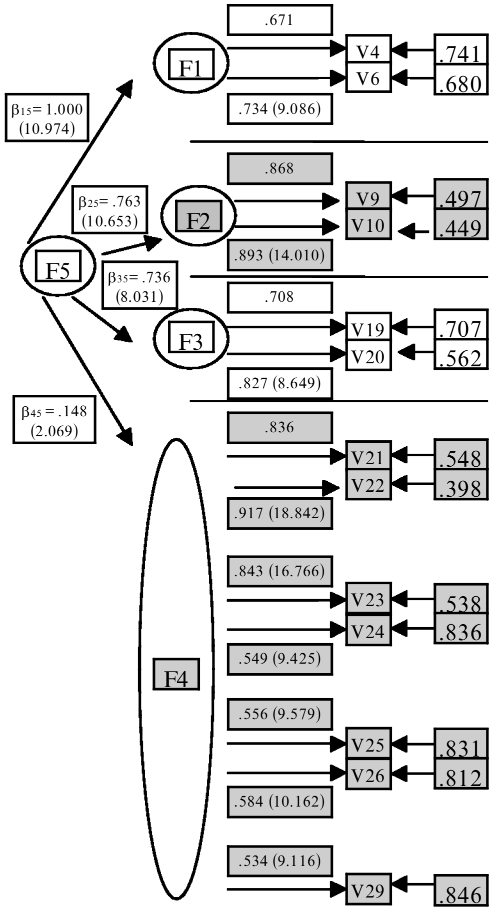
This model essentially has the same first-order factor structure. However, the higher-order factor, TQM (F5), is hypothesized as accounting for or explaining all variance and covariance related to the first-order factors. It is important to note that TQM does not have its own set of measured indicators. Rather, it is linked indirectly to those measuring the lower-order factors. The first-order factors now operate as dependent variables, which means that their variances and covariances are no longer estimated parameters in the model. Such variation and covariation is presumed to be accounted for by the higher-order factor.

The overall model statistics indicate that the fit of the second-order model is as good as that of the first-order model (see table 9). This is as it should be, given that the second-order model merely specifies a higher-order factor to account for the correlations among the lower-order factors, rather than of these factors among

n	269 (Group 1: First wave of responses)
χ^2 statistic	158.105
Degrees of freedom (df)	61
p-value	0.001
χ^2/df	2.59
Normed fit index	0.905
Nonnormed fit index	0.922
Comparative fit index	0.939

† All of the standardized residuals were below 0.219.
‡ Distribution of standardized residuals was symmetric and centred on zero.

Table 9. Goodness-of-fit indices for the second-order TQM CFA model.



* Each first-order factor has a disturbance term associated with it. D1=.375, D2=.599, D3=.787, and D4=.965.

( ) t-values in parentheses

F5 = Factor 5 = TQM

F1 = Factor 1 = TQM Strategic Systems

V4: Adequate resources are provided to carry out quality improvements within your plant (2.2)
V6: Key factors for building and maintaining customer relationships are identified and used by your plant (3.1)

F2 = Factor 2 = TQM Operational Systems

V9: An adequate amount of training in quality awareness is provided to hourly/direct labor employees within your plant (5.2)
V10: An adequate amount of training in quality awareness is provided to managers and supervisors within your plant (5.2)

F3 = Factor 3 = TQM Information Systems

V19: Procedures have been developed for monitoring key indicators of plant performance (4.2)
V20: Procedures have been developed for monitoring key indicators of customer satisfaction (4.3)

F4 = Factor 4 = TQM Results

V21: After-sales customer complaints (7.1)
V22: Customer rejection of our products (e.g., manufacturing defects) (7.1)
V23: Defect rates/cost (7.2)
V24: Employee absenteeism (7.3)
V25: Cost of quality (e.g., inspection and testing) (7.2)
V26: Employee grievances (7.3)
V29: Total cost of purchased parts (7.4)

Figure 3. The final full TQM CFA measurement model.

themselves (as is the case with the first-order structure). The results given in figure 3 (the final full second-order TQM CFA measurement model) show that the loadings of all four first-order factors on the second-order factor are positive and significant.

5.5. *Cross-validation and non-response bias*

This section addresses the issue of non-response bias and, more importantly, cross-validation. In other words, is the TQM CFA model equivalent across the first wave of responses (Group 1; $n = 269$) and the second wave of responses (Group 2; $n = 257$).

In testing for invariance across both groups, the factor loadings and covariances for each measurement model were tested for their equivalence across groups. The results of this test are presented in table 10. Shown first are the goodness of fit statistics relative to the entire model, which comprises the two baseline models with equality constraints between them. As indicated by a CFI of 0.943 for the TQM CFA model, the multigroup models represent excellent fit to the data. The hypothesized equality of the specified factor loadings and covariances held. Given these findings, all measures of TQM are operating in the same way for both groups.

Methods of non-response bias are often based on the assumption that subjects who answered later are more like non-respondents. The most common type of extrapolation is carried over successive waves of a questionnaire. A wave refers to the response generated by a stimulus, which was a reminder postcard in this study (Armstrong and Overton 1977). Persons who responded in the second wave are assumed to have responded because of the stimulus. The evidence of a lack of non-response bias exists since the TQM CFA model proved to be equivalent across the first and second waves of responses.

5.6. *Summary*

In summary, all of the causal paths specified in the hypothesized model were found to be positive and statistically significant. These paths reflected the impact of: (1) TQM (F5) on TQM Strategic Systems (F1), TQM Operational Systems (F2), TQM Information Systems (F3), and TQM Results (F4). In other words, all of the structural paths from TQM to: TQM Strategic Systems ($\beta_{15} = 1.000$, $t = 10.974$), TQM Operational Systems ($\beta_{25} = 0.763$, $t = 10.653$), TQM Information Systems ($\beta_{35} = 0.736$, $t = 8.031$), and TQM Results ($\beta_{45} = 0.148$, $t = 2.069$) were positive and significant as hypothesized. Paths not specified *a priori* did not prove to be essential components of the causal structure; therefore, they were not added to the model. No paths were found to be misspecified. Also, none of the hypothesized paths

Degrees of freedom (df)	137
χ^2 statistic	313.840
p-value	0.001
χ^2/df	2.29
Bentler–Bonett normed fit index	0.903
Bentler–Bonett non-normed fit index	0.935
Comparative fit Iidex	0.943

† All of the standardized residuals were below 0.238.
‡ Distribution of standardized residuals was symmetric and centered on zero

Table 10. Goodness-of-fit Indices for the TQM multigroup model.

were non-significant and were subsequently left in the model. The next section builds on these findings by discussing the results and using them to act as a basis for future research.

6. Discussion

At the beginning of this study, we started by noting that our interest was in examining two sets of relationships, both involving the MBNQA. The first involved the relationship between the measures developed by the MBNQA and their fit with the categories to which they were assigned by the MBNQA programme. The second involved the extent to which the MBNQA really did capture this unobserved but critical construct known as Total Quality Management. The analyses presented in the previous section positions us to examine and discuss the various results, as they pertain to these two sets of relationships. We begin our discussion by examining the first set of relationships.

6.1. *Assessing the fit between the measures and the constructs*

In the first set of relationships, we were interested in examining the fit between the measures and the constructs that these were intended to capture. This analysis can be answered at two levels. At the first, we focus on determining the extent to which there is a unique and theory-driven relationship between measures and constructs. As noted previously, this study had operationalized the MBNQA in terms of four major constructs or categories: (1) TQM Strategic Systems; (2) TQM Operations Systems; (3) TQM Informational Systems; and (4) TQM Results. For each construct, a number of metrics or measurement variables were generated, consistent with the MBNQA criteria. The results show that the measures do indeed load on the appropriate constructs. More important, these measures load on no other construct. As a result, we can say that the metrics generated by the MBNQA do indeed measure the underlying unobserved behaviour that they are supposed to measure. To this end, there is nothing really unique about this first set of findings.

However, what is more interesting is the second set of findings involving the number of measurement or manifest variables that were ultimately identified as best loading on each construct. Initially, for each for the four constructs, we introduced four or more measures. During the analysis, we noted that some of the measures were dropped from further consideration.

TQM Strategic Systems. For this system (F1), only two of the original seven measures were retained in the final TQM CFA measurement model. These two measures were: (1) 'Adequate resources are provided to carry out quality improvements within your plant' (V4); and (2) 'Key factors for building and maintaining customer relationships are identified and used by your plant' (V6). These two measures include issues as they pertain to strategic planning (V4) and customer/market focus (V6). The other four measures were eliminated as not providing as strong a relationship as these two remaining measures. This study provides empirical support that a TQM-based system must encompass an organization's knowledge of its customers, responsiveness and ability to meet customer requirements and expectations.

TQM Operational Systems. For the TQM Operational Systems construct (F2), we found that only two of the original seven measures associated with this construct were retained in the final TQM CFA measurement model. These two were: (1) 'An adequate amount of training in quality awareness is provided to hourly/direct labor employees within your plant' (V9); and (2) 'An adequate amount of training in quality awareness is provided to managers and supervisors within your plant' (V10). Both of these measures deal with human resource development. The other measures associated with process management were dropped. This last finding is interesting in that it was not expected. Given the heavy emphasis in the TQM literature on process awareness and the need for better or more effective process management, we had expected some degree of process management to be present in the retained set of measures.

One possible reason for this observed result can be found in the need for training, as it pertains to human resources. Human resource development is defined as the success of an organization's efforts to realize the full potential of its workforce in implementing TQM (MBNQA 1997). Many authors have used case studies and anecdotal examples to contend that TQM can best be achieved when there is a high level of commitment and involvement from people (Cole 1980, Juran 1981a, 1981b, Oliver 1988, Harber *et al.* 1993, Longnecker and Scazzero 1993). Several other articles identify training and development at all levels as the single best strategy to improve quality (Ebrahimpour 1985, Lee and Ebrahimpour 1985). Training can shape employees' positive attitudes about their company's commitment to TQM. This study provides empirical support that TQM demands that human resource management, in the form of training and development, assumes strategic roles.

TQM Information Systems. This construct (F3) retained two of the original six measures in the final TQM CFA measurement model. These were: (1) 'Procedures have been developed for monitoring key indicators of plant performance' (V19); and (2) 'Procedures have been developed for monitoring key indicators of customer satisfaction' (V20). These two measures refer to the quality of the metrics, both internally (within the plant) and externally (at the customer). These measures point the need for managers to develop and have a balanced view of performance and to have the information necessary to maintain this view. Internal performance must be linked and balanced against external performance. It is not enough to have a well-running plant, if that efficiency is not translated into increased customer satisfaction.

Furthermore, the importance of timely, reliable, and adequate information in the development and implementation of TQM has been noted by several researchers (Garvin 1983, Babbar 1992). Fundamental to TQM is collecting relevant information from all phases of an organization's operations and using it to monitor and improve quality. The importance of the expanded role of information technology and systems in integrating information from inside and outside the company (e.g. customers) has also been identified in the literature (Willborn 1986, Riehl 1988). This study supports that information, and analysis can help reinforce the implementation of TQM practices.

TQM results. Of the four constructs assessed within this study, this construct retained the largest number of measures in the final TQM CFA model. Of the nine original measures, seven were kept: (1) after-sales customer complaints (V21); (2) customer rejection of our products (e.g. manufacturing defects) (V22); (3) defect rates/cost (V23); (4) employee absenteeism (V24); (5) cost of quality (e.g. inspection and testing) (V25); (6) employee grievances (V26); and (7) total cost of purchased parts (V29). This more complex relationship between the metrics and the underlying latent variable can be explained as follows.

TQM requires that companies monitor and improve their quality performance based on objective measures of quality results. Ensuring customer-driven quality requires measurement of quality results. Researchers have suggested that superior quality performance is the result of understanding the factors that determine quality performance (Reddy and Berger 1983, Fortuin 1988). Other studies have similarly concluded that firms need to focus on improving the quality of every work process as measured by the needs of internal and external customers (Cole 1990, Fisher 1992). This study provides empirical support that monitoring performance is integral to TQM. What this means is that measures should be identified and used to determine if the system is delivering the desired results.

Simplifying the underlying model. The results presented in the preceding section are interesting when taken construct by construct. However, when taken as a whole, another interesting finding emerges from the empirical data. Originally, the MBNQA framework had been operationalized as consisting of four constructs or latent variables. These four constructs were measured using 30 measurements or manifest variables (TQM Strategic Systems – 7; TQM Operational Systems – 7; TQM Information Systems – 6; and TQM Results – 9). Of these 30 measures, only 13 measures were ultimately retained in the TQM CFA model (TQM Strategic Systems – 2; TQM Operational Systems – 2; TQM Information Systems – 2; TQM Results – 7). This is an interesting result because the ultimate TQM model can be viewed as a simpler, less complex model.

Simplification is an important goal of any research study. Simpler models are easier to study and explain. They are also more powerful. As noted within the well-known theorem of 'Occam's Razor,' parsimony is a desired trait of every model. What these empirical results point to is a more parsimonious model of the relationship between the measures and the various constraints.

6.2. *Assessing the fit between the MBNQA framework and TQM*

From the outset, this study was interested in the extent to which the various categories and measures contained in the MBNQA really do capture this higher-order but unobserved construct commonly referred to as TQM. The empirical evidence strongly supports this relationship between the MBNQA framework and TQM. This relationship is supported in terms of the relationships between the four constructs of the MBNQA and the second-order construct of TQM. We can see these linkages in the relationships between each of the four constructs and TQM (see figure 3).

The structural path from TQM (F5) to TQM Strategic Systems (F1) was positive and significant as hypothesized ($\beta_{15} = 1.000$, $t = 10.974$). Similarly, the structural path from TQM (F5) to TQM Operational Systems (F2) was positive and significant as hypothesized ($\beta_{25} = 0.763$, $t = 10.653$). For the third relationship between TQM and TQM Information System, we observed a positive and significant relationship, as hypothesized ($\beta_{35} = 0.736$, $t = 8.031$). Finally, the structural path from TQM and TQM Results (F4) was positive and significant, as hypothesized ($\beta_{45} = 0.148$, $t = 2.069$). In short, these findings indicate that by following the MBNQA framework, you will achieve the implementation of this system known as TQM. What makes this finding so strong is that it is not based on simply anecdotal or case data, nor is it based on the experiences of 'leading edge' practitioners. Rather, what we have here are findings based on a cross-section of experiences of firms within a specific industrial sector.

More important and of greater interest, the findings presented in this study draw a unique picture of the TQM *process*. To attain TQM status, a firm must first attain certain performance requirements, as captured by the various measures. As these performance traits are attained, the firm is able to achieve the implementation of certain critical TQM subsystems (i.e. TQM Strategic Systems, TQM Operational Systems, TQM Information Systems, and TQM Results). Each subsystem is required to be in place before the firm can hope to claim that it has a complete TQM system present. In short, we can argue that the MBNQA is more than simply a framework, it is a process model.

7. Future research

Before the findings reported in this paper can be meaningfully generalized, more work is required. These efforts must begin with the TQM measures presented in this paper. The TQM measures were an attempt at operationalizing a higher-order concept and must therefore be replicated and refined. Alternative measures must be formulated and compared with the results of this study to clarify the theoretical foundations of this higher-order construct. Also, the structural equation model developed requires validation on another data set. This is particularly critical because the model was developed as well as tested using the same data set. Additionally, given the perceptual nature of the data used to reflect the theoretical constructs, it is important to recognize the problems associated with the 'key informant' approach (Phillips 1981, Huber and Power 1985, Sethi and King 1994). Thus, a logical extension would be to use multiple informants to verify perceptions regarding the features and impact of a TQM-based system.

Furthermore, since over 80% of the people employed in the US are employed in services, it would not make sense to confine TQM research to the manufacturing sector of the economy, especially since the service sector also has to deal with quality issues. To do this, future research will need to recognize how service businesses are similar and different from manufacturing businesses in terms of quality management. This translation of manufacturing TQM-based systems to service organizations cannot be presumed and would have to be carefully thought out.

The exact structure of TQM must also be explored in more detail. An explicit consideration of the causal relationships among the first-order factors will be necessary. To date, there has been very little research aimed at providing empirical evidence of the linkages between implementation of TQM criteria. An explicit model is required in establishing the relationships that exist between the various criteria (e.g. Strategic Systems, Operational Systems, Information Systems, and Results). This type of study should provide a framework that enables researchers and managers to gain improved insight into TQM by providing a generic framework for implementation. An empirical test of the linkages between the TQM factors can help to provide a roadmap for firms seeking to progress towards total quality cultures.

8. Conclusions

Given that reliable and valid measures were needed for TQM constructs, which are in turn important for theory building, this study has developed and validated an appropriate measurement instrument for TQM. To overcome the weaknesses of classical approaches to instrument validation (e.g. multitrait/multimethod and exploratory factor analysis), a newer approach, namely structural equation modelling, was used in this study. This approach provided a more rigorous validation of instruments for unobserved constructs and testing of the research model than the classical approach (e.g. exploratory factor analysis). The ability of structural equation modelling techniques to assess relationships comprehensively has provided a transition from exploratory to confirmatory analysis (Bollen 1989).

This study has illustrated how this approach can be used to test instrument validation and modify instruments for more rigorous psychometric properties. This study has moved from anecdotes and case studies to a testable model and specific research hypotheses, linking the theoretical concept of TQM to empirical indicants. This study has contributed to TQM theory building by identifying the constructs associated with TQM, developing scales for measuring these constructs, and empirically validating the scales. More important, it has shown empirically that the MBNQA framework does capture the concept of Total Quality Management. This should be encouraging news for all those organizations and firms that have based their quality improvement efforts (i.e. programmes, self-assessment projects, and awards) on this programme.

Appendix A. Operational measures

Appendix A (Operational measures) deals with generating items that represent manifestations of the four factors/latent variables associated with TQM. Multi-item scales for each factor serve as parsimonious representations of unidimensional factors, corresponding in similarity to each of the four factors associated with the MBNQA framework (e.g. TQM Strategic Systems, TQM Operational Systems, TQM Information Systems, TQM Results). Definitions of these factors and the selection of items were developed directly from the MBNQA criteria. Each manifestation is measured with an item in a scale. The corresponding MBNQA criterion is shown in parentheses following each measure. Factor items were constructed through confirmatory factor analysis (using loadings of 0.50 as a minimum value for inclusion of an item in a factor) and reliability analyses (Cronbach's coefficient alpha). Items that did not have significant λs or that did not demonstrate sufficient discriminant validity were, by definition, not representative of the factor, and were subsequently removed from the model (as indicated by an asterisk in the descriptions that follow).

The focus of the measures is on real decisions made by plant managers and the ultimate effects of those decisions, as viewed by them, irrespective of the theoretical correctness or incorrectness of those decisions. Consequently, all data will be based on managers' perceptions. While one could argue that focusing on managerial perceptions may miss the truth, this approach provides a balance by focusing on the real world approach of making decisions on educated perceptions. The following sections list the factors and the scales used in the study.

Factor 1 (F1): TQM Strategic Systems

The TQM Strategic Systems factor includes and examines: senior executives' personal leadership and involvement in creating and sustaining a customer focus and clear and visible quality values; how the values and expectations are integrated into the company's management system; the company's planning process and how all key quality requirements are integrated into overall business planning; the company's short- and long-term plans and how quality and performance requirements are deployed to all work units; how the company determines requirements and expectations of customers and markets; and how the company enhances relationships with customers and determines their satisfaction (MBQNA 1997). The questions for the TQM Strategic Systems factor will be measured using an 11-point bipolar scale (e.g. 0 = strongly disagree, 10 = strongly agree).

V1—Quality goals are clearly communicated to all plant personnel (1.1).*
V2—Quality is emphasized through a well-defined set of quality policies and procedures within your plant (1.1).*
V3—Customer quality requirements are used to establish a plant level quality strategy (2.1).*
V4—Adequate resources are provided to carry out quality improvements within your plant (2.2).
V5—Plant and/or other company personnel actively interacts with customers to set reliability, responsiveness, and other standards for the plant (3.1).*
V6—Key factors for building and maintaining customer relationships are identified and used by your plant (3.1).
V7—Formal and informal customer complaints are evaluated by your plant (3.2).*

Factor 2 (F2): TQM Operational Systems

The TQM Operational Systems factor includes and examines: how the workforce is enabled to develop and utilize its full potential, aligned with the company's objectives; the company's efforts to build and maintain an environment conducive to full participation, and personal and organizational growth; the key aspects of process management, including customer-focused design, product and service delivery processes, support services and supply management involving all work units, including research and development; and how key processes are designed, effectively managed, and improved to achieve higher performance (MBQNA 1997). The questions for the TQM Operational Systems factor will be measured using an 11-point bipolar scale (e.g. 0 = strongly disagree, 10 = strongly agree).

V8—Human resources management within your plant is affected by quality plans (5.1).*
V9—An adequate amount of training in quality awareness is provided to hourly/direct labor employees within your plant (5.2).
V10—An adequate amount of training in quality awareness is provided to managers and supervisors within your plant (5.2).
V11—The work environment within your plant is conducive to employee well-being and growth (5.3).*
V12—The manufacturability of products built within your plant is considered during the product design process (6.1).*
V13—Easy access for customers seeking information or assistance and/or comment and complaint is provided (6.2).*
V14—Suppliers' facilities are visited regularly by plant and/or other company personnel (6.3).*

Factor 3 (F3): TQM Information Systems

The TQM Information Systems factor includes and examines: the scope, validity, analysis, management, and use of data and information to drive quality excellence and improve competitive performance; and the adequacy of the company's data, information, and analysis system to support improvement of the company's customer focus, products, services, and internal operations (MBNQA 1997). The questions for the TQM Information Systems factor will be measured using an 11-point bipolar scale (e.g. 0 = strongly disagree, 10 = strongly agree).

V15—Quality data within the plant are made visible – displayed at work stations (4.1).*
V16—Quality data within the plant are provided in a timely manner (4.1).*
V17—Quality data are made available to all employees within your plant (4.1).*
V18—Benchmark data are used to improve quality practices within your plant (4.2).*
V19—Procedures have been developed for monitoring key indicators of plant performance (4.2).
V20—Procedures have been developed for monitoring key indicators of customer satisfaction (4.3).

Factor 4 (F4): TQM Results

The TQM Results factor includes and examines: the company's performance and improvement in key business areas—product and service quality, productivity and operational effectiveness, and supply quality (MBNQA 1997). The questions for the TQM Results factor will be introduced by: 'Please estimate the magnitude of change experienced in each quality measure over the last three years'. These measures will then be converted into an 11-point bipolar scale ranging from 0 to 10, where '0' = +100% or >, '1' = +80%, '2' = +60%, '3' = +40%, '4' = +20%, '5' = no change, '6' = −20%, '7' = −40%, '8' = −60%, '9' = −80%, '10' = −100% or >.

V21—After-sales customer complaints (7.1).
V22—Customer rejection of our products (e.g. manufacturing defects) (7.1).
V23—Defect rates/cost (7.2).
V24—Employee absenteeism (7.3).
V25—Cost of quality (e.g. inspection and testing) (7.2).
V26—Employee grievances (7.3).
V27—Employee turnover (7.3).*
V28—On-time delivery of purchased parts (7.4) (reversed scale).*
V29—Total cost of purchased parts (7.4).

References

Ahire, S. L., Golhar, D. Y. and Waller, M. A., 1996, Development and validation of TQM implementation constructs. *Decision Sciences*, **27**, 23–55.

Ahire, S. L., Landeros, R. and Golhar, D. Y., 1995, Total Quality Management: a literature review and an agenda for future research. *Production and Operations Management*, **4**, 277–306.

Armstrong, J. S. and Overton, T. S., 1977, Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, **14**, 392–402.

Babbar, S., 1992, A dynamic model for continuous improvement in the management of service quality. *International Journal of Operations and Production Management*, **12**, 38–48.

Bagozzi, R. P., 1981, An examination of the validity of two models of attitude. *Multivariate Behavioral Research*, **18**, 133–145.

Bentler, P. M., 1989, *EQS: Structural Equations Program Manual* (Los Angeles: BMDP Statistical Software); 1990, Comparative fit indexes in structural models. *Psychological Bulletin*, **107**, 238–246; 1992a, *EQS: Structural Equations Program Manual* (Los Angeles: BMDP Statistical Software); 1992b, On the fit of models to covariances and methodology to the bulletin. *Psychological Bulletin*, **112**, 400–404.

BENTLER, P. M. and BONETT, D. G., 1980, Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, **88**, 588–606.

BLACK, S. A. and PORTER, L. J., 1996, Identification of the critical factors of TQM. *Decision Sciences Journal*, **27**, 363–379.

BOLLEN, K. A., 1989, *Structural Equations With Latent Variables* (New York: Wiley).

BYRNE, B. M., 1994, *Structural Equation Modeling With EQS And EQS/Windows* (California: Sage Publications).

CALANTONE, R. J., SCHMIDT, J. B. and SONG, M. X., 1996, Controllable factors of new product success: a cross-national comparison. *Marketing Science*, **15**, 341–358.

CAPON, N., KAYE, M. and WOOD, M., 1994, Measuring the success of a TQM programme. *International Journal of Quality and Reliability Management*, **12**, 8–22.

CHAU, Y. K., 1997, Re-examining a model for evaluating information center success using a structural equation modeling approach. *Decision Sciences*, **28**, 309–334.

COLE, R. E., 1980, Learning from the Japanese: Prospects and pitfalls. *Management Review*, **69**, 22–28; 1990, U.S. quality improvement in the auto industry: close but no cigar. *California Management Review*, **32**, 71–85.

CURKOVIC, S., SWARTZ, S. and VICKERY, S. K., 1997, An empirical analysis of quality performance in the automotive supply industry. *Proceedings of the 28th Annual Meeting of the Midwest Decision Sciences Institute*, Indianapolis, Indiana.

DEAN, J. W. and BOWEN, D. E., 1994, Management theory and total quality: improving research and practice through theory development. *Academy of Management Journal*, **19**, 392–418.

DEMING, W. E., 1981, Improvement of quality and productivity through action by management. *National Productivity Review*, **1**, 12–22; 1982, *Quality, Productivity, and Competitive Position* (Cambridge, MA: MIT Center for Advanced Engineering); 1986, *Out of the Crisis* (Cambridge: Massachusetts Institute of Technology, Center for Advanced Engineering Study).

DILLMAN, D., 1978, *Mail and Telephone Surveys* (New York: Wiley).

EBRAHIMPOUR, M., 1985, An examination of quality management in Japan: implications for management in the United States. *Journal of Operations Management*, **5**, 419–431.

EBRAHIMPOUR, M. and WITHERS, B. E., 1992, Employee involvement in quality improvement: a comparison of American and Japanese manufacturing firms operating in the U.S. *IEEE Transactions on Engineering Management*, **39**, 142–148.

ETTORRE, B., 1996, Is the Baldrige still meaningful? *Management Review*, **85**, 28.

EVANS, J. R., 1992, *A Report of the Total Quality Leadership Steering Committee and Working Councils* (Milwaukee: Proctor and Gamble); 1997, The 1997 Baldrige: making quality an adjective. *POM Chronicle*, **7**, 4.

FISHER, T. J., 1992, The impact of quality management on productivity. *International Journal of Quality and Reliability Management*, **9**, 44–52.

FLYNN, B. B., SCHROEDER, R. G. and SAKAKIBARA, S., 1994, A framework for quality management research and an associated instrument. *Journal of Operations Management*, **11**, 339–366.

FORTUIN, L., 1988, Performance indicators—why, where, and how? *European Journal of Operations Research*, **34**, 1–9.

GARVIN, D. A., 1988, Quality on the line. *Harvard Business Review*, **61**, 64–75.

GERBING, D. W. and ANDERSON, J. C., 1988, An updated paradigm for scale development incorporating unidimensionality and its assessment. Journal of Marketing Research, 25, 186–192.

GRIFFIN, R., 1988, Consequences of quality circles in an industrial setting: a longitudinal assessment. *Academy of Management Journal*, **31**, 338–358.

HAIR, J. F., ANDERSON, R. E., TATHAM, R. L. and BLACK, W. C., 1996, *Multivariate Data Analysis* (Englewood Cliffs, New Jersey: Prentice Hall).

HARBER, D., BURGESS, K. and BARCLAY, D., 1993, Total Quality Management as a cultural intervention: an empirical study. *International Journal of Quality and Reliability Management*, **10**, 28–46.

HARTWICK, J. and BARKI, H., 1994, Explaining the role of user participation in information system use. *Management Science*, **40**, 440–465.

HOCKMAN, K., 1992, Does the Baldrige award really work? *Harvard Business Review*, Jan–Feb, 126–148.

HUBER, G. P. and POWER, D. J., 1985, Retrospective reports of strategic-level managers: guidelines for increasing their accuracy. *Strategic Management Journal*, **6**, 171–180.

JORESKOG, K. G. and SORBOM, D., 1993, *New Features in LISREL 8* (Chicago: Scientific Software).

JURAN, J. M., 1981a, Product quality—a prescription for the West: Part I, *Management Review*, **70**, 8–14; 1981b, Product quality—a prescription for the West: Part II. *Management Review*, **70**, 57–61.

JURAN, J. M. and GRYNA, F. M., 1980, *Quality Planning and Analysis* New York: McGraw-Hill).

KLASSEN, R. D., 1995, The implications of environmental management strategy for manufacturing performance. Doctoral Dissertation, University of North Carolina at Chapel Hill.

LEE, S. M. and EBRAHIMPOUR, M., 1985, An analysis of Japanese quality control systems: Implications for American manufacturing firms. *SAM Advanced Management Journal*, **50**, 24–31.

LOGOTHETIS, N., 1992, *Managing for Total Quality: From Deming to Taguchi and SPC* (Hertfordshire, UK: Prentice-Hall International), 5.

LONGNECKER, C. O. and SCAZZERO, J. A., 1993, Total Quality Management from theory to practice: a case study. *International Journal of Quality and Reliability Management*, **10**, 24–31.

MARDIA, K. V., 1970, Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **57**, 519–530.

MBNQA, 1997, *Malcolm Baldrige National Quality Award 1997 Award Criteria* (Milwaukee, WI: The Malcolm Baldrige National Quality Award Consortium, Inc).

MELNYK, S. A. and DENZLER, D. R., 1996, *Operations Management: A Value-Driven Approach* (Chicago, Illinois: Irwin).

MUTHEN, B. and KAPLAN, D., 1985, A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, **38**, 171–189.

OLIVER, N., 1988, Employee commitment and Total Quality Control. *International Journal of Quality and Reliability Management*, **7**, 21–29.

PETER, P. J., 1981, Construct validity: a review of basic issues and marketing practices. *Journal of Marketing Research*, **18**, 133–145.

PHILLIPS, L. W., 1981, Assessing measurement error in key informant reports: a methodological note on organizational analysis in marketing. *Journal of Marketing Research*, **18**, 395–415.

POWELL, T. C., 1995, Total Quality Management as a competitive advantage: a review and empirical study. *Strategic Management Journal*, **16**, 15–37.

REDDY, J. and BERGER, A., 1983, Three essentials of product quality. *Harvard Business Review*, **61**, 153–159.

RIEHL, J. R., 1988, Planning for Total Quality: the information technology component. *SAM Advanced Management Journal*, **53**, 13–19.

SARAPH, V., BENSON, P. G. and SCHROEDER, R. G., 1989, An instrument for measuring the critical factors of quality management. *Decision Sciences*, **20**, 810–829.

SCHONBERGER, R. J., 1983, Work improvement programs: quality control circles compared with traditional western approaches. *International Journal of Operations and Production Management*, **3**, 18–32.

SEGARS, A. H. and GROVER, V., 1993, Re-examining perceived ease of use and usefulness: a confirmatory factor analysis, *MIS Quarterly*, **17**, 517–525.

SETHI, V. and KING, W. R., 1994, Development of measures to assess the extent to which an information technology application provides competitive advantage. *Management Science*, **40**, 1601–1627.

STEEPLES, M. M., 1993, *The Corporate Guide to the Malcolm Baldrige National Quality Award* (Homewood, IL: Business One-Irwin).

SWAMIDASS, P. M. and NEWELL, W. T., 1987, Manufacturing strategy, environmental uncertainty and performance: a path analytic model. *Management Science*, **33**, 509–524.

VAN DE VEN, A. and FERRY, D., 1979, *Measuring and Assessing Organizations* (New York: Wiley).

VENKATRAMAN, N., 1989, Strategic orientation of business enterprises: the construct, dimensionality, and measurement. *Management Science*, **35**, 942–962.

VICKERY, S. K., DROGE, C. and MARKLAND, R. E., 1993, Production competence and business strategy: do they affect business performance? *Decision Sciences*, **24**, 435–455.

WHEATON, B., MUTHEN, B., ALVIN, D. F. and SUMMERS, G. F., 1977, Assessing reliability and stability in panel models. In *Sociological Methodology* edited by D. R. Herse (San Francisco, CA: Jossey-Bass), 84–136.

WHYBARK, D. C. and VASTAG, G., 1993, *Global Manufacturing Practices* (New York: Elsevier).

WILLBORN, W., 1986, Quality assurance and computer assisted production. *International Journal of Quality and Reliability Management*, **2**, 15–24.

WOMACK, J. P., JONES, D. T. and ROOS, D., 1990, *The Machine that Changed the World* (New York: Rawson Associates).