



Advanced Performance Institute

Management Case Study

Measuring and Managing Performance in the Royal Air Force



For more information please see: www.ap-institute.com

Measuring and Managing Performance in the Royal Air Force

By

Bernard Marr* and Ian Shore

Abstract: *This management case study illustrates the implementation of a strategic performance management system in the Royal Air Force. It outlines how a scorecard reporting approach was designed and implemented, in particular, it highlights the importance of integrating performance management with risk management, it discusses the problem of objective versus subjective measurement and it outlines the challenges of automating the system using a custom-built performance management software application.*

Release Version: Sep 2007

** corresponding author*

Bernard Marr is the Chief Executive and Director of Research at the Advanced Performance Institute. E-mail: bernard.marr@ap-institute.com

Ian Shore is a Wing Commander at the Royal Air Force Air Command.

Disclaimer: Any views expressed herein are personal and are not necessarily the views of the Ministry of Defence.

The **Advanced Performance Institute** is a world-leading independent research and advisory organisation specialising in organisational performance. It provides expert knowledge, research, consulting and training to performance orientated companies, governments and not-for-profit organisations across the globe. For more information please visit: www.ap-institute.com

How to **reference** this case study:

Marr, B. and Shore, I. (2007) Measuring and Managing Performance in the Royal Air Force, Management Case Study, The Advanced Performance Institute.

Measuring and Managing Performance in the Royal Air Force

Introduction

The United Kingdom's war-fighting capability in the air environment is provided by the Royal Air Force (RAF), one of the nation's three Armed Services. The Royal Air Force's aim is to achieve precise effect at range and in time when called upon to do so. Air Command¹ is the front line of the Royal Air Force. It is charged with meeting the Royal Air Force's aim by training the personnel and operating the cutting edge aircraft, weapons and supporting equipment, supplied by the Defence Equipment and Supply Organisation.

Air Command has 42,000 service personnel, 9,800 civilian staff and more than 1000 aircraft located on over 70 stations or units. Air Command supports operations in the Gulf region, Kosovo and Afghanistan as well as maintaining a RAF presence in Cyprus, Gibraltar, Ascension Island and the Falkland Islands. Its key peacetime responsibility is to maintain the required readiness levels of its forces (e.g. the Harrier, C17 Globemaster, and Sentinel aircraft and their crews) in support of the requirement to operate as an expeditionary air force.

While this case study illustrates the implementation of a performance measurement and management system, it particularly highlights the importance of integrating performance management with risk management, discusses the problem of objective measurement and reviews the challenges of automating the system using a custom-built software application.

Why Performance Measurement?

In 1968 the Fulton Committee on the Civil Service declared that "Accountable management means holding individuals and units responsible for performance measured as objectively as possible. Its achievement depends upon identifying or

¹ On 1 April 2007 the then two RAF Commands, Strike Command and Personnel and Training Command were amalgamated into a single Command, called Air Command

establishing accountable units within government departments – units where output can be measured as objectively as possible and where individuals can be held personally responsible for their performance.”

Expenditure on defence, funded by the taxpayer, is controlled by a process of delegated budgets. The Commander in Chief of Air Command is a delegated budget holder. The planning of expenditure is heavily influenced by the long lead times for equipment procurement, and for the implementation of new technology and procedures as well as the lag represented by the staff recruitment and training timescales (up to 3 years for an operational pilot). As a result commanders place a high premium on knowing the forecast performance levels as well as the operational risks alongside the planned costs of their forces. Furthermore, once budgets are set, senior commanders are tasked with achieving agreed levels of force readiness, thus they need to be informed about the latest levels of achievement compared against the agreed targets. In addition, given the interdependency between the front-line and the supporting organizations who supply equipment, front line commanders need to be able to measure whether their suppliers are meeting (and will be able to meet) their requirements and how any shortfall is going to impact on their ability to deliver the required outputs. Lastly, but of most importance, when called upon to respond to a crisis somewhere in the world, commanders need to know the state of readiness of their forces which they will need to deploy.

Deciding What to Measure

The first challenge was to establish what needed to be measured. Priority was placed on establishing the measurement of the Department's outputs. Unlike a commercial concern, where everything can be translated to a common currency ('the bottom line'), defining the outputs of an operational area within the Ministry of Defence has been a challenge. The final output is an achieved effect, which could be as simple as constraining the supply of munitions to an enemy's forces, which might be achieved by a highly-focused bombing mission on a single electricity sub-station supplying the factory. Measuring effects in war time is problematic enough, in the (very welcome!) vacuum of peacetime a proxy has to be sought. The most obvious proxy would be to measure the capabilities, for example the capability to drop a bomb within a defined area of a target. However, aircraft have multiple roles and capabilities can therefore be supplied in more than one method. So this

attractive option has proved very difficult to measure. The next nearest proxy has been to determine the numbers of aircraft and crews that are to be available against a range of readiness states, from immediate readiness, in increments out to a number of months before the given unit can be brought up to operational status.

While analysis of readiness by force was deemed essential, shortcomings within forces also need to be analysed against a common framework to reveal trends applicable to more than one force. A common framework was developed by which analysis is undertaken using eight perspectives – the so called Defence Lines of Development (DLOD) (see Figure 1). Thus, the Department's, and thus Air Command's, outputs are measured by the readiness of its Forces (called force elements at readiness, FE@R for short) analysed by DLOD.

FE@R	
Personnel	Organisation
Equipment	Training
Logistics	Infrastructure
Information	Doctrine and Concepts

Figure 1: Defence Lines of Development (DLOD)

The measurement of FE@R by DLOD, the Outputs, was considered essential. It was also recognised that the detailed analysis of the achievement of the output needed to be supported by an understanding of the contribution being made by both the Resources (eg. cash, staff, contractors, equipment) and Processes (eg. training, logistics, personnel management) to the delivery of the output. It was also clear that some of the Department's resources and processes went towards the procurement and introduction of new capabilities (e.g. Typhoon). Therefore, for a comprehensive view, in addition to measuring Outputs, it was necessary also to measure: Resources, Processes and the progress towards delivery of Future Capabilities.

Developing a Scorecard Reporting Framework

Reporting on a wide range of Outputs, Resources, Processes and Future Capabilities in a manner that assisted commanders to identify actions they needed to take was a significant challenge. The source data was extremely varied and needed to be

compared against the targets at a detailed level to reveal shortfalls. The nature of the data demanded that it was summarised before presentation to the Commander, at the same time, essential detail could not be lost as that could point to the most appropriate action to resolve the shortfall. To solve this conundrum the Department wanted to design a performance dashboard or scorecard. However, popular tools such as the Balanced Scorecard developed by Kaplan and Norton, which has a commercial business focus, did not fit the Department's not-for-profit model. Nonetheless, the concept of presenting data by traffic lights and being able to drill down layers of such traffic lights to the raw data was widely accepted as the best approach for informing commanders. The Department therefore adopted a scorecard approach using the four perspectives: Outputs, Resources, Processes/Enablers and Future Capabilities (see Figure 2). To give Commanders a 'look ahead' capability to address the long lead times involved in resolving adverse situations, the concept was soon extended to provide a view of the trends in the traffic lights over time – looking back over the past two reporting periods (quarterly) and forward for 4 years.

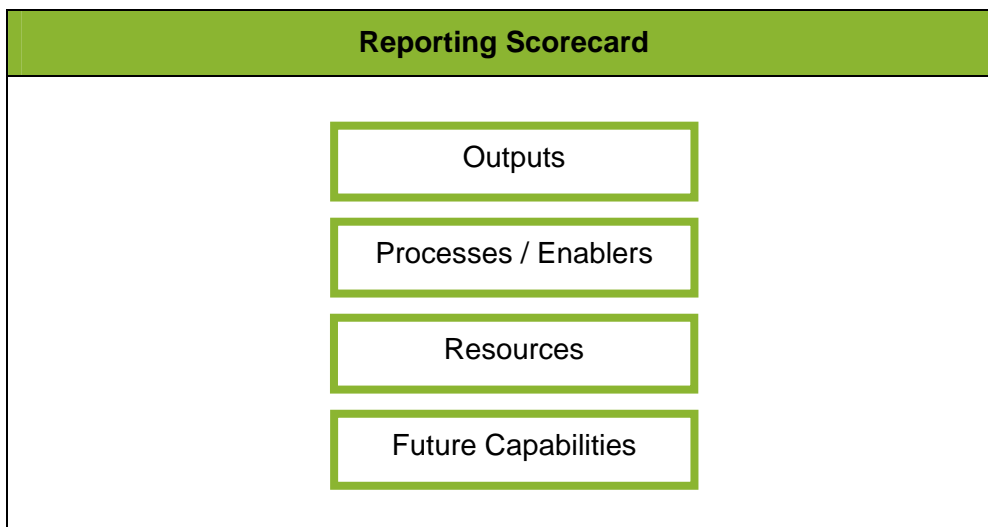


Figure 2: Reporting Scorecard

The Objectives and their Performance Indicators (PIs) that supported assessments under each of the four scorecard perspectives were developed over time, with each new year seeing a refinement of the previous year's set. These measures were then cascaded down the organisation hierarchy in Air Command (Command, Group, Station).

Assessing Performance – Balancing Objectivity and Subjectivity

Initially, all traffic lights were populated as subjective judgement of the subject matter experts (SMEs). While this was the easiest method of collecting the data, subsequent investigation of underlying reasons was hampered by the lack of objective detail. Thus, over time, targets for each FE@R were developed as extensive tables, with associated rules to define the levels of shortfall that would trigger different traffic lights. It was recognised that having only subjective assessments was sub-optimal but initially the appetite for a scientific approach to performance management was limited and many areas, classically an assessment of 'fighting spirit', were always seen as best handled purely subjectively.

For the steadily increasing numbers of performance indicators which were being objectively measured appropriate data sources had to be identified. Where it was possible already existing data sources were used. The number of individual performance indicators used to measure an Output averaged some 70. In order to present these to senior commanders in a format that was easy to read, these 70 performance indicators had to be grouped into a hierarchy. As a result each Output is measured against a standard set of 8 headings (Training, Equipment, Personnel, Infrastructure, Doctrine/Concepts, Organisation, Information and Logistics), each heading then being further sub-divided until the lowest level (ie. that at which the data is collected) is reached. This consolidation of traffic lights up hierarchies of performance indicators forms a key component of the process.

The need to consolidate a wide range of measures generated the requirement to be able to define how subordinate performance indicators should consolidate into the superior single performance indicators. A range of rules were developed to consolidate indicators (see Figure 3).

The provision of automated data feeds, traffic light generation and consolidated scorecards generated a concern that such a 'hard-wired' set of information would not reflect the true position of performance. Senior commanders place high value on the views of the 'commander on the ground'. These views are informed by a mix of

factual information, experience and judgement. While automating data capture would provide factual information, it would not always be comprehensive. Furthermore, the local commander's 'applied military judgement' (AMJ) was seen as an essential input into the reporting process. Thus the system was designed to provide both a pure view based on the factual data (known in the system as the Calculated (CALC) view) and then a second, parallel view that took into account the AMJ of the local commanders. This second subjective AMJ view also ensured that the senior commanders were not misled by traffic lights generated from flawed data or rules.

Consolidation Rules	
Best Case	upper level equals best colour of lower level
Worst Case	upper level equals worst colour of lower level
Average	upper level equals arithmetic average of lower level
Weighted Average	lower level given weights to reflect importance, upper level equals arithmetic weighted average
Conditional	predefined truth table defines upper colour to be used given each combination of lower level colours
Reassessment	lower levels ignored and upper level treated as base level performance indicators with own target and achievement data

Figure 3: Indicator Consolidation Rules

The system now had a set of performance indicators organised in different perspectives with both calculated and subjective assessments triggering traffic lights. However, in order to be useful it is critical that the key insights could be extracted. To aid the extraction of insights and to facilitate meaningful performance review meetings, a group of performance analysts was established. Dedicated analysts serve particular parts of the organisation and help them to make sense of the data. Their role is to analyse the reported performance data and highlight the key insights. This allows analysts to flag up any issues and propose questions that should be

discussed and answered in the quarterly Defence Management Board review meeting and its equivalent at lower organizational levels.

The Automation of Performance Reporting

Initially, scorecard reporting was done using MS PowerPoint, supported by MS Word-based reporting templates. This prevented any automation of the data collection and consolidation process and required high demand for human input to interpret and re-present data. This had the effect of constraining the extent to which the benefits of the reporting process could be delivered to subordinate levels in the organisation. To address these limitations, Air Command initiated a major project to define, develop and implement an IT system based on a relational database infrastructure.

The first key decision was that the system needed to bring together into a single database information on performance, risk and cost. The project objectives were to integrate Performance Management, Resource Management, Supplier Management and Risk Management (later Risk and Issue Management) across all levels of the command, providing a single version of the truth and utilising existing data sources as far as practical. This represented a significant challenge as each of the associated business processes were governed by different parts of the organisation. The benefits of an integrated approach were self-evident to many, but the decision did raise the stakes as the complexity of the scale of integration was significant.

An associated decision, which influenced the number of users and the geographic spread of the system, was to involve RAF stations in these integrated processes. Prior to automation stations had very little involvement in Performance Management, heavy involvement in Resource Management, but not in an integrated manner and the only formally documented risk management was generally of the non-operational sort.

It was clear from the outset that the complexities of reporting across multiple hierarchies, across changing timescales, using multiple targets and across a large number of users would require an 'industrial strength' database. Development was thus outside the scope of in-house capabilities. So it was decided to contract out the

work and as a result a bespoke software application (based on ORACLE) was developed by Fujitsu Services Ltd. This application was called SAPPHIRE.

Prior to automation the complexities of measuring using SMART² performance indicators had generally defeated the capabilities of the supporting systems. SAPPHIRE was designed from the start to allow for SMART performance indicators, automatically capturing data from source systems where practical. In recognition of the shortcomings of hardwiring such a reporting process, SAPPHIRE also recognised the need for the 'man on the ground' to provide his 'Applied Military Judgement', in effect a parallel set of traffic lights reflecting the commander's view of the situation notwithstanding the 'facts' being reported through the automated process. This decision was made in recognition of the increase in workload and complexity for software developers. Subsequent feedback from users has indicated that this was the most important decision made.

The considerable technical and user complexity of the system drove the decision to both develop and implement incrementally. This approach, which is in line with Ministry of Defence procurement best practice, allowed early benefits to be gained without the full level of investment. It also allowed cost control to ensure that sufficient value had been delivered before further investment was authorised. However, the overall scale of the project was recognised earlier on and a decision was taken to seek appropriate high level approval up front and to obtain the necessary funding approvals without under-stating the potential cost, in particular for the ongoing maintenance/upgrades and for the support staff. The scale of the project also triggered the decision to involve specialist support for project management alongside a properly resourced implementation team. Without these essential components the programme would not have made the required progress.

Today, the SAPPHIRE system provides automation to the functions of performance management, risk/issue management and supplier management. In due course it will also provide some support to resource management. Its key functions are to:

² The SMART abbreviation comes from goal setting theory which argues that any goal should be: Specific, Measurable, Achievable, Realistic and Timely

- Capture day-to-day achievement data from source systems to avoid inefficient dual-keying where possible
- Compare captured / entered achievement against pre-defined and stored targets to generate performance indicator traffic lights
- Consolidate performance indicator traffic lights to generate higher level traffic lights against scorecard objectives
- Record risks and relate them to scorecard objectives
- Present summarised data in traffic light form with commentaries to management supported by a drill down capability
- Store data for future use to allow trends to be identified
- Interface with external systems (Defence Balanced Scorecard Application and MS Excel)

Aligning Performance and Risk Management

Once the scorecard element was complete, the next step was to link Performance Management with Risk and Issue Management. Performance Management is about measuring, reporting and making decisions about achieving objectives. Risk management is about identifying and managing the risks to achieving the same objectives. Thus there is a common denominator – an organisation's objectives.

For example, key objectives are expressed as the delivery of the FE@Rs. Performance Management identifies shortfalls between the targets that have been set (e.g. 10 crew members at 7 days readiness) and the achievement (e.g. 8 crew members at 7 days readiness). The shortfall of 2 crew members, if it is certain it is an issue, if it is only probable it is a risk. The shortfall, which will be revealed by the Performance Management process, needs to be identified, assessed and managed as in all good risk management practices.

Performance Management also includes the forecasting of the position in the future, extrapolating from the known baseline of the current position. For such forecasting to be realistic it needs to take into account the known risks and issues and modify the extrapolation to reflect the assessed most likely impact of the combined effect of these risks and issues. Thus judgements being made within the performance management process cannot be separated from data available from the risk management process. They are inextricably linked.

The Royal Air Force has been doing Risk Management for many years. However, unlike a commercial company, the risk focus is not primarily on the financial 'bottom line' but rather on the delivery of war-fighting capability. Key risks for the Royal Air Force are those which impact on war-fighting capability, termed, loosely, 'operational risk'. Some of these are generated by deliberate decisions shaped by available funding. For example, to limit the level of spares purchased to avoid overspending a budget raises the risk that equipment will be unavailable due to unserviceability. Again providing a lower (cheaper) level of protection for personnel, risks increasing the danger to them in a hostile environment. Over time these risks accumulate and their combined effect may be greater than would be recognised when looking at the individual risks.

SAPPHIRE was designed to meet Commanders' need for greater visibility of their 'operational risks', and to maintain both an authoritative audit trail of reviews and the current state on ameliorating actions, all to aid their decision making. It also had to meet the corporate governance requirements, however these were considered to be secondary to the prime objective.

SAPPHIRE has therefore been designed to provide both performance and risk information in a manner that allows the performance assessments to be made in the light of the risks that have been recorded. The Risk and Issue Management implementation was an enduring success. After an initial 3-month pilot period on selected stations, the system was introduced Command-wide. To avoid the important being swamped by unhelpful detail, commanders were encouraged to concentrate on their top 5 to 10 concerns.

Addressing Key Implementation Challenges

As with any project of this scope, there are countless challenges and hurdles to overcome. Here, we highlight three major challenges: (a) Managing the implementation and the important role of a dedicated implementation team, (b) overcoming scepticism and (c) the significant task of collecting and integrating performance data.

The Role of a Dedicated Implementation Team

The scale of the programme, with over 1300 users on geographically dispersed sites across the UK, generated its own challenge as each required training, for many in two separate sessions one for Performance Management the other for Risk/Issue Management. There has also been a need to keep the community up-to-date with developments without swamping them with unnecessary detail. To ensure that the training was professional, a Training Needs Analysis was undertaken; a Computer Based Training Course was created and a range of courses and briefings were designed, and provided. The programme was also supported by a web site.

It was recognised at the start of the project that success would be dependent on creating an in-house team to manage the process from gestation to full operational capability. The team benefited from being able to build on some early work being undertaken on performance measurement and separately on the provision of a command-wide management information system. The team consisted of:

- A User Champion responsible for owning the vision, defining the requirements, and delivering the benefits.
- A Project Manager responsible for managing the development and implementation programme.
- An in-house operational researcher, responsible for defining and introducing an objective measurement process and optimising the use of information and knowledge.
- Specialists, who were students doing their 'year in industry' who formed the backbone of the help-desk and training capability.

A key role of the team was to maintain momentum during the long implementation period required due to the staged implementation approach. The team served as the focal point for all matters associated with automating the performance management and risk management processes. At times the team also had to establish the doctrine and concepts to support the business processes. In addition, it provided all the training, communications programme, developed and maintained the web-sites, manuals, Quick Reference Cards & computer based training packages, as well as the help desk.

Mention of the team would not be complete without covering the essential work provided by another party – the Tactical Communications Wing staff. They had the responsibility of providing the database administration services by maintaining the server and shared network infrastructure, managing the configuration of the software and setting up the data structures in the system to support the business requirement.

Overcoming Scepticism

Within the Ministry of Defence there is some scepticism about the science of management, which in some circles is seen as only helpful in a commercial environment, particularly for those involved in repetitive production line activities. The application of these scientific methods to war fighting and preparing for war are not always recognised as beneficial. The implementation of scorecards, and later SAPPHIRE, had therefore to overcome the drag caused by such scepticism. This implementation benefited from the support and direction from the Commander-in-Chief which provided the necessary 'senior management sponsorship' considered to be an essential component of a successful project. Without strong leadership from the top progress would have been impossible.

While top management focused on the benefits that could be gained from taking a more structured approach to management, middle management often had to be persuaded about the benefits to them. This proved the highest hurdle to be crossed. Furthermore, the hurdle kept re-appearing as the comparatively short posting cycles meant that convinced commanders were replaced by sceptical commanders, presenting a constant need to educate and persuade middle management of the benefits to them of the system.

The Command's stations are widely geographically spread, stretching from Morayshire in Northern Scotland to Cornwall in the far South West of England. As one of SAPPHIRE's aims was to enable reporting to be done by this front line the distances from the HQ in High Wycombe near London represented a significant challenge as station staffs needed to be trained, supported and updated. While some reliance was placed on web-sites and emails, it was recognised that face-to-face briefings and question and answer sessions were the only reliable method of communicating the key messages. This placed a high demand on the core team

who had to travel to each location to conduct briefing and training sessions, all designed to win hearts and minds.

Overcoming the scepticism and cynicism that was borne from seeing a range of initiatives that had fallen short on their promised benefits, combined with a general reluctance to accept the disciplines of public accountability was probably the biggest challenge of the entire implementation. In response the SAPPHIRE team sought to emphasise:

- There was no choice – accountability was a requirement of being the beneficiaries of taxpayers' munificence!
- Senior commanders wanted better information to improve their decision-making.
- By sharing data efficiency would improve through the reduction of decisions made on flawed assumptions.
- By contributing to the process they could ensure that their concerns were better recognised by managers.
- The most senior commander in their command chain was sponsoring the change.

The Challenge of Collecting and Integrating Data

Collecting indicators and automating the process posed a considerable challenge. The drive for SMART performance indicators led to the need to capture and consider data at a more detailed level. This gave the more junior levels of staff considerable concern about workload, with the distraction it created from other (more important in their view) tasks. This was particularly so when the junior users rarely saw personal benefit from their efforts.

A major challenge was the very complex nature of the data structures with dimensions of: Organisation, function, performance measure, capability, risk, reporting frequency, time, readiness state, security classification, user access, currency of data. The interactions between these various aspects of the data all needed to be determined, designed and tested. A very close working relationship was required between the experienced programmers of Fujitsu Services Ltd and the core SAPPHIRE team, built around with a comprehensive test regime involving both parties.

Once the system was installed the team were able to address the complex task of matching the source data captured from existing computer systems with the expected data structures within the SAPPHIRE system. This involved the creation of a number of rules for transforming the data, which required special technical as well as business knowledge. This work proved detailed and technically complex.

The PM reporting structures are based on the business units and their reporting chains up the organisational hierarchy. The links in these chains are re-aligned on a frequent basis in response to the formation of new units to deliver new capabilities (eg Typhoon sqns), closure of bases as the Ministry of Defence estate is rationalised, and re-brigading to more efficiently align related business units. This continuous 'shuffling of the pack' has placed a heavy demand on the database administration function, who have had to maintain a close co-ordination with the business units.

Final Remarks and a Look into the Future

The implementation of an integrated performance management and risk management approach has been a success and the aim of providing commanders with better information and insights has been achieved. Today, senior Commanders have greater confidence in the information provided as they are able to drill down into data in a consistent manner to establish the explanations for unexpected results. Furthermore, there has been an increased visibility of their risks/issues providing a common basis for discussion.

More junior commanders find that they are included in the key performance reporting and risk management processes. They therefore experienced having access to a new range of information, and are thus able to be better informed about the past and predicted performance achievements, risks/issues and support by their suppliers.

Looking to the future, there are a number of areas that will be addressed, listed here in no specific order:

- It is intended to broaden the benefits of the scorecard approach and the SAPPHIRE application to the stations, who at the moment are mostly data providers instead of information consumers. One step is to design station

scorecard based on strategic maps (Value Creation Maps) which will identify their strategic priorities, the key performance questions they want to have an answer to, as well as the performance indicators that will help them to answer their questions. This work is ongoing and it is hoped that eventually it will lead to more focused Scorecards at all levels.

- One improvement would be to create more flexibility in reporting frequencies. Performance reporting is currently being undertaken on a quarterly cycle, a 'battle rhythm' dictated by the frequency of the Defence Management Board's review of its scorecard. There is a strong argument that the lower levels of the organisation would benefit from a more frequent review.
- There is also scope for wider use of the SAPPHIRE software application, which has been written to ensure that it can be used by any organisation. Thus there is potential for the system to be adopted more widely, both across the Ministry of Defence, across UK government and more widely in other organisations.
- Finally, an area that has proved to be a frustrating and enduring challenge, has been the integration of performance management and resource management. This demands the alignment of performance data and cost data at a more detailed level than is currently available. The integration of resource and performance management will be a challenge going forward...