

An empirical study on the relationships within the categories of the EFQM model¹

*Iñaki Heras-Saizarbitoria**

Associated Professor
Department of Management
University of the Basque Country

iheras@ehu.es

* Corresponding author

Address: Escuela Universitaria de Estudios Empresariales, University of the Basque Country, Plaza Oñati 1, 20018, San Sebastian, Spain. Email: iheras@ehu.es; Telf.: +003443018354; Fax: +003443018361.

Frederic Marimon

Associated Professor
Department of Management
Faculty of Economics and Social Sciences
Universitat Internacional de Catalunya

fmarimon@cir.uic.es

Address: Faculty of Economics and Social Sciences, Universitat Internacional de Catalunya, Immaculada, 22, 08017 Barcelona. Spain. Email: fmarimon@cir.uic.es; Telf.: +0034932541800; Fax: +0034932541850.

Martí Casadesús

Professor
Department of Management & Product Design
Polytechnic School
University of Girona

marti.casadesus@udg.edu

Address: Department of Management & Product Design, University of Girona, Av. Lluís Santaló, s/n, 17071 Girona, Spain. Email: marti.casadesus@udg.edu; Telf.: +0034972418286.

¹ This article has been written as part of a research project entitled “Mejora de la satisfacción de los clientes en las empresas españolas mediante modelos de gestión de la calidad total” (ECO2009-12754-C02-02) financed by the Ministry of Science and Innovation of Spain.

An empirical study on the relationships within the categories of the EFQM model

Abstract

The relationships within the categories of the EFQM self-assessment model are analysed in this article, based on 242 independent assessments carried out in the European region with the highest density of EFQM awards (the Basque Autonomous Community, in Spain). The main finding of the article is that the relationships within the categories of the EFQM are robust, despite the fact that there exist relationships among some of its enablers and results that fail to reach a suitable level of validity. These findings coincide with the conclusions drawn from studies carried out previously for the Malcom Baldrige model. The conclusions drawn in the article may be of interest both for academic and professional spheres of activity.

Keywords: Total Quality Management, self-assessment, EFQM model, relationships .

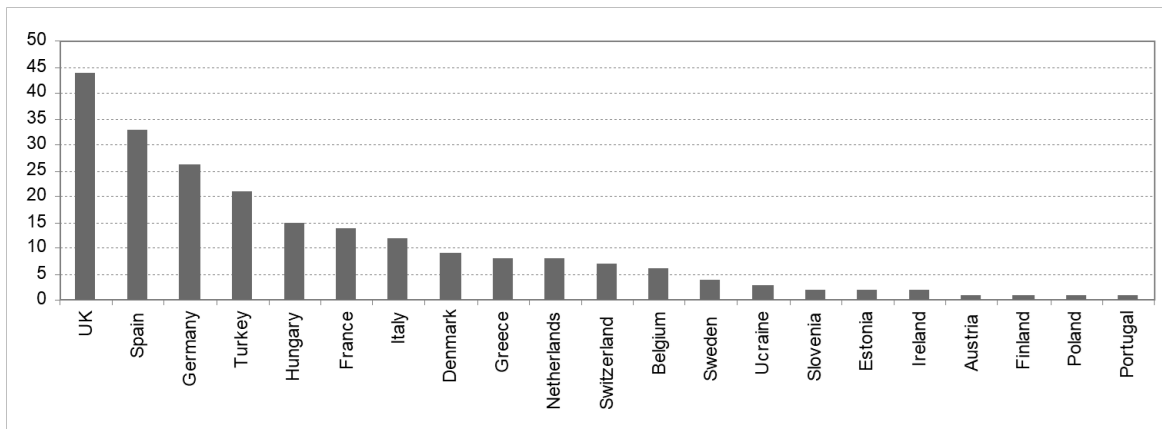
Classification: Research paper.

1. Introduction

Total Quality Management (TQM) may be defined as something that is both complex and ambiguous. Nevertheless, some key elements or principles are common to all TQM models (Dahlgaard-Park, 1999; Reed *et al.*, 2000; Sousa and Voss, 2002): customer satisfaction, continuous improvement, commitment and leadership on the part of top management, involvement and support on the part of employees, teamwork, measurement via indicators and feedback.

The TQM self-assessment models, such as the EFQM model—the leading quality award model together with the Malcolm Baldrige Model (Dahlgaard-Park, 2008)—, have contributed immensely towards clarifying and disseminating TQM in Europe. According to José Ignacio Wert, the former President of EFQM, there were around 30,000 European organisations that were using the EFQM model (Wert, 2006). Regarding the dissemination of EFQM Excellence Awards,

as can be seen in graph 1, United Kingdom, Spain and Germany ranked among the countries with the greatest number of recognitions.



Graph 1. EFQM Excellence Awards by country (1992-2009)

Source: put together by the author from information obtained from EFQM (2010).

However, despite the unprecedented success in the practical application of the model, empirical academic research regarding its reliability has not been developed parallel to this (Bou-Llusar *et al.*, 2005; Williams *et al.*, 2006; Bou-Llusar *et al.*, 2009), and, as Eskildsen *et al.* (2001) pointed out several years ago, there are clear shortcomings existing when analysing the consistency of the model. As Dahlggaard-Park (2008) underlined, clear indications of cause and effect relationships in terms of enabler and results criteria may be questioned. Furthermore, as Williams *et al.* (2006) stressed, there is a major lack of academic work that contrasts the relationships within the EFQM model, a basic issue for the legitimisation of any management model. In this respect, this article constitutes a contribution to the aforementioned.

The article is structured as follows: following this introductory section, the literature review and the conceptual framework are included in the second section; in the following – third – section, the research model and its corresponding hypotheses are articulated; in the fourth, the methodology and data used are analysed; the fifth section contains the results of the empirical research; in the sixth are to be found the discussion and conclusions drawn from the article, with

their practical implications and limitations; the seventh and last section contains the bibliographical references.

2. Literature review and conceptual framework

The EFQM model can be considered as a holistic and integrative approach, where strategic, managerial and operational control processes are integrated in the model (Dahlgaard-Park *et al.*, 2001).

In the literature, some of the internal relations existing in the EFQM model have been analysed in previous research. Analysis tended to focus on the study of the inter-relation existing between some of the elements or categories (theoretical constructs) that make up the model (Dijkstra, 1997; Eskildsen *et al.*, 2001).

More recently, Bou-Llusar *et al.* (2005), analysed the EFQM model in depth, based on the information supplied by a further set of companies, in order to try and assess the causal inter-relation existing between the *enabler* and *results* criteria; the authors ascertained that the *enabler* criteria are indeed related in a balanced way to the *results*. In another interesting work by these same authors (Bou-Llusar *et al.*, 2009), they also ascertained that the EFQM model reliably reflects the premises of TQM.

From the practitioners point of view Tejedor-Panchón (2004) carried out a study in 168 companies' candidates to obtain two regionals quality awards in Spain, in order to test the existent relations on the EFQM model from 1999. Later, Carmona *et al.* (2010) follows the study to test the 2003 version of the EFQM model, using more than 300 evaluations candidates to six different regional awards in Spain.

From the academic field Calvo de Mora and Criado (2005) analysed the reliability, validity and predictive power of adaptation of the EFQM model applied to the state university sphere of activity, based on a sample of 111 Spanish university centres. This is a work which, despite

focusing on a very specific sector of activity for which purpose the EFQM model has been adapted, constitutes a background and key reference point for this research. On the same way that our previous exploratory works (Blind reference) that contributes to this final paper.

Similarly, in a very recent investigation Gómez-Gómez *et al.* (2011) carried out an exploratory analysis of the relationships in the 2003 version of the EFQM model, using data from 68 self-evaluations of both public and private Spanish organizations. In this study the authors also analyse if there are possible differences in the EFQM implementation between public and private organizations.

However, despite these interesting contributions that have been quoted, no study has been detected among those reviewed that has empirically contrasted the relationships within the EFQM model based on reliable primary sources like the ones used in this article. Specifically, it is information deriving from the external assessments themselves made using a very rigorous protocol by independent professionals, on the scores of the categories and subcategories of the EFQM model. Without considering academic studies like the ones provided by Calvo de Mora and Criado (2005) and Gómez-Gómez *et al.* (2011) , the rest of studies are using always data obtained generally from a survey addressed to company managers.

On the contrary, similar studies based on external assessments has been carried out in academic literature for other TQM models such as the Malcolm Baldrige model (e.g. Wilson and Collier, 2000; Flynn and Saladin, 2001; Pannirselvam and Ferguson, 2001; Jayamaha *et al.*, 2011; He *et al.*, 2011) and other Business Excellence models used in the Asia Pacific region (Su *et al.*, 2003; Jayamaha *et al.*, 2008; Jayamaha *et al.*, 2009) or in South America (e.g. González *et al.*, 2009) .

As stressed by Williams *et al.* (2006) some years ago, after so many years during which the EFQM model has been used, it is time for it to be analysed. However, the EFQM model involves so directional paths between constructs or boxes, which makes its study very difficult. Considering that, researchers as Jayamaha *et al.* (2009), have develop parsimonious models that help on

their objective. This will be the case in order to analyse which relationships on the categories of the EFQM model can be considered robust and significant from the statistical point of view, and consequently, give some lights about the underlying theory.

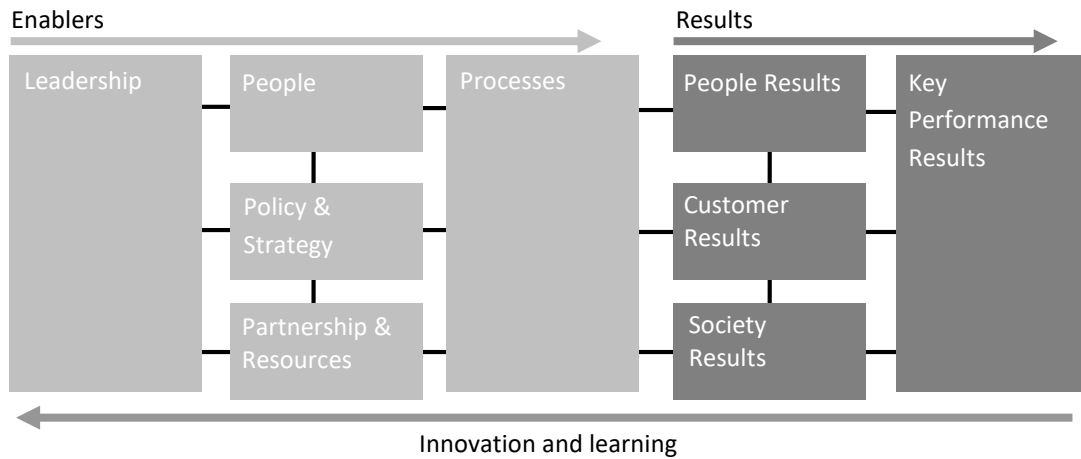
It should be pointed out that two aspects are new ones in this type of study of the EFQM model: on the one hand, the contribution of the point of view of the assessor and, on the other, the adjustment to the EFQM model itself. As stressed by Jayamaha *et al.* (2009) there is still insufficient evidence of the validity of Business Excellence models such as EFQM, due to the lack of available data, namely scores secured by award applicants on the measurement items. There is a huge difficulty in obtaining data related to self-assessment in accordance with the EFQM model, a source of data of a confidential nature with major exploratory potential (Heras *et al.*, 2009).

In this case, it must be underlie the specific characteristics of the used data in order to contrast the EFQM relations, that differentiates this work from previous academics studies (e.g. Calvo de Mora and Criado, 2005; Gómez-Gómez *et al.* 2011). First of all, the number of used observations (242 assessments), and the large period of time included (from 1998 to 2008), both numbers higher than previous works. Secondly, as it is described in section 4, it is relevant the quality of the observations, because they came from external valuations carried out by different professionals with an evaluation criteria establish by an independent and rigorous organization. And finally, because in this research the work is done using scores from criteria and sub-criteria from the EFQM model, presenting a more desegregated work than any previous one that uses only the criteria boxes.

3. Research model and hypothesis

When analysing the relationships within the categories of the EFQM model² (EFQM, 2003), see figure 1, the objective of this article is to explore the extent to which the *agent* or *enabler* criteria are to be found in practice, related to the *results* criteria.

Figure 1. EFQM model



Source: EFQM, 2003.

However, and after considering the objective of the model and difficulties on this kind of analysis, it is very interesting to study first whether the relationships insinuated by the model when pinpointing the different categories or *boxes* of criteria from left to right truly refer to the impact each group of *boxes* has over the criteria located on the right.

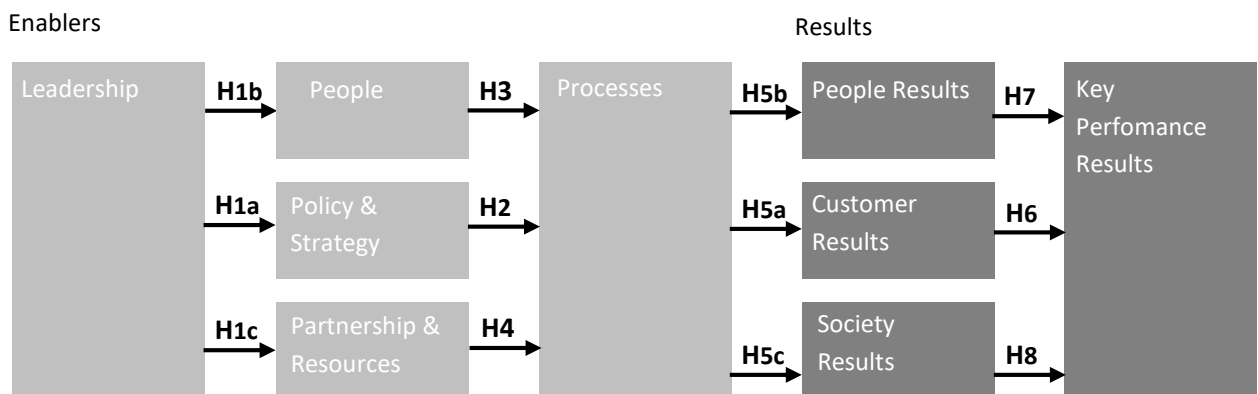
Additionally, and considering that the model suggests a causal relationship among the different criteria that comprise it from left to right, ranging from the criteria of a more strategic nature

² Formally, it should be pointed out that the 2003 version of the EFQM will be the one subject to analysis (this being adapted in case no data should happen to be available for some sub-criteria pertaining to the aforementioned version in the empirical part).

(leadership) to operative results (key results), this article will focused on it. Basically, the study analyse leadership as the strength that drives organization to their results, through the provision of people, resources, partnerships and the formulation of policies and strategy; and the relationship between its implementation on processes and the obtained results. Thus, the first criterion (leadership) has an impact on criteria of a tactical nature (criteria 2, 3 and 4) and the latter, in turn, on operative criteria. In this way, the processes explain the results in customers, people and society and all these in turn ultimately explain the operative results.

It is important to note that figure 1 seems that policy and strategy should not be detached from the utilization of people and physical resources and partnerships from organizational; however, in order to maintain the model simple, it will not be considered this direct relation. On the same way, having customer results in the middle seems to imply that customer outcomes should be balanced against the outcomes for other two key stakeholders: people and the wider society. Both concepts have not been included in the model presented on figure 2.

Figure 2. Research model of the relationship among the criteria of the EFQM model



Source: Own elaboration.

Additionally, it is necessary to stress that this model doesn't include rearranging or regrouping of criteria, the theoretical construct of the EFQM model that appears on figure 1, or the sub-criteria, the concepts that capture the essence of these constructs, according to any others possible latent constructs.

Therefore, criteria or sub-criteria will not be treated as has been done in other works among the literature available that analyse other quality management models, even though this may be to the detriment of the reliability of the constructs used and also subsequently to the detriment of the fitness of the sample to the model. In short, we shall assess the model solely with the aim of detecting any possible limitations in the sample, rather than eliminating or rearranging any items into different criteria.

To sum up and taking the inter-relations put forward by the EFQM model itself as a reference (EFQM, 2003), a research model (see Figure 2) is proposed for the purpose of analysing the impact of *enabler* criteria on *results*. It includes twelve working hypotheses, each one corresponding to a link or inter-relation existing between some category or element of the model, whether an element that may belong to *enabler* or *results* criteria. They will be analysed by means of a structural equation model using SmartPLS software.

Following the work of Gómez-Gómez *et al.* (2011), our hypotheses assumes the sense of the causal relationships implicit in the model, from left to right. The model proposed is sufficiently explicit if the content and objectives of the EFQM self-assessment model are analysed, and we shall therefore draw up the twelve hypotheses it suggests. Specifically, we shall clarify the list of hypotheses that are set out in the arrows that go from right to left, given that the direction of these relationships is determined by the EFQM model itself (EFQM, 2003). Likewise, we should take into account that this main direction of the relationships between *enablers* and *results* is also posited by the academic literature (e.g. Reiner, 2002; Bou-Llusar *et al.* 2005) and by other models such as the Malcolm Baldrige model (Flynn and Saladin, 2001; Pannirselvam and Ferguson, 2001) and the scheme proposed by Anderson *et al.* (1994), based on theories proposed by Deming. Furthermore, these causal relationships among the criteria of the EFQM model are the main relations that have been analysed in the specialized academic literature of

the field (e.g. Bou-Llusar *et al.*, 2005; Eskildsen and Dahlggaard, 2000; Eskildsen *et al.*, 2001; Calvo de Mora and Criado, 2005).

Additionally to this direct reference to the causal relationships implicit in the model, in order to support the hypothesis of our work presented in Figure 2, we use the academic literature focused in the analysis of Business Excellence models. Consequently, we present the hypothesis of our work as follows:

The leadership and commitment of the management have a positive influence on policy and strategy (e.g. Dijkstra, 1997; Eskildsen *et al.* 2000; Pannirselvam and Ferguson, 2001; Calvo-Mora *et al.* 2005; Jayamaha *et al.*, 2008).

H1a. Leadership has a positive relationship with Policy and Strategy

The leadership and commitment of the management is a driver of the enabler People (e.g. Eskildsen *et al.* 2000; Meyer and Collier, 2001; Calvo-Mora *et al.* 2005; Badri *et al.* 2006).

H1b. Leadership has a positive relationship with People

The leadership and commitment of the management has a positive influence on Partnerships and Resources (e.g. Calvo-Mora *et al.* 2005; Badri *et al.* 2006; Gómez-Gómez *et al.*, 2011).

H1c. Leadership has a positive relationship with Partnerships and Resources

The enabler Policy and strategy have a positive influence on process management (e.g. Calvo-Mora *et al.* 2005; Eskildsen and Dahlggaard, 2000; Eskildsen *et al.* 2001; Eskildsen *et al.* 2002).

H2. Policy and Strategy has a positive relationship with Pocesesses

The enabler People management has a positive influence on process management (e.g. Su *et al.*, 2003; Calvo-Mora *et al.* 2005; He *et al.*, 2011)

H3. People has a positive relationship with Pocesesses

The enabler Partnership and Resources has a positive influence on process management (e.g. Eskildsen and Dahlggaard, 2000; Tejedor-Panchón, 2004; Calvo-Mora *et al.* 2005; He *et al.*, 2011).

H4. Partnerships and Resources has a positive relationship with Pocesesses

Process management affects positively to Customer satisfaction (e.g. Wilson and Collier, 2000; Su *et al.*, 2003; Tejedor-Panchón, 2004)

H5a. Processes has a positive relationship with Customer results

Process management affects positively to People results (e.g. Eskildsen and Kanji, 1998; Eskildsen and Dahlgaard, 2000; Gómez-Gómez *et al.*, 2011).

H5b. Processes has a positive relationship with People results

Process management affects positively to the Society results (e.g. Eskildsen and Kanji, 1998; Westlund, 2001; Calvo de Mora and Criado, 2005).

H5c. Processes has a positive relationship with Society results

The satisfaction of the customers has a positive influence on the key performance results of the organizations (e.g. Eskildsen and Kanji, 1998; Eskildsen and Dahlgaard, 2000; Gómez-Gómez *et al.*, 2011)

H6. Customer results has a positive relationship with Key performance results

People-related results have a positive influence on the key performance results of the organizations (e.g. Eskildsen and Kanji, 1998; Prabhu *et al.* 2000; González *et al.*, 2009)

H7. People results has a positive relationship with Key performance results

Society-related results have a positive influence on the key performance results of the organizations (e.g. Westlund, 2001; Reiner, 2002; Gómez-Gómez *et al.*, 2011)

H8. Society results has a positive relationship with Key performance results

It is necessary to now make a final observation about the model being analysed. In accordance with Calvo de Mora and Criado (2005), we shall use latent constructs with reflective indicators for the *enablers* and with formative indicators for the *results*. In fact, the *enabler* sub-criteria evidence and display the latent construct that encompasses them. The sub-criteria of a specific enabler are affected by the same latent construct (Chin, 1998). However, according to Collier and Bienstock (2006), we shall consider the *results* criteria to be formative: they are the result of adding the respective items in order to obtain a global value. Indeed, formative items generate or give rise to the latent variable (Fornell, 1982). Each of these results criteria comprises two sub-criteria: one, which measures perception and another, constructed by the indicators themselves used by the organisation to measure the criterion. Therefore, these indicators do not necessarily have to be correlated. They may manifest themselves as being separate from each other (Chin and Gopal, 1995).

4. Methodology and data

The empirical analysis has been based on data provided by Euskalit, the Basque Foundation for Quality, referring to scores that have been obtained in external assessments of organisations from the Basque Autonomous Community (BAC) in Spain, for the years between 1998 and 2008, inclusive. Attention should be drawn to the strong dissemination of the EFQM model in the BAC: organisations from this region awarded 21 of the 29 cases of recognition between 2001 and 2009 of those awarded to Spanish organisations by the EFQM.

As for the reliability of the data, it is interesting to point out that the theoretical reliability of data obtained from external assessment processes has been highlighted in academic literature (e.g. Pannirselvam and Ferguson, 2001). By focusing on the case of Euskalit, it should be pointed out that the EFQM assessors who took part in the field work are not EFQM licensees (neither from Euskalit nor from any other similar organisation). The assessors belong to the Euskalit Assessors' Club; they are people who have received specialist formal training in the EFQM self-assessment model and who, without any financial gain at all, are committed to improving the management quality of organisations within their milieu. To sum up, these assessors constitute a very reliable, independent source of information owing to their training and specialisation in EFQM model self-assessment and assessment work.

On the other hand, it is also interesting to add that only international EFQM recognition obtained by companies from the BAC evidence the rigorous work carried out by external assessors from Euskalit; attention should also be drawn to the fact that the companies externally assessed by external assessors from the EFQM Foundation have always obtained higher scores than those obtained in external assessments made by Euskalit. In our opinion, this evidence corroborates the reliability of the data used.

The customary work process for finding a model that adapts to a sample involves two stages. In the first is carried out an exploratory analysis until a model is determined that can then be

validated in the second, confirmatory phase. In our case, we consider the EFQM model to be good as it is, without removing or adding anything. In any case, we shall then also go on to analyse the subscales – not with the aim of refining these scales as has been stated, but rather, to ascertain their degree of reliability and validity. This will provide criteria when drawing conclusions from the subsequent analysis.

A structural equation model will be used for this subsequent analysis using the Partial Least Squares (PLS) technique, which enables the path analysis among latent constructs to be carried out (Ringle *et al.*, 2005). Smart-PLS software will be used for such purpose. The aim of this technique is to predict the latent variables and is based on covariance, to the extent that it is applied in order to explain the variance of the independent variables.

The main advantages of this technique over those based on covariance lie in the fact that it is less demanding with the distribution of the sample variables and with the size of the sample. Indeed, PLS enables latent constructs to be modelled under conditions of non-normality (Compeau and Higgins, 1995). In contrast, the main disadvantage involves the fact that it proves to be not so sufficient in analyses of an exploratory nature. In fact, rather than taking on equivalent weights for all the indicators of a single latent variable, PLS permits greatest weights for those items with a stronger correlation with the latent variable. That is why it is suitable for application in our study, as our aim is not to search for a new model, but rather, to analyse the causality of an existing model that has been widespread and used for over a decade now (Eskildsen *et al.*, 2001).

Specifically, the *path analysis* has been used to estimate the robustness of the relationships existing among the new constructs. This is a multi-variant analytical method for examining groups of relationships established by linear causal models (Li, 1975; Jöreskog and Sörbom, 1993). The EFQM model represents the causal relationships among the different sub-criteria,

and so this methodology is suitable for the purpose of our analysis (Pannirselvam and Ferguson, 2001).

5. Results

5.1. Sample and statistical description of the variables

The study sample is made up of 242 assessments of companies from the BAC made by Euskalit according to the EFQM model between the years 1999 and 2008. Some of the companies were assessed more than once during this period. The elements that make up the sample are assessments rather than companies. We are unable to identify each of the companies assessed owing to data confidentiality.

The average scores in the different sub-criteria of the EFQM model are within a range of between 25.72 and 49.84, with the score range being between 0 and 100 in the case of all criteria. In no case is the average value of the scale exceeded. Most of the average scores of the sub-criteria are within a range of between 40 and 50. The average scores of each criterion have also been calculated, and these values are between 42.02 and 45.80 in the case of the *enabler* criteria. On the other hand, the average values of the *results* criteria are 44.65 for *results in customers*, 42.37 for *results in people*, 28.78 for *results in society*, and 45.06 for *key results*. It is noted that the *results* criteria for society are far lower than the other criteria.

As regards variance, it is observed that this is between 46.08 and 163.76. It should be noted that variance in the items pertaining to criterion 8 (*results in society*) is also very different compared to variance in the other sub-criteria: the latter is far higher. All this leads one to draw the conclusion that the criterion *results in society* may prove difficult to fit in to a model that lists EFQM criteria.

5.2. Assessment of the measurement model

We shall now proceed to analyse four aspects in this section: the individual reliability of the items; the reliability of the subscales or internal consistency; the convergent validity and, lastly, the discriminant validity of the constructs.

The individual reliability of the item for constructs with reflective indicators is guaranteed by a load value of over 0.707. Carmines and Zeller (1979) point out that a higher value than this enables the fact that the indicator forms an integral part of the construct to be ascertained. Table 1 shows loads of the external model (in the diagonal in bold) and also includes the cross-loadings. As is noted, four of the sub-criteria do not reach this threshold. Although other authors Barcklay *et al.* (1995) accept lower values, we have not pursued the usual procedure for refinement of the subscales since, as has been previously stated, our aim has been to test the relationships implicit in the EFQM model, rather than seeking the best model of relationships that is adapted to the sample. Despite this, a high degree of individual reliability of the items is noted. On the other hand, it can be observed as the cross-loadings are lower than the figures in bold of the diagonal.

The sub-criteria with load on their corresponding factor below 0.707 are:

1e. Refers to motivation, support and recognition of people by the leaders of the organisation. Data is only available for companies audited in 2004 and subsequent years. This is a criterion that is incorporated in the 2003 version.

2e. Refers to communication and introduction of policy and strategy. In reality, this is just on the limit and in fact 56 companies have only answered this indicator, which explains such a weak load.

3c. Measures the involvement and extent to which responsibilities are assumed: this is a value that is very close to the boundary value established.

4b. Its load is 0.7060, just below the established limit of 0.707

Table 1. Loads of the external model and cross-loadings.

	1 Leadership	2 Policy & strategy	3 People	4 Alliances & resources	5 Processes	6 Customer results	7 People results	8 Society results	9 Key results
1a	0,862337	0,649935	0,612329	0,412193	0,569921	0,190902	0,267771	0,252195	0,255406
1b	0,868487	0,601667	0,615381	0,387993	0,594410	0,199252	0,274194	0,157130	0,198375
1c	0,810808	0,540306	0,449904	0,477224	0,547413	0,114461	0,173746	0,260787	0,200443
1d	0,731644	0,360064	0,506374	0,225561	0,404066	0,054629	0,186368	0,102840	0,108040
1e	0,592900	0,313813	0,338968	0,313155	0,407039	0,068361	0,101311	0,002411	0,231409
2a	0,522227	0,828280	0,476533	0,579623	0,543213	0,164696	0,188920	0,235477	0,385395
2b	0,501113	0,827361	0,473116	0,567407	0,495882	0,161882	0,099248	0,177513	0,394067
2c	0,516045	0,813988	0,540452	0,535269	0,490135	0,197538	0,198039	0,251970	0,379794
2d	0,595685	0,764908	0,527075	0,405635	0,593597	0,126935	0,225122	0,175011	0,233122
2e	0,294278	0,492084	0,260649	0,240713	0,298331	0,073924	0,031817	0,091398	0,048144
3a	0,498287	0,554316	0,797021	0,441089	0,380444	0,175737	0,259334	0,062188	0,317818
3b	0,472917	0,497005	0,788284	0,436744	0,446211	0,248033	0,264931	0,127057	0,263431
3c	0,484024	0,349446	0,699835	0,160049	0,409132	0,096674	0,222494	0,146223	0,054509
3d	0,547333	0,470382	0,768599	0,364639	0,448884	0,124072	0,240292	0,202194	0,135039
3e	0,478558	0,443118	0,708039	0,377483	0,307242	0,120184	0,287956	0,179970	0,261925
4a	0,340048	0,458875	0,330106	0,728077	0,403737	0,064070	0,096169	0,228048	0,296571
4b	0,317085	0,478544	0,369767	0,706028	0,355778	0,224095	0,139434	0,089247	0,459446
4c	0,275613	0,448646	0,281671	0,743846	0,431502	0,219939	0,082839	0,245579	0,389055
4d	0,377796	0,427771	0,342285	0,756283	0,431124	0,038753	0,086283	0,141031	0,307567
4e	0,439440	0,539684	0,438865	0,791643	0,503569	0,124910	0,118499	0,144626	0,346084
5a	0,479218	0,514392	0,430807	0,315550	0,692735	0,093850	0,165762	0,251899	0,054750
5b	0,532065	0,561884	0,413161	0,456214	0,829758	0,176945	0,185402	0,208776	0,197828
5c	0,537221	0,473935	0,397355	0,476089	0,751087	0,153535	0,155036	0,101027	0,317012
5d	0,455140	0,366453	0,351968	0,412080	0,707213	0,142974	0,207911	0,074999	0,260979
5e	0,500391	0,549888	0,430439	0,518873	0,822487	0,220170	0,178187	0,187016	0,300960
6a	0,172243	0,109202	0,097546	0,004090	0,157065	0,446949	0,223041	0,058842	0,160044
6b	0,146705	0,187970	0,199130	0,192309	0,189811	0,977516	0,479116	0,225563	0,372690
7a	0,218050	0,092936	0,328557	-0,01166	0,105115	0,273146	0,435267	0,119769	0,013156
7b	0,258442	0,208530	0,320320	0,146321	0,231136	0,484307	0,996916	0,253059	0,287169
8a	0,168982	0,174922	0,182074	0,191858	0,162357	0,221306	0,260490	0,840257	0,091173
8b	0,212713	0,263797	0,164830	0,212840	0,222809	0,186987	0,211887	0,943691	0,137585
9a	0,238807	0,374772	0,241199	0,452179	0,300364	0,312056	0,228191	0,158962	0,833196
9b	0,225698	0,360583	0,251906	0,421694	0,247642	0,360110	0,265427	0,096454	0,944284

Note: all the loads of the *enabler* criteria are significant (t-value>1.96) Source: put together by the authors from data supplied by Euskalit.

Regarding to the items of the results constructs, just two of them are below 0.5: 6a and 7a.

Both are measurements of perception for the results.

The robustness of these loads is analysed below using a *bootstrapping* process. Those that are below a t value of 1.96 - and in which their robustness is therefore not assured - are items 6a, 7a, 8a, 8b and 9a. We wish to put on record here that in view of these results that in spite of the

weakness found in these four items of enablers criteria, we proceed with the analysis, because we insist once again that our aim is to try out the unaltered EFQM model.

The second point to be analysed in order to assess the measurement model is the internal consistency of the subscales of enabler criteria, i.e. the reliability of the subscales.

The five constructs evidence satisfactory values according to the criteria proposed by Hair *et al.* (1998). Five factorial analyses were also carried out in order to research the one-dimensional nature of the enabler constructs. In all cases, a single factor was extracted and the amount of variability captured ranges from 57.86% to 65.50%.

Another rate used to assess the reliability of the reflective constructs is the composite reliability. Nunnally and Bernstein (1994) suggested a minimum 0.7 for valid modest reliability for the first stages of the research, although the recommended value is 0.8 for basic research purposes. The five values obtained are within a range of 0.862 and 0.884 (see table 4).

Table 2. Reliability analysis of the enabler constructs

Construct	Items	Cronbach's alpha	Range of Cronbach's alpha by eliminating an item	Range of correlations of items and subscale total	Type of construct	Unidimensionality analysis	
						KMO	% variance captured by the factor
1 Leadership	1a, 1b, 1c, 1d, 1e	0.866	0.822 – 0.864	0.586 – 0.748	Reflective	0.855	65.50%
2 Policy and strategy	2a, 2b, 2c, 2d, 2e	0.854	0.803 – 0.872	0.532 – 0.748	Reflective	0.841	65.47%
3 People	3a, 3b, 3c, 3d, 3e	0.816	0.761 – 0.801	0.535 – 0.667	Reflective	0.787	58.11%
4 Alliances and resources	4a, 4b, 4c, 4d, 4e	0.814	0.766 – 0.796	0.554 – 0.642	Reflective	0.832	57.86%
5 Processes	5a, 5b, 5c, 5d, 5e	0.829	0.760 – 0.834	0.515 – 0.753	Reflective	0.771	60.76%

Source: put together by the author from data supplied by Euskalit.

The third point to be analysed is that of convergent validity (it also only applies for enabler criteria; the result criteria are formative). To this end, the average variance extracted (AVE), which provides the amount of variance obtained via its indicators related to variance due to measuring error. Fornell and Larcker (1981) recommend values over 0.5. The AVE indicators for

the five *agent* or *enabler* criteria are between 0.5561 and 0.6084 (see table 4). Convergent validity is therefore assured.

The fourth and final aspect to be analysed in order to assess the measurement model is that of discriminant validity. We use the criteria used by Fornell and Larcker (1981): the square root of the AVE should be higher than the correlations evidenced by this construct with the other constructs. Table 3 shows the square root diagonal of the AVE, while the other cells show the correlations. The initials N.A. indicate the fact that the procedure is not applicable to formative constructs – in our case, those referring to *results*.

Table 3. Discriminant validity

	<i>1 Leadership</i>	<i>2 Policy and strategy</i>	<i>3 People</i>	<i>4 Alliances and resources</i>	<i>5 Processes</i>	<i>6 Customer results</i>	<i>7 People results</i>	<i>8 Society results</i>	<i>9 Key results</i>
<i>1 Leadership</i>	<i>0.7800</i>								
<i>2 Policy and strategy</i>	0.659048	<i>0.7560</i>							
<i>3 People</i>	0.660150	0.615709	<i>0.7532</i>						
<i>4 Alliances and resources</i>	0.475616	0.628333	0.472770	<i>0.7454</i>					
<i>5 Processes</i>	0.656164	0.656767	0.532744	0.575270	<i>0.7628</i>				
<i>6 Customer results</i>	0.178827	0.196964	0.203636	0.168447	0.213556	N.A.			
<i>7 People results</i>	0.230745	0.202988	0.274266	0.155414	0.223468	0.452885	N.A.		
<i>8 Society results</i>	0.216968	0.262638	0.177850	0.218836	0.225075	0.198674	0.222729	N.A.	
<i>9 Key results</i>	0.254630	0.398992	0.269880	0.474326	0.295095	0.377769	0.300357	0.136355	N.A.

Note: correlations between latent variables under the main diagonal. In the diagonal are the square roots of the AVE, in italics.

Source: put together by the authors from data supplied by Euskalit.

It is noted that the reflective constructs comply with the criterion used by Fornell and Larcker (1981) to guarantee discriminant validity. For their part, the formative indicators also exceed the condition put forward by Fornell and Larcker (1981) and by Luque (2000), as the maximum correlation is 0.45, far from the maximum 0.9 threshold recommended.

5.3. Assessment of the structural model

PLS does not use fit indices: the fitness is established with significant path coefficients and high R2 values. The variability explained by the model for the dependent latent variables on the left part of the model (*enabler* criteria) is higher than 0.40 in four cases. In the case of *process* criterion, it reaches nearly 50%. However, the model fails to explain so well the constructs on the right part that refers to the *results* criteria. In fact, the reliability analysis for these constructs already reveals possible problems in this part of the model. However, we once again insist on the fact that the initial purpose of this analysis is to study the EFQM model *as it is*, without any alteration (table 4).

Table 4. Overview of the model

	AVE	Composite reliability	R2	Communality	Redundancy
1 Leadership	0.608446	0.884195		0.608446	
2 Policy and strategy	0.571514	0.866546	0.434345	0.571514	0.244197
3 People	0.567289	0.867378	0.435798	0.567289	0.246166
4 Alliances and resources	0.555639	0.861835	0.226211	0.555639	0.123675
5 Processes	0.581812	0.873684	0.493242	0.581812	0.209878
6 Customer results			0.045606	0.588419	0.029022
7 People results			0.049938	0.502893	0.030212
8 Society results			0.050659	0.741018	0.037966
9 Key results			0.165413	0.790739	0.107985

Source: put together by the authors from data supplied by Euskalit.

Table 5 shows the coefficients of the internal model. A *bootstrapping* process has been used to test the robustness of these coefficients consisting of 500 samples of 100 elements each. In each box is noted down whether the corresponding hypothesis is accepted or rejected.

Table 5. Coefficients of steps between internal variables

	1 Leadership	2 Policy and strategy	3 People	4 Alliances and resources	5 Processes	6 Customer results	7 People results	8 Society results	9 Key results
1 Leadership		0.6590 (10.4058) H1a Accepted	0.6610 (11.3460) H1b Accepted	0.4756 (6.5347) H1c Accepted					
2 Policy and strategy					0.3969 (3.2827)				

					H2 Accepted				
3 People					0.1723 (1.7243)				
					H3 Rejected				
4 Alliances and resources					0.2422 (2.0561)				
					H4 Accepted				
5 Processes						0.2136 (2.0166)	0.2234 (1.8404)	0.2250 (1.8557)	
						H5a Accepted	H5b Rejected	H5c Rejected	
6 Customer results									0.2989 (2.2357)
									H6 Accepted
7 People results									0.1555 (1.0427)
									H7 Rejected
8 Society results									0.0423 (0.3276)
									H8 Rejected
9 Key results									

Source: put together from data supplied by Euskalit.

Note: the t-value is in brackets. The significant coefficients at a 0.05 level are in bold. Each of these results obtained are used in order to contrast one working hypotheses.

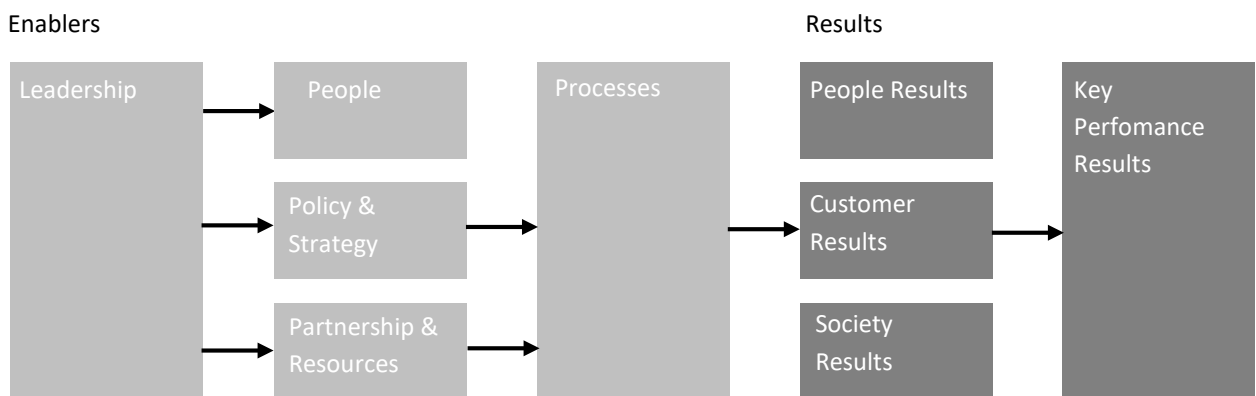
Figure 3 displays the results from table 5. This figure only shows the significant paths between criteria. A greater density of robust coefficients is noted on the left part. Indeed, the *leadership* criterion goes a long way to explain the results obtained in the *agent* criteria of *policy and strategy*, *people* and *alliances and resources*. The processes depend to a large extent on previous criteria (*policy and strategy* and *alliances and resources*). However, they only impact on one of the *results* criteria (*results in customers*).

There is only one path from the *leadership agent* to the *key results*. If one may be permitted to refer to the classic name used in project management, we might say that the “critical path”

traverses *customer results*. These criteria are especially determinant, as the model indicated the fact that they are a necessary step on the way to obtaining key results.

The left part of the model (the *enabler* criteria) shows robust coefficients: only one of the six is not statistically significant, although it should be pointed out that the t-value associated with the relationship between the *people enabler* and the *process enabler* is 1.72, close to the boundary value established by 1.96. In other words, although this relationship is not significant to a level of 5%, it is so when slightly relaxing it.

Figure 3. Significant coefficients



Source: Own elaboration. Note: coefficients significant at a level 0.05.

To sum up, it is noted that the *enabler* criteria are closely correlated. On the other hand, the *results* criteria are not so inter-related as the *enablers*. The prior analysis involving measuring assessment already enabled the results to be disclosed as shown in table 5: the existence of a major number of rejected hypotheses in the bottom right area of the table, which refers to the relationships among *results*. Analogously, the same phenomenon is observed in the up right area, regarding to people results.

6. Conclusions

In the course of the analysis it has been ascertained that there is a major impact of the *leadership* enabler on the pursuit of policy and strategy in organisations, and also on the *people* criteria and

on *alliances and resources*. The importance of leadership in accordance with what is described in classical literature about TQM is clearly in evidence. It should also be pointed out that both the *policy and strategy* criterion and *alliances and resources* impact on the *process* criterion; however, the *people* enabler criterion does not have a significant impact on an improvement in processes.

On the other hand, the *process* enabler only impacts on *customer results*. This criterion, in turn, is the only one that explains the *key results* criterion. In this sense, attention should be drawn to the fact that both the results in the *people* criterion and the *results in society* criterion are excluded from the model, given that no significant relationships have been detected with other criteria.

In short, the left side of the EFQM model is quite better supported by our data than the right side. Our data provides evidences that the relationships among enabler criteria work well; nevertheless, the data do not support the relationships among results, neither between enablers and results. Therefore, with our data the EFQM model fails explaining the right side. It could imply that enablers really do not cause results. Future research should confirm these exploratory results that we provide in the article.

To sum up, several of the relationships among the constructs proposed by the EFQM model are significant: seven of the twelve suggested by the model. These conclusions would seem to coincide with the conclusions drawn from studies carried out previously by Pannirselvam and Ferguson (2001) for the Malcom Baldrige model, and Calvo de Mora and Criado (2005) and Bou-Llusar *et al.* (2005, 2009) for the EFQM model. Indeed, Pannirselvam and Ferguson (2001) proved the existence of significant relationships among the categories and confirmed the consistency of the Malcolm Baldrige National Quality Award framework, based on data obtained from external assessments. Calvo de Mora and Criado (2005) and Bou-Llusar *et al.* (2005, 2009)

also detected strong evidence of the causal relationship between the *enabler* and *result* criteria of the EFQM model based on perceptual data.

Attention should be drawn to the fact that another of the contributions made by this article is without doubt the proposal for using data obtained from external assessments of the EFQM model made by independent assessors, based on a training and assessment protocol such as that defined by Euskalit. As Pannirselvam and Ferguson (2001) point out in their study – and Calvo de Mora and Criado (2005) and Bou-Llusar *et al.* (2005, 2009) also stress when referring to the limitations of their respective studies based on perceptual variables – the information deriving from a third party who assesses this type of TQM model guarantees objectivity, rigour and less characteristic bias introduced than information obtained from the directives of the organisations themselves that adopt these models.

This work has several limitations that need to be fully considered when interpreting the conclusions drawn from it. The main one is related to the relationships that have been analysed. The parsimonious model used proposes relationship between criteria from “left to right”, it is to say, from the criteria of a more strategic nature to operative results. This is the main direction of the relationships between enablers and results proposed by the EFQM and the majority of the academic literature, however these are not the only one proposed in the model. Consequently, other different and/or complementary interpretations of causal relationship have to be analysed on the future.

Another limitation of our work is related to the methodology used to contrast the model. As Calvo de Mora and Criado (2005) point out, structural equations refer to the linearity of the relationships existing among the latent variables – in our case, the criteria pertaining to the EFQM model. In any event, we understand that the tool used is particularly suitable as it is geared towards a predictive causal analysis in situations of great complexity, albeit with sufficient theoretical knowledge in order to develop analyses of a confirmatory nature.

Moreover and as Diamantopoulos and Winklhofer (2001) note, the PLS technique is suitable for assessing models with latent variables with formative and reflective indicators.

A final limitation of the article that we'd like to mention is related to the limited geographic scope of the sample of data used. It would be very interesting to extend this scope to Spain as a whole or even to a series of European Union countries. In this sense, the analysis could be greatly enriched by being able to include data obtained from external assessments presented at awards themselves granted by EFQM.

7. References

Ahmad, S. and Schroeder, G. (2002). The Importance of Recruitment and Selection Process for Sustainability of Total Quality Management. *International Journal of Quality and Reliability Management*, 19 (5), 540-550.

Anderson, J., Rungtusanatham, M., and Schroeder, R. (1994). A theory of quality management underlying the Deming management method, *Academy of Management Review*, (19), pp. 472-509.

Badri, M.A., Selim, H., Alshare, K., Grandon, E.E., Younis, H. and Adsulla, M. (2006). The Baldrige education criteria for performance excellence framework. Empirical test and validation, *International Journal of Quality and Reliability Management*, 23 (9), 1118–1157.

Barclay, D., Higgins, C. and Thompson, R. (1995). The Partial Least Squares (PLS) Approach to Causal Modelling: Personal Computer Adoption and Use as an Illustration. *Technology Studies*, 2 (2), 285-309.

Bou-Llusar, J.C., Escrig-Tena, A.B., Roca-Puig, V. and Beltrán-Martin, I. (2009). An empirical assessment of the EFQM excellence model: evaluation as a TQM framework relative to the MBNQA model. *Journal of Operations Management*, 27 (1), 1-22.

Calvo de Mora, A. and Criado, F. (2005). Análisis de la validez del modelo europeo de excelencia

para la gestión de la calidad en instituciones universitarias: un enfoque directivo. *Revista Europea de Dirección y Economía de la Empresa*, 14 (3), 41-48.

Bou-Llusar, J.C.; Escrig-Tena, A.B.; Roca-Puig, V. and Beltrán-Martín, I. (2005). To What Extent do Enablers Explain Results in the EFQM Excellence Model? *International Journal of Quality & Reliability Management*, 22 (44), 337-353.

Carmines, E.G. and Zeller, R.A. (1979). Reliability and Validity Assesment, Stage University Paper Series on Quantitative Applications in the Social Sciences, nº. 7010, Sage, Beverly Hills, California.

Carmona, M, Rivas, M.A. and Martí, E. (2010). Análisis de las relaciones intravariabes del modelo EFQM mediante Modelos de Ecuaciones Estructurales, *4th International Conference on Industrial Engineering and Industrial Management, XIV Congreso Ingeniería de Organización (CIO)*, 8-10 September, San Sebastian, Spain.

Chin, W. W., and Gopal, A. (1995). Adoption Intention in GSS: Relative Importance of Beliefs. *Data Base*, 26 (2-3), 42-63.

Chin, W.W. (1998). Issues and opinion on structural equation modelling, Commentary in *MIS Quarterly*, 22 (1), 7-16.

Collier, J.E. and Bienstock, C.C. (2006). Measuring Service Quality in E-Retailing. *Journal of Service Research*, 8 (3), 260-275

Compeau, D.H. and Higgins, C.A. (1995). Application of social cognitive theory to training for computer skills. *Information Systems Research*, 8 (6), 118–143.

Dahlggaard-Park, S. M. (1999). The evolution patterns of quality movement. *Total Quality management*, 10 (4-5), 473-480.

- Dahlggaard-Park, S. M. (2008) Reviewing the European excellence model from a management control view. *The TQM Journal*, 20(2), 98 – 119.
- Diamantopoulos, A. and Winklhofer, H. M. (2001). Index Construction with Formative Indicators: An Alternative to Scale Development. *Journal of Marketing Research*, 38, 269-277
- Dijkstra, L. (1997). An Empirical Interpretation of the EFQM Framework. *European Journal of Work and Organizational Psychology*, (6), pp.321-41.
- EFQM (2003). *EFQM Excellence Model*, EFQM, Brussels.
- EFQM (2010). General information from the webpage [www.efqm.org].
- Eskildsen, J.K. and Dahlggaard, J.J. (2000). A causal model for employee satisfaction, *Total Quality Management*, (11), pp. 1081-94.
- Eskildsen, J.K., Kristensen, K. and Juhl, H.J. (2000). The causal structure of the EFQM Excellence Model. In: MAAOE Conference Proceedings, Estes Park, Colorado, US.
- Eskildsen, J.K., Kristensen, K. and Juhl, H.J. (2001). The Criterion of the EFQM Excellence Model. *International Journal of Quality & Reliability Management*, (18), 783 – 795.
- Flynn, B.B. and Saladin, B. (2001). Further Evidence on the Validity of the Theoretical Models Underlying the Baldrige Criteria. *Journal of Operations Management*, (19), 617-652.
- Fornell, C. (1982). A second generation of multivariate analysis: an overview, in C. Fornell [ed.]: *A second generation of multivariate analysis*, 1-21, Praeger, New York.
- Fornell, C. and Larcker, D.F. (1981). Evaluating Structural Equation Models with unobservable variables and measurement error. *Journal of Marketing Research*, 18 (5), 39-50.
- Gómez-Gómez, J., Martínez-Costa, M. and Martínez-Lorente, A.R. (2011). A critical evaluation of the EFQM model. *International Journal of Quality & Reliability Management*, 28(5), 484–502.

González, A., Miles, J., Sorondo, A. and Zeballos, F. (2009). *El Modelo de Mejora Continua del Uruguay Análisis de las relaciones estructurales subyacentes y de la evolución de la puntuación*, Documento de Trabajo, Facultad de Ciencias Empresariales, Universidad Católica Dámaso A. Larrañaga, Uruguay.

Grigg, N.P. and Mann, R.S. (2008). Rewarding Excellence: An International Study into Business Excellence Award Processes. *Quality Management Journal*, 15(3), 26-40.

Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1998). *Multivariate Data Analysis* (5th edn). Prentice Hall International, Upper Saddle River, NJ.

He, Z., Hill, J., Wang, P. and Yue, G. (2011). Validation of the theoretical model underlying the Baldrige criteria: Evidence from China. *Total Quality Management & Business Excellence*, 22(2), 243-263.

Jayamaha, N.P., Grigg, N.P. and Mann, R.S. (2008). Empirical validity of Baldrige criteria: New Zealand evidence. *International Journal of Quality & Reliability Management*, 25(5), 477-493.

Jayamaha, N.P., Grigg, N.P. and Mann, R.S. (2009). A study of the validity of three major business excellence models in the Asia Pacific region. *Total Quality Management & Business Excellence*, 20(11), 1213-1227.

Jayamaha, N.P., Grigg, N.P. and Mann, R.S. (2011). Empirical analysis of the Baldrige Criteria as both an organizational performance measure and a theoretical model. *Measuring Business Excellence*, 15(1), 20-33.

Jöreskog, K.G. and Sörbom, D. (1993). *LISREL 8: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrument Variables and Least Squares Method* (8th ed.). Scientific Software, Morresville, IN.

Li, C.C. (1975). *Path Analysis: A Primer*, The Boxwood Press, Pacific Grove, CA

Nabitz, U., Klazinga, N. and Walburg, J. (2000). The EFQM excellence model: European and Dutch experiences with the EFQM approach in health care. *International Journal for Quality in Health Care*, (12), 191-201.

Nabitz, U., Severens, P., Brink, W.V.D. and Jansen P. (2001). Improving the EFQM Model: An empirical study on model development and theory building using concept mapping. *Total Quality Management*, (12), 69-81.

Nunnally, J.C. and Bernstein, I.H. (1994). *Psychometric Theory*, McGraw-Hill, New York.

Pannirselvam, G.P. and Ferguson, L.A. (2001). A study of the relationships between the Baldrige categories. *International Journal of Quality and Reliability Management*, 18 (1), 14–34.

Prabhu, V., Appleby, A., Yarrow, D. and Mitchell, E. (2000). The impact of ISO 9000 and TQM on best practice/performance. *The TQM Magazine*, 12 (2), 84–91.

Reed, R., Lemark, D.J., & Mero, N.P. (2000). Total quality management and sustainable competitive advantage. *Journal of Quality Management*, 5, 5–26.

Reiner, G. (2002). Analysis of critical factors of company success based on the EFQM Excellence model. In: Proceedings of the 7th World Congress for Total Quality Management, vol. 2, Verona, Italy, pp. 361–366.

Ringle, C.M., Wende, S. and Will, A. (2005). *SmartPLS 2.0 (beta)*. Retrieved from <http://www.smartpls.de>

Sousa, R. and Voss, C. A. (2002). Quality Management Re-visited: a reflective review and agenda for future research. *Journal of Operations Management*, 14 (20), 91–109.

Su, C.T., Li, S.C., and Su, C.H. (2003). An empirical study of the Taiwan National Quality Award causal model. *TQM and Business Excellence*, 14(8), 875–893.

Tejedor-Panchón, F. (2004). *Análisis del Modelo Europeo de Excelencia mediante la aplicación de Modelo de Ecuaciones Estructurales*, Instituto Andaluz de Tecnología (IAT), Seville, Spain.

Wert, J.I. (2006). Prólogo, in Camisón, C., Cruz, S. and González, T. (2006). *Gestión de la calidad: conceptos, enfoques, modelos y sistemas*, Pearson Educación, Madrid.

Westlund, A.H. (2001). Measuring environmental impact on society in the EFQM system. *Total Quality Management*, 12 (1), 125–135.

Williams, R., B. Bertsch, A. Van der Wiele, J. Van Iwaarden, and B. Dale. (2006). Self-assessment against business excellence models: a critique and perspective. *Total Quality Management*, 17, (10), 1287-1300.

Wilson, D.D. and Collier, D.A. (2000). An empirical investigation of the Malcolm Baldrige National Quality award causal model. *Decision Sciences*, 31(2), 361–383.