What do we really mean by “Balanced Scorecard”?

Mike Perkins, Anna Grey and Helge Remmers
The York Management School, University of York, York, UK

Abstract

Purpose – The balanced scorecard (BSC), first created by Kaplan and Norton in 1992, has been developed over the last 20 years, resulting in numerous iterations. There is currently no agreed taxonomy for these iterations, making comparison of research findings difficult. The purpose of this paper is to propose a framework for describing the different iterations.

Design/methodology/approach – In order to assist in understanding the numerous changes, the concept of revision control was introduced. A full review of the literature describing new iterations of the BSC was undertaken and these iterations were classified as minor developments or major generational evolution.

Findings – Eight subsets of the Balanced Scorecard are proposed, grouped into three generations; the first being the initial generation of scorecard, the second generation including strategy maps, and the final generation which includes destination statements.

Practical implications – It is planned that this simple classification will prove useful to practitioners, for describing which generation of the scorecard they propose to implement, and for academics to describe more precisely the scorecard that they are analysing.

Originality/value – This paper is the first to provide a taxonomy for the different versions of the BSC, through a process of identifying and labelling the major and minor changes that have occurred. This allows a more nuanced analysis of the BSC as a tool for managing performance and adding precision to any critique, in that it is clear which version has been used.

Keywords Balanced scorecard, Strategy map, Third generation, Version control

Paper type Literature review

Introduction

A citation analysis conducted by Marr and Schiuma (2003), concluded that the balanced scorecard (BSC) seems to be one of the most influential concepts in the field of performance management and measurement (PMM). Downing (2001) also reports on information from the BSC Collaborative, stating that by the middle of 2001, over 50 per cent of surveyed firms worldwide had adopted the concept along with a further 25 per cent considering its adoption. Other research by Silk (1998) suggests that around 60 per cent of the Fortune 1000 companies in the USA have experimented with the scorecard concept whilst Malmi (2001) encountered high levels of adoption within Finnish businesses. More recently, Rigby (2011) also found that 54 per cent of the 1,230 global firms sampled had reported the use of a BSC. These studies all show how ubiquitous the concept of the BSC has become.

Since the BSC was first created by Kaplan and Norton (1992), numerous other authors have contributed to the idea, leading to the evolution of the BSC concept from a performance measurement tool to a performance management system (PMS). However, as this paper will show, there is much confusion in the literature as to what is meant exactly by a “Balanced Scorecard”. The idea that one author has in mind when discussing the term may be entirely different to that conceptualised by another. This paper examines the term “Balanced Scorecard” as used in the literature, and provides an evolutionary-based framework for categorising the scorecard concept from its origins with Kaplan and Norton (1992), all the way to the “third-generation scorecards” proposed by Lawrie and Cobbold from 2002 onwards.
It is evident from recent studies of the BSC that the distinctions between different versions of the scorecard are still not being made explicit. This is shown in Taylor and Baines’ study (2012), which assessed the reasons why UK universities are increasingly making use of strategic management tools such as the BSC. Whilst they conclude that by using the BSC, universities manage to enhance their competitiveness through following strategies more rigorously, it is not made clear which generation of BSC the universities actually implemented, which is fundamental in understanding the process of adoption. Another recent example is Amando et al. (2011), who assessed the usefulness of BSCs in combination with Data Envelopment Analysis (DEA). The paper refers to a third-generation BSC, but does not explain what is meant by this. In light of such potential confusion, this paper will assist future users of the BSC, both academic and practitioner, by making their discussions on the BSC more precise, and in particular, in understanding the exact type of scorecard that is being used.

It is not just that BSCs have followed a linear evolutionary pathway. As the context and operating environment of individual organisations vary significantly, the BSC has continued to evolve and branch as it crosses industries, sectors, and even countries. These offshoots of the BSC include Non-Profit or Governmental Organisations (NPGO) Scorecards, Public Sector Scorecards and Swedish Scorecards. For example, scorecards have previously been designed for use in NPGOs (Wisniewski and Dickson, 2001; Kaplan and Norton, 2001). The Public Sector Scorecard (Moullin, 2002) was developed specifically to encompass the culture and values of the public and voluntary sectors, and to take into account the wider range of stakeholders. It has also been applied in a number of areas including healthcare (Moullin, 2009, 2011) and voluntary services (Moullin, 2006). Of particular interest has been the adoption of BSCs in Nordic countries, with Kald and Nilsson (2000) reporting that 27 per cent of the 236 major Nordic companies surveyed have implemented a scorecard approach, with the Swedish BSC (Ax and Bjørnenak, 2005) being recognised as a “sub-species” of scorecard. Whilst this paper will not categorise these offshoots, it is important to recognise that they exist as separate entities.

The remainder of this paper is structured as follows: the extant literature relating to the BSC is discussed first, followed by a description of the methodology used to classify the various generations of the scorecards. The results of this classification are presented in a table that will enable BSC practitioners and academics alike to classify and label their scorecard tools appropriately. A thorough discussion of the key characteristics that make up the different versions of the BSCs is also presented.

Literature relating to the BSC concept

The BSC was first developed and proposed by Robert Kaplan and David Norton in 1992 and aimed to combine the use of financial and non-financial measures and provide managers with richer, more detailed information than financial measures alone. The scorecard concept has evolved over a number of years through a series of papers and books by Kaplan and Norton (Kaplan and Norton, 1992, 1993, 1996a,b, 2000, 2004a,b, 2006) transforming the scorecard concept from an innovative, but relatively simple performance measurement tool, through to a complex PMS.

The initial aim of the BSC was to provide managers with a “fast but comprehensive view of the business” (Kaplan and Norton, 1992, p. 71) with the implication that this will be a starting point for improved managerial performance. Throughout the evolution of the scorecard concept, the emphasis has been on ensuring that the executives of an organisation are involved in the process of implementation, therefore ensuring that the
BSC puts strategy and vision, rather than control, at the heart of the PMS (Kaplan and Norton, 1992). Kaplan and Norton reported that companies' early experiences with the BSC fulfilled two managerial needs; namely that of performance information availability and sub-optimisation. By presenting all the disparate information required into one, concise report, the BSC increased the effectiveness of the performance information available to managers and ensured that managers consider operational measures as a whole, allowing them to see whether improvement in one area may have come at the expense of another.

Empirical evidence in support of the BSC

The BSC has been applied in many different areas of the private sector; from the electrical retail industry (Neely et al., 2004), various manufacturing companies (Malina and Selto, 2001; Lohman et al., 2004), and the petrochemical industry (Varma and Deshmukh, 2009) amongst many others. It has also been applied not only in large organisations, but in SME’s (Fernandes et al., 2006; Manville, 2007) as well. Although there are a large number of examples of BSC usage in many different sectors, it must be considered whether the benefits of a BSC outweigh the time and costs of an implementation. To do this, we will investigate the empirical evidence behind the BSC in both the private and public sectors.

In order to evaluate the efficacy of the BSC, it needs to be examined whether there is any proof that BSC implementations have led to an improvement in financial performance ex-post. Considering the huge amount of literature in this area, it is surprising to note that there is little empirical research performed on the impacts of BSC implementations.

It is important to separate the two major concepts of “success” in the BSC literature. One version of success may simply be a BSC successfully being applied to a company and used as a PMS. However, Cinquini and Mitchell (2005) argue that that “success” should be multi-dimensional and dynamic; what is a measure of success now may not be suitable in the near future. A successful BSC implementation, like the implementation of any other PMS, should improve the measurable performance of a firm by improving its ability to manage its assets, whilst at the same time allowing costs to be reduced through an increased understanding of the business environment in which it is operating.

Hendricks et al. (2004), performed a study of Canadian firms that had adopted the BSC. They examined the financial performance as measured by Return on Assets (ROA) and Return on Sales (ROS), both before the implementation of the scorecard as well as after. They found that up to three years after the BSC was implemented, there were no significant performance improvements on either ROA or ROS. However, limitations in the data prevented them from drawing any firm conclusions about post-implementation performance.

Neely et al. (2004) performed a comparison study of two electrical retail firms with 35 pairs of matched branches by location; one firm using a BSC, the other using traditional financial-based performance measures. Over a period of 12 months, the BSC appeared to have no significant impact in terms of sales growth or gross profit growth. They conclude that collecting performance data over a longer period may further assist in understanding the issues around why the BSC implementations appeared to show no improvement in performance, and recommend that further research be carried out in this area.

Empirical evidence in support of the view that Balanced Scorecards lead to improvements in performance in private sector organisations is reported to be scarce.
The majority of reported examples of value creation emerging from a BSC implementation have only been demonstrated in the anecdotal work of Kaplan and Norton. Hard evidence in support of the BSC comes from the work of Geuser et al. (2009). From an empirical survey study of 76 business units, it was found that the BSC did have a positive effect on organisational performance. This was evidenced by an improved translation of strategy into operational terms, enhancing the process so that it is continuous, and ensuring a greater alignment between the processes, services, competencies and units of an organisation. Another example of BSC success comes from a case study analysis by Malina and Selto (2001) of an international manufacturing firm, who found that a BSC implementation resulted in an increased ability to control corporate strategy. Although this BSC example was labelled a success by the authors of the study, there was no comparison carried out on the performance of the company prior to the implementation.

Whilst the typical benchmarking indicators of BSC success ex-post in the private sector such as ROS or ROA (Hendricks et al., 2004), cannot be used in NPGO’s, evidence exists in the literature showing tangible benefits of an implementation in a diverse range of NPGO’s. Inamdar and Kaplan (2002) interviewed nine hospital managers who had implemented a BSC in their organisations. They found a measurable performance improvement in the hospitals in terms of the competitive market positioning, customer satisfaction, and financial results. Greilling (2010) reports on a study of 20 German non-profit organisations in the field of social services who had implemented a BSC. This study asked managers what they felt the perceived benefits to the organisation from a BSC implementation were. Perceptions from study participants were that the scorecards led to an increased ability to clarify and communicate strategy, an improvement in the alignment of strategic objectives with actions, and that they provided a base for building up a performance measurement system.

The use of BSCs in police forces worldwide has also been widely reported[1], however, the evidence relating to tangible benefits from scorecards is limited. A long-term implementation of BSCs in the Dumfries and Galloway Constabulary (Wisniewski and Dickson, 2001; HMIC for Scotland, 2007) showed evidence of an increased ability to assess strategic and operational performance despite a number of issues relating to the implementation procedures such as clashes with police culture. Silk (1998) also reported on the use of an automated BSC system in the West Mercia Constabulary, which showed that users of the system could more easily relate to how personal targets fed into overall force targets, therefore enabling a greater understanding of the contributions required by individuals, whilst at the same time showing potential for reducing costs.

It is surprising that considering the large body of literature regarding the BSC, there is relatively little evidence showing whether an implementation of the BSC leads to an increase in performance. Whilst some studies (Silk, 1998; Wisniewski and Dickson, 2001; Geuser et al., 2009) have shown that managers attribute improved elements of performance to the BSC, with the exception of Neely et al. (2004), there have been very few studies examining the implications of a BSC implementation ex-ante and ex-post. Further research in this area is recommended to fully understand the performance effects of a BSC implementation (Geuser et al., 2009; Burkert et al., 2010). However, in order to undertake such an exercise, it is clear that the specific BSC being implemented would need to be described. If not, it would be difficult to assess whether a failure to show improvements in performance was due to the implementation strategy, or perhaps stemmed from the choice of an inappropriate BSC.
Criticisms of the BSC

As shown in the previous section, whilst the BSC has, on occasion, proven successful in assisting companies with their strategic and PMSs, it has also attracted a certain degree of criticism. Despite these criticisms, or perhaps because of them, the BSC has continued to evolve over the past two decades since its inception. Whilst some of these criticisms can be considered general to the whole range of BSCs, other criticisms can be directly assigned to a specific generation or version of scorecard.

One of the most fundamental criticisms of the BSC relates to the causality linkages within the scorecard Norreklit (2000, 2003). This criticism is complex in nature but revolves around three principal issues, the lack of a time dimension, the lack of clarity regarding the interrelationship between the different perspectives in a score case and the lack of evidence relating to causality of the measures to improve performance. The first criticism is the lack of an adequate time dimension to the scorecard As a BSC has an inherent time lag between a propose change and any result being seen in performance, the lack of a time dimension means that it is difficult to demonstrate the direct link between cause and effect. Although targets and the time taken to achieve objectives is discussed from the BSC 1.1 onwards, and the concept of destination statements (which consider the long-term objectives of an organisation) are introduced in BSC 3.0, this issue is still not resolved as neither of these improvements consider the short-term lag problem. The second of Nørreklit’s criticism regards the proposed relationships between perspectives. She argues that as the relationships described between the four perspectives are ambiguous, it cannot be proven that improvements in one perspective can lead to improvements in another. The final criticism follows on from this and is related to the inter-dependencies between the four perspectives. The validity of all of the discussed generations of the BSC as a performance measurement model relies, in part, on a cause and effect relationship between the different elements. Whether this is between perspectives, measures or strategic objectives, this concept is key to the scorecard approach to PMM. Norreklit argues that this assumption is flawed, due to the inability of these causal relationships to be proven either empirically or logically. She claims that rather than a causal relationship existing in the model, the relationship is more one of interdependence of the perspectives in a logical relationship; therefore, to infer unidirectional influences between the perspectives is incorrect. Moreover, Nørreklit (2003) challenges the concept of the BSC as such. She raises the point that the BSC did not became such a popular management tool because of its convincing concepts, but rather because of the persuasive style the book is written in. She describes the arguments proposed “unconvincing, untenable and unsound” (Nørreklit, 2003, p. 610).

Neely et al. (1995) also criticise the BSC primarily based on the lack of a competitive dimension. They propose that a manager creating measures based solely on the four perspectives of Financial, Customer, Internal Business, and Innovation and Learning, would not be able to answer the key question of what the organisations competitors are doing. Whilst this holds true for un-modified scorecards, it is a moot point for scorecards such as NPGO BSCs.

Neely (2002) claims that the failure of the BSC to consider more dimensions such as a competitive perspective, HR/employee satisfaction perspective, supplier performance perspective, product/service quality perspective, and environmental/community perspective, limits the comprehensiveness of the model because not all measures can be included. Brignall (2002) also claims that there should be a Social and Environmental Aspect (SEA) to the BSC. Second-generation BSCs resolve this issue...
to some extent as they include the addition of both tangible, organisational assets, and intangible assets into the scorecard mix. However, as the BSC has evolved, it has moved away somewhat from the consideration of these types of measures; even the third-generation scorecard does not consider all of the above discussed perspectives, as the model itself has changed to become more of a strategic management system, rather than a performance measurement tool.

An overarching criticism of the BSC from Nørreklit (2000) is the lack of a solid theoretical foundation to the scorecard approach. She discusses how this lack of theoretical underpinning is reflected in the weaknesses of the causal linkages, and recommends that further discussion should take place about the theoretical underpinnings and the research methodologies concerning the BSC.

From a more practical perspective, Lipe and Salterio (2000) examine the problems that can occur during BSC implementation. In an empirical investigation using MBA students, they show how unique measures in a business unit level BSC can be underweighted by managers designing them. They claim that this occurs due to the agency effect of measures that are not being used in the evaluation or reward of managers affecting their decision making. In an extension of this work, Lipe and Salterio (2002) found that by arranging measures that were to be used in performance evaluation in a BSC format, managerial judgement could be affected due to the subconscious links between measures made by the managers.

Kasurinen (2002) criticises the BSC concept due its lack of focus on change management and implementation. He claims that Kaplan and Norton do not foresee any sort of analyses of potential barriers; therefore, structural barriers are not fully understood when they are encountered in BSC implementations. He concludes that because of this criticism, concepts such as the BSC are very likely to be pushed in the wrong direction due to the number of people involved at all levels of the organisation. This criticism unveils the problems of implementing the BSC and perhaps provides an explanation for the general lack of empirical evidence in BSC studies.

Methodology
As part of a detailed literature search to track the development of BSCs, it was recognised that difficulties were arising in relation to describing which version of a scorecard was being considered. Whilst there were many papers which discussed the application of the BSC in various settings, a comparatively small number were identified which actually developed the concept of the BSC. However, in comparing the descriptions of the BSCs discussed in the “application” literature, it soon became clear that the descriptions of the particular version BSCs being used were either unclear, or markedly different from one another, in line with the changes described in the “development” literature. This opaqueness demonstrated the need for clarity when authors used the term “Balanced Scorecard”. In order to understand the context, it was proposed that a standard way of describing the different iterations was attempted in order to achieve this clarity. The classification system proposed was one of Version Control (VC).

The concept of VC, also known as Revision Control (RC), is used to express the intensity of change that has emerged with each new development of the BSC. Whilst the concept of VC has been used widely since the 1980s in the fields of software design (Tichy, 1985) and engineering (Beech and Mahbod, 1988) it is generally unused in management literature. In simple terms: “Version Control is the task of keeping [software] systems consisting of many versions and configurations well organised”
It uses a basic numbering system to indicate changes to the original software/design and ranges from a simple, single-level system (Version 1, Version 2, Version 3, etc.) to a more complex, multi-level system (Version 1.0, Version 1.0.1, Version 1.1, etc.), which allows more minor changes to be indicated.

The authors applied this concept to BSCs by examining the minor and major changes proposed to the BSC in the literature over time. This allowed the different versions of scorecards to be identified and each individual iteration allocated a different version number. Whilst undertaking this process, the benefits of VC were clear and allowed the changes made to the original “baseline” scorecard over the years to be clearly identified and categorised. Were this model more widely adopted it would allow BSC applications to be more easily assessed both internally and externally.

In order to categorise the different versions of the BSC, the extant literature relating to the BSC was divided into “application literature” and “development literature” as discussed above. The development literature (shown in Table I) was ordered from oldest to most recent by their publication dates and the process of categorisation was initiated. The first published instance of the BSC in the literature (described in Kaplan and Norton, 1992) was assigned the revision number “1.0”. In order to avoid the classification becoming overly complex it was agreed to use only three levels; indicating major, significant, and minimal changes (x.x.x).

Having agreed the numbering system, the development literature post “Balanced Scorecard 1.0” was then assessed, and from this, the key differences between each different version were identified. Where a minor difference from the previous version was identified, such as the addition of a new component that did not alter the previous version in any significant manner, the version number would be changed at the third level. Where a more significant change was identified that changed the scorecard in a more substantial manner, the version number would be changed at the second level. Finally, if the change identified was so large in that there was a modification to the entire concept, the version would be changed at the first level, and a new “generation” of scorecard was categorised.

**Results**

Table I shows the results of how the BSC has evolved over time, and how each revision can be identified based upon its defining characteristics.

**Discussion**

The analysis of the studies shown in the table above proved to be very illuminating concerning the identification of different versions of the scorecard concept. During its development, the scorecard can be shown to have three major generations, as well as a number of minor versions (or revisions), which developed as academics and practitioners altered the BSC to suit their own purposes. The first two generations of the BSC were conceptualised by Kaplan and Norton and remain the most cited and discussed versions of the scorecard: when case-studies of BSC implementations are mentioned, it is usually first or second-generation scorecard (or variations of) that are used. By acknowledging the developments in the field from other academics and practitioners, Cobbold and Lawrie (2002) (along with Speckbacher et al., 2003) discuss the development of the BSC and highlight a third generation of the BSC. For the purpose of this work, the three generations of the BSC as identified by Cobbold and Lawrie (2002) have been sub-divided and expanded upon in order to take into account the additional theoretical and practical work undertaken by Kaplan and Norton (2004a) and others since 2002.
<table>
<thead>
<tr>
<th>First-generation balanced scorecard</th>
<th>Second-generation balanced scorecard</th>
<th>Third-generation balanced scorecard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developed in […]</td>
<td>Kaplan and Norton (1992)</td>
<td>Kaplan and Norton (2004a, b)</td>
</tr>
<tr>
<td></td>
<td>Kaplan and Norton (1996b)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kaplan and Norton (2000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kaplan and Norton (2004a, b)</td>
<td></td>
</tr>
<tr>
<td>Interactions between four different perspectives of organisational performance</td>
<td>Shows an early “how to” guide for building a BSC</td>
<td>How intangible assets can determine the performance of critical internal processes</td>
</tr>
<tr>
<td>Number of measures balanced</td>
<td>Focus on strategy: the four processes of managing strategy</td>
<td>More focus on strategic linkage model, less on specific measures</td>
</tr>
<tr>
<td></td>
<td>Introduction of strategic objectives</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Strategic linkages between measures</td>
<td></td>
</tr>
</tbody>
</table>

Table I. The evolution of balanced scorecards
The following sections describe the identified versions of the scorecard along with the key components that conceptualise them.

**First-generation scorecards**

*BSC 1.0*

The original example of a first-generation BSC by Kaplan and Norton (1992) is shown in Figure 1. Kaplan and Norton (1992) claim that the links between the four perspectives highlight the need for managers to consider all of the four perspectives as one whole.

In the initial article, Kaplan and Norton (1992) suggested that for clarity, the number of measures contained within the scorecard itself should be limited in number, and focus on the measures that are most critical to the organisation.

The concept of setting a time-period for companies to meet the goals and objectives set down in the four perspectives is only touched upon very briefly and is not a major part of the composition of first-generation scorecards.

The four different perspectives described in the BSC are as follows:

**Customer perspective**

This perspective requires the organisation to consider what the customer expects or needs and to set performance measures that ensure that the company is neither under nor over performing. According to Kaplan and Norton (1992), customers’ concerns generally fall into the four main categories of time, quality, performance and service, and cost. By using this perspective correctly, managers can effectively match their performance to the expectations of the customer.

**Internal business perspective**

This perspective requires the organisation to focus on the effectiveness of the internal processes and to use only those indicators that focus on activities which have the greatest impact on the operation of the business. These measures could include product cycle time, productivity, cost, and time to market amongst others. However, these measures must relate activities that can be influenced by the employees or managers.

**Innovation and learning perspective**

This perspective requires the organisation to consider how well it can develop and improve on its performance and how good it is at adapting to change, given the ever-increasing external pressures of global competition. The measures tend to focus on employee satisfaction and development, and the monitoring of new products or services with the emphasis being placed on innovation.

![Figure 1. The Balanced Scorecard 1.0](image-url)
Financial perspective
The final perspective is one that traditionally has been the main measure for organisations, that of financial performance. The other three perspectives in the BSC can be considered the drivers of future performance, whereas the financial perspective can be considered as a check of past performance, allowing managers and executive to consider the bottom-line impact of any changes made Silk (1998). Kaplan and Norton (1992) note that financial indicators remind executives that achieving the objectives laid out in the other perspectives only benefit the company when translated into improved sales or market share, reduced operating expenses or a higher asset turnover. Ideally, companies should specify how the improvements are going to lead to changes in financial performance.

Balance in the BSC
In the BSC 1.0, Kaplan and Norton (1992) proposed that the number of measures within each perspective used in a scorecard should be the same. The idea of this was to force managers to move away from the traditional, financially focused view of performance measurement towards a broader view on performance management; encouraging managers to consider the (equally important) non-financial aspects of the business as well as the financial aspects.

In more recent versions of the scorecard approach, there is less of a focus on the exact balance between measures from different perspectives, and more of an emphasis on the links between the desired performance outcomes and the drivers necessary to achieve these outcomes; an idea only briefly touched upon in the original article. Art Schneiderman, who was one of the creators of the concept that was to become the BSC, claims that balance in a scorecard can actually be harmful, and that measures used should mainly consist of non-financial, short term, leading measures (Schneiderman, 2001).

BSC 1.0.1, 1.1, and 1.2
The basic principle of this four-box approach to performance measurement has not changed throughout the evolution of the scorecard; however, as the concept has grown and changed, more attention has been focused on how the BSC can be used as more of a strategic tool, enabling the transition from performance measurement tool to PMS. This led to the next set of minor iterations as described below.

The BSC as a strategic management tool
As the concept of the BSC developed, Kaplan and Norton began to propose the BSC as a framework that should sit at the heart of a company’s strategic management system. The concept of linking the measures chosen in the BSC to the overall strategy was proposed by Kaplan and Norton (1993) as the first step in implementing the BSC. To do this, the corporate goals are broken down to the strategic business-unit level and then translated into strategic business-unit objectives, where each strategic objective has an expected connection to the business unit and corporate financial results. These more precise strategic objectives are subsequently set as targets that could be achieved by certain actions taken (Kaplan and Norton, 1993). Thus, each strategic objective contributes logically to achieving the overall corporate strategy. Figure 2 illustrates the concept of the BSC 1.0.1, whose major characteristic is the translation of corporate strategy into simple business-unit level objectives and measures. As this change does not affect the BSC in any significant manner, the change is identified as a minimal one.
The idea of translating the corporate strategy was further developed by the introduction of four processes that Kaplan and Norton (1996a) articulated for managing strategy: first, “Translating the Vision”; second, “Communicating and Linking”; third, “Business Planning”; and fourth, “Feedback and Learning”. These four processes aimed to provide a framework for companies to develop their ability to link long-term strategic objectives with short-term actions in an iterative manner, with the BSC sitting at the heart of this process, shown in Figure 3.

Source: Adapted from Kaplan and Norton (1996a, p. 77)
As such, the four identified processes revolve around the perspectives of the BSC and are interrelated with each perspective. Thus, a continuous cycle of adaptation and learning is generated that results in adjusted strategic objectives and changing measures, while constantly optimising the business unit levels’ performances. As this does not change the underlying principles, it is a substantial but not major change.

Kaplan and Norton (1996a) claim that the BSC enables strategic learning in an organisation by supplying three key elements. They describe these as articulating the company’s vision, supplying a strategic feedback system, and facilitating a strategy review. By sitting at the heart of the strategic management system, the authors claim that: “The BSC provides a framework for managing the implementation of strategy while also allowing the strategy itself to evolve in response to changes in the company’s competitive, market, and technological environments” (Kaplan and Norton, 1996a, p. 85).

This change from a performance measurement system to a strategic management tool sets the scene for the future development of the BSC towards the second-generation as discussed later in this paper.

**Introduction of targets**

The major difference between the original BSC (Figure 1) and the second substantial revision (Figure 4) is the addition of targets and initiatives, which reflects the change from a theoretical performance measurement tool to a more practical approach to performance management.

The illustration above shows how the BSC has become a technical tool with Version 1.2 allowing managers to include and make use of more detailed information on objectives, measures, and initiated actions. Thus, each strategic objective can be targeted more precisely and consequently is more likely to be fulfilled. The additional information will also give managers a “dashboard-overview” at any time to derive the current status of strategic objectives and actions taken to achieve those.

**Causality and strategic linkages**

A further change to the scorecard concept in this iteration of the BSC is the focus on the scorecard as being more of a top-down reflection of a company’s strategy. This is enabled by linking the chosen performance measures to the overall strategic objectives of a company as shown in Figure 4.

---

**Source:** Adapted from Kaplan and Norton (1996a, p. 76)
Although the idea of links between the different perspectives had been introduced in the Balanced Scored 1.0, later writing by Kaplan and Norton (1996a) introduced the idea of strategic linkages between the measures themselves across different perspectives.

Further developing this idea of causality was the concept of linkages between the strategic objectives of the organisation, rather than the measures (Kaplan and Norton, 1996b). This concept was the key change between BSC 1.1 and 1.2 and laid the foundations for the next generation of BSCs.

Bukh and Malmi (2005) argue that the cause and effect principle is one of the core features of a BSC but note the criticism that there is no evidence of a relationship between the different measurements. However, they suggest that one of the issues is that Kaplan and Norton do not define what they mean by causality and using a looser definition does allow for cause and effect. As such, a BSC and strategy maps is developed using the senior managers assumptions as to what drives a company’s performance and there should be no expectation that such assumption are based on hard evidence or are completely accurate. Strategy maps can only be based on a hypothesis and were it not so, management would be a more straightforward process. This focus on the role of causality is important, given that the strategy maps move the BSC from a performance measurement tool to a performance management tool. However, recognising the limitations of the strategy approach is important, because whilst the cause and effect relationships may not be perfect, it does highlight that the development of a BSC should be an iterative process; one that allows an organisation to gain a better understanding of what are the key drivers of performance. However, Ittner and Larcker (2003) argue that one of the difficulties with this approach is that managers do not actually focus on what is important, but merely select the measures that are easiest to capture. In this case, the lack of causality is based on inappropriate measures rather than any issues with the theoretical underpinning of the BSC.

Othman (2006) examined the causal model of strategy. He analysed the effects that having a causal model of strategy has on the development and implementation of a BSC and found that ultimately, the development of the causal model benefits the Scorecard and contributes to a successful implementation.

Second-generation BSCs

**BSC 2.0**

The major change that heralded the evolution of the first-generation BSC into the second-generation BSC was the introduction of the Strategy Map (Kaplan and Norton, 2000).

A strategy map is a tool built upon the BSC, which provides a visual framework for the corporate objectives within the four key perspectives of the BSC. They were designed with the aim of bringing into focus the line of sight between corporate strategy and the everyday actions of employees, therefore enhancing collaboration and coordination (Kaplan and Norton, 2000).

Strategy maps developed the idea of strategic linkages by looking not only at the strategic linkages between measures (Kaplan and Norton, 1996a), but also examining the linkages between strategic objectives (Kaplan and Norton, 1996b) and developing this idea into a conceptual framework. Strategy maps made the concept of strategic linkages much more useful, by examining the causal links between them. The aim was to allow managers to identify what changes had led to the desired organisational outcomes. From a broader, organisational viewpoint, Strategy maps allow managers to
see how the company can convert its “raw materials” such as initiatives, resources, and intangible assets, into tangible outcomes.

The key point to note in this iteration of the BSC (Figure 5) is that it is designed to be developed from the top-down, starting at the overall objective of the organisation, which is used to create the measures and goals in the rest of the scorecard, working down to the individual perspectives and their measures.

**BSC 2.1**

In the BSC 2.0, the idea of intangible assets was mentioned in the context of how Strategy maps allow managers to see how the assets of a company can be translated into tangible outcome. The BSC 2.1 develops this idea, by examining how the intangible assets fit into the overall Strategy map.

These intangible assets are described by Kaplan and Norton (2004a, p. 55) as follows:

- **human capital**: the skills, talent and knowledge that a company’s employees possess;
- **information capital**: the company’s databases, information systems, networks and technology infrastructure; and
- **organisation capital**: the company’s culture, its leadership, how aligned its people are with its strategic goals, and employees’ ability to share knowledge.

The key difference between BSC 2.0 and BSC 2.1 is the focus on the foundations of the Strategy map in a bottom-up analysis of strategic performance management. By doing this, Kaplan and Norton (2004a) claim that it can easily be shown how intangible assets determine the performance of the internal organisational processes. These causal links can then be used to see how the intangible assets relate to the company’s strategy and overall performance. Kaplan and Norton (2004a) claim that
it is then possible to align the intangible assets with the strategy and measure the contribution of the assets to it.

A direct benefit of being able to measure the contribution of an intangible asset to a company's overall strategy is that resources can be focused on the areas that need the most attention. Either this could be where the intangible assets in question are not providing enough benefit to the strategy and need to be improved, or it could be simply ensuring that a critically important area of strategy has the best assets assigned to it.

It is important to note that the design elements that make up the second-generation BSC are reported to be the “mainstream” definition of the BSC throughout a range of both academic and practitioner texts (Cobbold and Lawrie, 2002).

The BSC 2.1 is shown in Figure 6.

Whilst the BSC 2.1 represents the most recent version of the scorecard as conceptualised by Kaplan and Norton, as interest in the scorecard grew, practitioners and consultants have become interested in adapting it for their own purpose, resulting in new iterations.

The third generation of the BSC is one such adaption. First identified in 2002 by Cobbold and Lawrie, it was formally introduced as an academic concept by them in 2002.

**BSC 3.0**

The key concept that defines a third-generation BSC is the “Destination Statement” (Cobbold and Lawrie, 2002). The destination statement describes, in one or two pages, what the organisation should look like at an agreed future date (Shulver et al., 2000; Lawrie and Cobbold, 2004). These destination statements were initially created at the end of the design process of the BSC by challenging managers involved in the process to imagine the impact of the strategic objectives on the organisation. It was soon

---

**Figure 6.**
The Balanced Scorecard 2.1
realised that the destination statements made it easier for managers to understand the design process, which meant strategic objectives and management consensus could be achieved more quickly (Shulver et al., 2000; Lawrie et al., 2004).

The aim of the destination statement was to identify inconsistencies in the measures chosen for the scorecard by requiring managers to consider the causality between the measures and targets chosen, and the overall objectives, thus ensuring a more suitable choice of measures and objectives (Lawrie and Cobbold, 2004).

An example of a third-generation BSC is shown in Figure 7.

**BSC 3.1**

BSC 3.1 (shown in Figure 8) differs from the BSC 3.0 as it reduces the number of perspectives considered in its approach down to just two measures. The “outcome” perspective replaces the financial and customer perspectives, whilst the “activity” perspective replaces the learning and growth, and internal business process perspectives (Lawrie et al., 2004).

As a result, Lawrie and Cobbold (2004) claim that the careful use of a destination statement can encourage the consideration of non-financial measures equally as well as using four perspectives in the BSC. This eliminates the need for public sector managers to be “bogged down” with the categorisation of measures; as long as the correct measures are chosen for measurement, and the causal links are shown, there is not a problem with “missing” perspectives (Barney et al., 2004).

Case studies of the third-generation BSC seem to show an improvement over previous generations in the time taken for management teams to achieve consensus in the design of the scorecard (Shulver et al., 2000; Shulver and Antarkar, 2001; Lawrie et al., 2004). As less time is spent in the design process due to the use of the destination statement, the ability of an organisation to support flexible PMSs and develop their strategy in a more iterative manner is increased.
Conclusions
BSCs have become a popular tool within performance management since the first Kaplan and Norton paper was published in 1992. Over time, the concept has evolved from a simple performance measurement instrument, to a complex, multi-faceted PMS. As organisations began to use Balance Scorecards, they realised the shortcomings of the existing framework and sought to adjust the scorecard to better suit their own purposes. In response, practitioners and academics began developing new versions of the scorecard. This pattern of development is in line with the process described by Kennerly and Neely (2002) who discussed the factors that may affect the evolution of performance measurement systems within an organisation and state that for a PMS to evolve, a trigger must occur that causes the organisation to undergo a process of reflection, modification, and deployment of the newly altered PMS. In addition, Kaplan and Norton (2004a) claim that the BSC movement has encouraged organisations to challenge the traditional view of performance measurement and move towards a process where they are able to measure whatever they want to measure, including intangible, “softer” assets.

Whilst BSC implementations have been labelled successes, it is important to consider what is meant by this. In theory, introducing a PMS such as the BSC should improve the measurable performance of a firm such as by improving its ability to manage its assets, whilst at the same time allowing costs to be reduced through an increased understanding of the business environment in which it is operating. Unfortunately, there is very limited evidence that the BSC can achieve this, especially as the time lags between changes and performance improvements are not well defined as discussed earlier in the paper. Instead, the successes that most case studies talk about revolve around less concrete measures of performance such as an increased ability to manage strategy, understand the organisation, or assess performance rather than showing a multi-dimensional measurable increase in performance.

However, as discussed in the opening paragraph, the BSC has been improved and refined over the last 20 years and the aim of these iterations has been to produce a BSC
that is most effective in improving performance. Given that the BSC was initially designed as a performance measurement system in the first instance, it is not unsurprising that it did not necessarily produce improved performance. It was only when the scorecard began to develop into a more focused PMS that measuring success via improved performance became an indicator of successful implementation.

As such trying to measure the success of these implementations is further complicated by a lack of clarity in papers as to which iteration of the scorecard has been implemented or assessed. This paper attempts to provide a simple taxonomy for the different forms of the BSC, through a process of identifying and labelling the major, significant, and minor changes that have occurred throughout the evolution of the BSC. This is useful as it allows a more nuanced analysis of the BSC as a tool for managing performance by adding precision to any critique of a scorecard implementation or the usefulness of BSC in general. It is perhaps assumed that the latest version of the BSC will always be used, hence why there has been no previous attempt to classify the different strands. However, looking to the literature it is clear that this is not the case (Amando et al., 2011; Taylor and Baines, 2012). This is critically important for both academics and practitioners alike, as without the information regarding which iteration of the BSC is being considered, any discussion is immediately stifled due to a lack of information.

Having reviewed the development literature of the BSC, the authors propose three generations of the BSC split into eight separate versions. The first generation of the BSC is split into four different versions; the first laying the groundwork for the scorecard concept, the second beginning to focus on strategy, the third introducing specific targets, developing the strategy and introducing causality and the fourth highlighting the role the scorecard plays as part of a PMS. The second generation of the BSC is characterised by the introduction of strategy maps in the first version, and the further development of this concept alongside the change to a bottom-up, ongoing approach to the scorecard concept in the second version. The third, and final generation of the BSC identified in the development literature has two versions, with the first characterised by the addition of the “destination statement” to the BSC and a greater focus on the strategic linkage model whilst the second version simplifies the scorecard by removing many of the perspectives that are considered in earlier versions.

It is hoped that this simple classification will prove useful to practitioners, who will be able to describe which generation they would suggest using to managers, and academics in being able to more precisely describe the scorecard that they are analysing. The adoption of revision/VC for the purposes of classification has been shown to be a useful tool in enabling a more structured and detailed understanding of BSCs and it is hoped that the methodology described in this paper can be adopted more widely in the academic community. As it is expected that BSC will continue to be developed and refined in the future, it is hoped that the framework could be used for future versions as well.

It is clear that a large number of organisations who have implemented a BSC have done so without considering the full spectrum of scorecards that have been identified in this paper. It is therefore hoped that these classifications will also be useful to managers seeking to implement a BSC independently in their own organisations by providing them with the knowledge as to what possibilities exist when it comes to the BSC. It is important to note that later versions of the BSC are not necessarily “better”, than earlier versions; they are simply different. Whilst a BSC 2.1 may be suitable for one organisation, a simpler BSC 1.0.1 might be more suitable for another. Although BSC practitioners or academics
may recommend one type of BSC, it is up to the organisations in question to ensure that the version of the scorecard that is implemented is the most suitable for them.

Therefore, we can conclude by taking the stance that the BSC is a powerful tool, that when applied in an appropriate manner may have significant benefits for the organisation in question. However, it must be understood that the BSC cannot be thought of as a miracle tool that will somehow improve the performance of a struggling firm. It must instead be considered as one element in a firm’s arsenal, which can help in effectively managing its performance by providing clearer answers around the question “How are we performing?” The particular version of the BSC chosen for implementation must be carefully matched to the needs of the organisation in order to ensure success, and the classification system provided in this paper should prove a useful tool in allowing managers, practitioners, and academics to achieve this.

Note

References


About the authors

Mike Perkins graduated from the University of York with a degree in management, before acting as an Independent Management Consultant for a number of organisations. After a brief stint working for the North Yorkshire Police, he started a PhD at the University of York examining how specific police KPIs can be improved in order to increase the performance of police forces. After completing his PhD he started a Lecturer post with the British University Vietnam, Hanoi. Mike Perkins is the corresponding author and can be contacted at: mgperkins@gmail.com

Anna Grey graduated with a degree in Psychology from the University of York and has since worked in Higher Education Management at various UK Institutions. Her current role focuses on research management and performance measurement. She successfully completed an MSc in management part time and is now registered for a part time PhD focusing on research performance measurement.

Helge Remmers graduated with a degree in Aviation Management from the International University of Applied Sciences, Bad Honnef-Bonn before taking on a role in DHL Global Forwarding (Basel) where he was involved in numerous projects involving the implementation of performance measurement systems. He then completed a research role at the University of York examining the use of performance measurement systems in financial institutions before returning to DHL.

To purchase reprints of this article please e-mail: reprints@emeraldinsight.com
Or visit our web site for further details: www.emeraldinsight.com/reprints
This article has been cited by:


8. Dag Madsen, Kåre Slåtten. 2015. The Balanced Scorecard: Fashion or Virus?. *Administrative Sciences* 5:2, 90-124. [Crossref]


