THE UTILITY OF METRICS IN TRAINING NEEDS ANALYSIS - LESSONS LEARNT FROM RN TNA

Lt Cdr C B Cook RN
NTE(TTD)
Portsmouth, UK
2slntettdso2tna@a.dii.mod.uk

ABSTRACT

Adoption of formal Training Needs Analysis (TNA) in the Royal Navy (RN) followed an investigation by the UK National Audit Office (NAO) into the use of simulators in training. The NAO report, released in 1992, recommended the implementation of rigorous methods for assessing the effectiveness of training solutions across the UK Armed Forces. Now in common use, TNA is the single methodology recommended by the UK MOD Acquisition Management System for determining the most cost-effective methods of meeting training requirements.

The aim of this paper is to research and expose the ways in which the RN, through its published guidance on the conduct of TNA, has sought to fulfil the requirements of the NAO report. In particular, the Author will investigate and justify the importance awarded to auditability and objectivity, common threads to the evolving TNA methodology, and conduct a review of existing metrics employed in TNA. This review will explore the utility of metrics, based on evidence from RN TNA, and will present a set of lessons learnt from the implementation of quantification techniques. Thus the Author will attempt to set the limits of achievable objectivity throughout TNA and seek to disprove the commonly-held misconception that auditability is confined exclusively to the domain of metrics.

The paper will conclude with recommendations formulated to assist TNA practitioners strike an objectivity balance, which seeks to avoid reliance on metrics alone. The Author's recommendations will be placed in context of the RN's latest guidance on TNA, which seeks to redress the balance generated by earlier over-prescription of quantification. In this way, a practicable approach for addressing the objectivity/subjectivity equilibrium will be presented, enabling the training analyst to generate more timely, meaningful and reliable information in support of the acquisition process.

ABOUT THE AUTHOR

Lieutenant Commander Cook joined the Royal Navy in 1992. His first appointment was to HMS COLLINGWOOD as an electronics and digital techniques instructor. In 1994, he was appointed to HMS ARK ROYAL as Deputy Education Officer, later returning to resume his duties at HMS COLLINGWOOD. In 1996, he joined the staff of the Computer Based Training Section at the Royal Naval School of Educational and Training Technology (RNSETT).

Whilst at the RNSETT, Cook gained an MSc in Mathematical Science, and attended a number of I/ITSEC and ITEC conferences, presenting papers on the development of synthetic training solutions in the Royal Navy. In 1999, he was appointed Computer Operations Manager to the CINCFLEET HQ, Northwood, with primary responsibility for the NSTN Message Handling System. In 2000, he was promoted to Lieutenant Commander.

In 2002, Cook was tasked to design and implement a programme of Continuing Professional Development for the Engineer (Information Systems) Specialisation. Later in the same year, he was appointed to the Defence Simulation and Modelling MSc, completing a full-time programme of study at the Royal Military College of Science in July 2003.

Lieutenant Commander Cook is currently Head of the RN Training Needs Analysis Cell, within the Directorate of Naval Training and Education, Portsmouth, UK.

THE UTILITY OF METRICS IN TRAINING NEEDS ANALYSIS - LESSONS LEARNT FROM RN TNA

Lt Cdr C B Cook RN
NTE(TTD)
Portsmouth, UK
2slntettdso2tna@a.dii.mod.uk

INTRODUCTION

The introduction of Training Needs Analysis (TNA) in the British Armed Forces resulted from an investigation by the UK National Audit Office (NAO) into the use of simulators in training. Released in 1992, the NAO report (NAO-92) identified a number of instances where high fidelity, high cost simulators had been procured in support of aircrew training, when more cost-effective solutions were available. Consequently, this report recommended that the UK Ministry of Defence (MOD) implement rigorous methods for assessing the effectiveness of training solutions across the Armed Forces, in order that value for money could be assured.

Within the UK military community today, TNA is captured in a number of policy documents, but principally in the Tri-Service Guide to TNA (MOD-01). Here, TNA is described as a methodology designed specifically to identify training requirements and to establish the most cost-effective means by which such training might be delivered. Intended to be a structured and objective assessment of training needs, supported by a validated and reliable audit trail, TNA is the single methodology recommended by the UK MOD Acquisition Management System (AMS-04) for determining the most cost-effective methods of meeting training requirements.

IMPLEMENTING THE NAO RECOMMENDATIONS

In the Royal Navy (RN), the NAO recommendations are embedded in an accompanying guide to the RN Systems Approach to Training (MOD-99). Entitled "RN Instructions for the Conduct of TNA" (MOD-01a), this guide describes the TNA methodology as a systematic, iterative, 3-phased, product-based approach which provides an auditable trail of analysis to justify and enable the design of specialist training and the acquisition of training equipment. The 3 phases (illustrated at Figure 1) described in the Guide are:

Phase One – Scoping Study

This defines the TNA management, and highlights those constraints that will apply throughout the conduct of the TNA.

Phase Two – TNA Development

This comprises three separate stages:

Operational Task Analysis (OTA)
The output of this stage is the Operational Task Inventory which identifies those additional tasks that contribute to delivery of the new or changed operational requirement.

Training Gap Analysis (TGA)
This represents the additional training requirement that needs to be satisfied in order to bring personnel up to a prescribed operational performance standard to meet the new operational need.

Training Options Analysis (TOA)
This stage considers the relative merits and costs associated with a variety of methods and/or media which will bridge or partially bridge the training gap.

Final Report
Here, a proposed training solution is offered for endorsement.

Phase Three – Post Project Evaluation (PPE)

This assesses and reports on the conduct of the TNA process and the capability of the endorsed training solution to meet the operational requirement.

Phase 1	Phase 2				Phase 3
Scoping Study	Deliverable 1 (OTA)	Deliverable 2 (TGA)	Deliverable 3 (TOA)	Deliverable 4 (Final Report)	PPE

Figure 1. The TNA Process

Substantiated in the description above, the TNA methodology, as applied in the RN, is a rigorous and comprehensive process, designed to be both auditable and objective, in order to mitigate any potential risks associated with the acquisition process. It achieves this objective by focussing only on the new training requirement (that comprising the training gap) and establishing the most cost-effective way of bridging that gap.

AUDITABILITY AND OBJECTIVITY

Following the original NAO report, it should be unsurprising to note that one of the principal outputs of TNA is an audit trial of evidence in support of analysis collated throughout the process. In this way, the recommendations presented to the key stakeholders (who form what is known as the TNA Steering Group) on completion of the Final Report, should be traceable to substantiated evidence in order that the justification for recommendations can be exposed if the TNA becomes subject to scrutiny.

Thus it is critical that all data captured in the TNA is referenced to an appropriate source in order that due credence can be given to the TNA's final recommendations. Indeed, it is an unwise training needs analyst, who, if they happen to be an expert in the subject of study, relies only upon their own knowledge of the problem as a basis from which to make their recommendations.

Therefore, in order to ensure credibility of the analysis, it is a quality criterion of all RN TNA deliverables that there be an audit trail in support of the work presented. This audit trail should reference meetings, interviews, policy and doctrine to provide evidence that analysis is based upon fact, within agreed constraints, and subject to acceptable assumptions. Nonetheless, the purpose of the TNA is not simply to gather evidence: while analysis requires to be suitably informed, it is the value added by the expert analyst and the application of the TNA methodology that transforms raw data into a training gap, against which a cost-effective means of overcoming that gap can be identified.

However, while evidence based on fact is by definition objective, it can be argued that as soon as these facts are subjected to any sort of treatment, they immediately migrate into the domain of subjectivity, where the reliability with which they were associated can become significantly undermined. So from one perspective, TNA is required to present an auditable trail linking objective fact to justifiable recommendations, yet on the other hand, the very process of linking appears to undermine the factual basis on which so much reliance is placed.

There is a clear conflict here – that the TNA is required to be objective and auditable, whilst it is recognised that the methodology necessarily removes data from the domain of fact-based reliability, and subjects it to evolution and extrapolation in order to derive suitable recommendations. The role of the analyst, therefore, is to gather the facts and apply the methodology in such a way that evolution from the known to the unknown is less a leap of faith, and more a logical, substantiated application of objective, intelligent analysis – this is the real challenge to TNA.

METRICS IN TNA

In rising to this challenge, and although indispensable in TNA, fact-based evidence is not the most critical factor. Rather, it is how the evidence is subjected to analysis, and the way in which the analytical tools and metrics are employed, that makes the greatest contribution to assuring an acceptable level of objectivity throughout the TNA.

Present in the RN TNA methodology are a number of tools designed to provide the training needs analyst with a means of taking evidence forward. Although the RN methodology does not mandate the use of tools (it does, however, require that conclusions are substantiated through a suitable process), reference is made to 3 common approaches, known as: Difficulty/ Importance/Frequency (DIF) Analysis (MOD-00), Fidelity Analysis (MOD-01a) and Measures of Training Effectiveness (MOTE) Scoring (MOD-01a). Each tool seeks to provide a measure, or metric, of the training requirement, which can be used to inform subsequent analysis and, ultimately, the

recommendation of a suitable and cost-effective training solution. Each of these methods are now described briefly, with the last tool (MOTE Scoring), subjected to a more detailed exposure later in the paper by means of a case study.

DIF Analysis

DIF analysis, conducted during Operational Task Analysis, derives a training priority for each operational task, by applying an assessment of task difficulty, importance and frequency to a suitable algorithm. Scoring is typically on a scale of 1-3, with the output of the DIF algorithm a singleton score of priority, intended to be more meaningful than an array of 3 individual values. The DIF priority is designed to provide an indication of the weight associated with each task. That is to say, if training resources are limited, training priorities can be used to determine which tasks should be trained first, and which could be deferred to a later date.

Fidelity Analysis

Conducted during Training Gap Analysis, Fidelity Analysis establishes the fidelity required of a training medium in terms of functional, physical and environmental parameters. If required, these 3 fidelity categories can be sub-divided further: functional into format, content and response; physical into spatial, tactile and appearance; environmental into sound, motion and ambience.

Taking the 9 fidelity sub-categories and employing a 4 level scoring metric (0-3), scores of between 0 and 9 can be derived for each of functional, physical and environmental fidelity. If required, a total fidelity score of between 0 and 27 can be established by summing individual scores.

The objective of fidelity analysis is to generate a fidelity requirements metric, in order to inform the training options analysis process of the degree to which the training medium needs to represent the actual operational equipment and the environment in which it will be operated and maintained. Thus, fidelity analysis should prevent the acquisition of expensive, overly-representative training aids, where less expensive but equally effective solutions may be available.

MOTE Scoring

The task of measuring the effectiveness of training methods and media, the objective of Training Options Analysis (TOA) in the TNA process, is acknowledged as being one of the most complex aspects of TNA. Despite significant attention in recent years, this challenge has not diminished, and Bowden's submission (BOW-99) that the selection of cost-effective methods and media for training is one of the main challenges to the TNA process, is still reflected in more recent work (ORT-04).

Selection of effective training media has traditionally been conducted through a range of methods. These include:

a. The practice of employing only the real equipment.

b. Tasking the equipment manufacturer to deliver supporting courseware.

c. Selecting a training solution based on historical evidence.

d. Selecting a training solution based on expert judgement.

e. Selecting a training solution based on a systematic approach.

While options (a) and (b) may sometimes deliver training solutions fit for purpose, they are either devoid of any analysis of the training need, or devolve responsibility for the evaluation of training effectiveness to a third party, whose solutions may not be cost-effective. Selecting training solutions based on historical evidence will clearly play a role in all training options analyses, but emergence of new technologies and the changing military operational and training environment, means that over-reliance on historical data can carry considerable risk. Similarly, employment of expert judgement can be problematic, as expertise tends to be based upon experience and opinion, and is therefore unavoidably subjective.

Thus a systematic approach to training options analysis, incorporating an auditable process of quantified judgements, appears to be the most appropriate way of identifying cost-effective training solutions. Use of such an approach does not preclude the application of historical or expert knowledge, but rather it tempers the injection of such knowledge with a stamp of quantification and traceability, which better meets the objective of the NAO than would any of the other alternatives alone.

It is not surprising therefore, that systematic approaches to TOA have been subject to much investigation. Examples of this include the application of Quality Function Deployment to the identification of cost-effective instructional products, examined by Jeffery et al (JEF-02), and the US Navy's TRADAM model, which seeks to provide an economic analysis of training delivery technology alternatives (HAS-00).

The RN approach, known formally as MOTE Scoring (MOD-01a), seeks to derive an objective measure of a training medium's ability to satisfy a set of training requirements. Examination focuses on the ability of media to meet the operational performance, conditions and standards associated with the tasks that constitute the training gap. Once each candidate training medium has been scored, media can be ranked according to their respective training effectiveness, in order to inform the process of selecting the most cost-effective solution.

OBJECTIVITY LIMITATIONS

Up to now, this paper has outlined the reasons for conducting TNA in the RN and has stressed the importance of presenting substantiated and auditable evidence. Furthermore, the paper has exposed in detail some of the metrics that, when applied to evidence, enable the TNA to progress and develop into a reasoned set of conclusions and recommendations. In these discussions, it has been proposed that a significant conflict exists within TNA: that the TNA is required to be objective and auditable, whilst it necessarily introduces subjectivity through the application of analytical tools.

Subjectivity is therefore introduced, principally, through the application of analysis (the process). It is equally arguable, however, that the input data on which TNAs rely is itself of limited reliability, as the information (data) gathered by the analyst is often based on subjective opinion, even before it is exposed to analysis.

Subjectivity of Data

Present in all 3 metrics described so far, is significant reliance upon the views of individuals. DIF, fidelity and MOTE scores all rely on the opinions of panels of individuals charged with relating their own assessments of, for example, the suitability of a particular medium to deliver training in physical skills. Whilst this evidence is mitigated by the forming of consensus wherever possible, it would be a brave analyst who suggested that the evidence gathered from SME interviews was "fact".

Subjectivity of Process

Arguably therefore, even before it is manipulated, the data gathered for the purpose of analysis is already subjective. Taking the example of DIF analysis, once the input data has been subjected to the DIF algorithm, which itself is a subjective interpretation of the interaction between task difficulty, importance and frequency, the valuable objectivity present initially in the raw data is reduced. Thus, the final priority assigned to a task is the mathematical product of a number of subjective assessments, and carries a lower degree of reliability than the source data upon which it was based.

The effects of compounded subjectivity are particularly evident in MOTE scoring. MOTE scores are based on the subjective views of individuals required to make assessments relating to the suitability of training media against training requirements. The subsequent reduction in objectivity, arising from the compounded subjectivity of data and process, will now be exposed in a metrics evolution case study, which will also describe how current practice seeks to increase and re-introduce objectivity at key stages of the TNA.

METRICS EVOLUTION – CASE STUDY

As has already been described, the MOTE scoring approach employed in RN TNA, is designed to establish, from a range of possible candidate solutions, the suitability of a particular medium to a particular training requirement (Figure 2). In TNAs conducted 5 or so years ago (MOD-99a), the approach taken to MOTE scoring involved considerable mathematical analysis, and was based upon the allocation of weightings to tasks, dependant on the training priority derived from earlier DIF analysis. Tasks with a higher training priority were allocated a higher weighting, in order that relatively unimportant tasks did not skew the selection of training media unrepresentatively.

Once scores had been obtained from SMEs, weighted scores were calculated by multiplying scores by weights. Weighted scores per task were then averaged according to the numbers on the MOTE panel. The suitability or otherwise of training media against task was then determined through comparison against a compliance threshold. Finally, the MOTE scores for each training media option were calculated by summing the weighted scores for each task/option and dividing by the number of tasks.

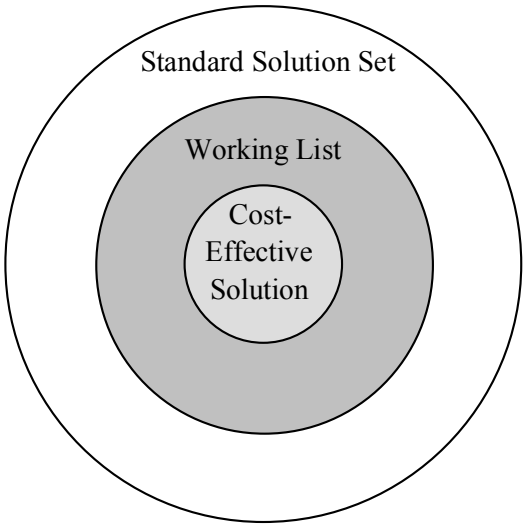

Figure 2. The MOTE Target

While this process of scoring and weighting is mathematically straightforward, it introduces subjectivity at a number of levels:

1. From the DIF analysis.
2. From the scores assigned by individual scorers.
3. From the weights assigned to DIF priorities.
4. From the position of the compliance threshold.

The degree of subjectivity is further compounded if training effectiveness is influenced by the degree to which potential media are judged to meet a range of criteria, such as risk and training management, which themselves can be prioritised in terms of importance. If incorporated in the MOTE scoring process, this adds a fifth layer of subjectivity. Thus, even with a confidence level as high as 87% at each stage, the effects of compounded subjectivity render the confidence of the combined MOTE process at less than 50%. On this basis, it is clear that the risks associated with subjectivity can be significant.

In order to mitigate these risks, a number of checks can be applied:

Range Check
The range check establishes the spread of task/media scores. This requires further analysis to be conducted where individual scores differ significantly.

Sensitivity Check
This assesses the sensitivity of MOTE results to the weights assigned to DIF scores, and suggests that where the allocation of weights produces a stable model, significant variation in DIF scores in unlikely.

Whilst these checks are designed to highlight any spurious results and thus limit subjectivity, the fact that they are required at all implies that the considerable volume of figures (8 separate calculations per task/medium pair) generated by this approach to MOTE scoring led to a rapid obscuring of the output data.

It is certainly not the view of the Author that the method described here was the basis of sub-optimal analysis, nor that the recommendations drawn from such evidence was flawed. Rather that this approach was at times over-complicated, time consuming, open to error and reverse-engineering and difficult to review. Although founded upon the rationale that mathematical manipulation converts subjectivity to objectivity, it is suggested that in the case of TNA MOTE scoring, this is not the case.

It is for these reasons that current advice is somewhat less prescriptive. The latest issue of the RN Instructions for the Conduct of TNA (MOD-01a) advise that:

"The MOTE seeks to derive an objective measure of a training medium's effectiveness in satisfying a set of training requirements. It is fundamental that this is a direct measure of effectiveness for each task rather than a calculated score derived from weighted combinations of a number of factors".

Thus the subjectivity inherent in earlier MOTE is reduced significantly, and MOTE scoring is now based upon assessment of the effectiveness of training only (that is the degree to which training prepares people for their jobs). Indeed, current guidance (MOD-01a), whilst not precluding the use of weightings, proposes that there is now some debate as to the worth of such a weighting process within MOTE scoring.

TNAs conducted today by the RN's own TNA Cell therefore employ a simpler approach to MOTE scoring, whereby 3 of the 5 levels of subjectivity present in previous techniques are eliminated. Scorers are asked to assess the effectiveness of media against tasks, with the raw results of MOTE scoring developed into cost-effectiveness recommendations principally through written analysis in the body of the TNA report.

Thus, the emphasis on manipulation of numbers is replaced by descriptions of the advantages and disadvantages of each media option (MOD-01a), with supporting arguments provided for the recommended

solution. In this way the subjectivity of data and process is reduced, whilst at the same time decreasing the level of obscurity previously introduced into MOTE scoring from the application of mathematical manipulation.

STRIKING THE OBJECTIVITY BALANCE

The challenge to TNA has already been presented: to combine facts and methodology in such a way as to present a logical, substantiated application of objective analysis in favour of 'a leap of faith'. Having been exposed to subjectivity through the case study, and the approach to MOTE scoring currently employed in the RN, it is now timely to describe the additional processes embedded in RN methodology, which mitigate the risk of subjectivity undermining the validity of the TNA.

Fundamentally, these processes fall into 4 categories, which individually or in combination, ensure that the recommendations of TNA are sound:

Training Expertise

The importance of limiting the conduct of TNA to professional training experts is not to be underestimated. As proposed (MOS-93) when TNA in the RN was still immature, no single options analysis model holds all the answers, and whatever approach is adopted will still require a skilled analyst to identify the appropriate attributes and to identify the many exceptions to the rules. A thorough understanding of the TNA methodology, the systems approach to training and training design is therefore a necessity.

Limiting the Subjectivity Risk

As evidenced in this paper, the simple but effective analytical tools now employed by the RN are designed to limit the risks of compounded subjectivity. In this way, the evolutionary distance between the inputs and outputs of a single stage is restricted to simple logical or mathematical abstractions, with the majority of analysis conducted within the analyst's own narrative.

Consensus

All DIF, fidelity and MOTE panels are formed of a group of SMEs, rather than reflecting the opinions of isolated individuals. In this way, the risk of individual views skewing results is limited.

Ownership

The TNA Steering Group (SG) are required to take corporate ownership of all analysis and are ultimately responsible for endorsing the work presented to them. In doing so, SG members are involved in a continual process of TNA product review and endorsement, in order to reduce the possibility that subjective judgements might unduly influence later analysis (Figure 3). Thus the product of a single TNA stage, before it is taken forward into subsequent analysis, is exposed to a process of subjectivity sanitisation, which, whilst not fool-proof, significantly reduces the risk of compounded subjectivity.

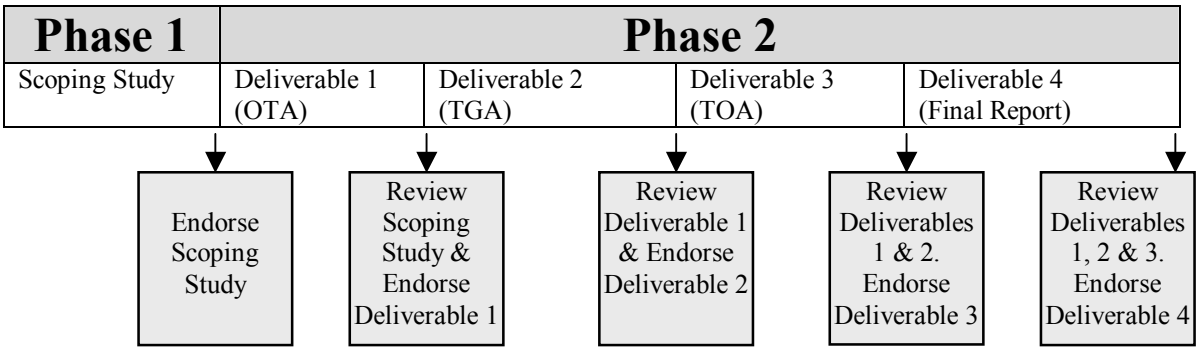

Figure 3. Subjectivity Sanitisation - the TNA Process with End-Stage Review

THE AUDITABILITY/METRICS DEPENDENCY

It is the dual intent of all TNAs to be both objective, and auditable: objectivity is required in order that recommendations can be robustly supported by evidence, and auditability is required in order that the evidence can be substantiated if required. Early approaches to MOTE scoring were based on the assumption that metrics gave both. This paper has already argued that metrics do not ensure objectivity.

In a similar manner, metrics are not the sole guarantor of auditability, that is to say, auditability is not confined exclusively to the domain of metrics.

A TNA will be equally auditable if it is void of any metrics whatsoever. Indeed, the work of Morris (MOR-02) suggests that selection by justification (that is to say, by a consensus of experts obtained through detailed, minuted discussion), is as auditable a process of conducting TOA as one supported by extensive metrics.

It is therefore the duty of the members of the TNA Steering Group, who are responsible for endorsing the output of the TNA, not to be blinded by metrics and figures, but to look to the evidence on which the TNA is based, and thereby assure themselves that the recommendations are sound.

CONCLUSION - LESSONS LEARNT

Many important lessons have been learnt since the introduction of TNA into the RN. In terms of the use of metrics, current guidance recommends a simplification of the methodology, by reducing the amount of mathematical manipulation. This has generated TNAs that are more understandable, less complex, easier to review, easier to re-visit, less prone to error and less open to reverse engineering. Emphasis has been placed on the employment of intelligent, experienced analysts and the presentation of sound argument based on informed analysis in order to support a TNA's recommendations. Emphasis has also been placed on the review of deliverables at the completion of each stage, to secure endorsement based on consensus, and to reduce the amount of subjectivity that would otherwise pass from one stage to the next.

In these ways, the risks of subjectivity introduced through process and data are mitigated, and current practice provides far greater transparency to TNA. Notwithstanding, TNA cannot be wholly objective. Indeed, the Tri-Service Guide (MOD-01) describes it as, by its nature, subjective, and recognises the importance of subject matter expert consultation throughout.

Ultimately, therefore, it falls to the TNA Steering Group to ensure that the methodology applied, and the metrics employed are fit for purpose, and that the recommendations for acquisition of a cost-effective training solution are based on sound and auditable research, and logical, intelligent analysis. Certainly, the process is not devoid of subjectivity, but the training community can undoubtedly be more confident today than ever, that the training systems being acquired are indeed optimum solutions within available funds.

REFERENCES

AMS-04. UK MOD (2004). *Acquisition Management System*, www.ams.mod.uk.

BOW-99, Bowden M N (1999). *Training Options Analysis*. A paper presented to the 10th International Training and Equipment Conference.

HAS-98, Hassen J E, Weickhorst R A (1998). *TRADAM: A Catalyst For Training Technology Infusion*. A paper presented to the 9th International Training and Equipment Conference.

JEF-02, Jeffery A B, Bratton-Jeffery M F (2002). *Using Quality Function Deployment in* Training Systems Design and Development*. A paper presented to the I/ITSEC Conference, Orlando, Florida, USA*.

MOD-99, UK MOD (1999). *Royal Navy Systems Approach to Training Quality Standard, Issue 2*.

MOD-99a, UK MOD (1999). *LPD(R) Command Team Training Needs Analysis*.

MOD-00, UK MOD (2000). *BR8420 Training Guide 2 – Analysis and Design of Training*.

MOD-01, UK MOD (2001). *JSP 502 - The Tri-Service Guide to Training Needs Analysis*.

MOD-01a, UK MOD (2001). *BR8420 Training Guide 1 – RN Instructions for the Conduct of Training Needs Analysis*, 2001.

MOR-02, Morris, P (2002). *Training Options Research*. A paper by the RNSETT.

MOS-93, Moses F L, Laveson J I (1993). *Changing Strategies for Training Military Units*. Plenum Press, New York, 1993.

NAO-92, UK Government National Audit Office, (1992). *Ministry of Defence: use of simulators in training: fifty-ninth report from the Committee of Public Accounts.*

ORT-04, Orton D M (2004). *Training Needs Analysis – Way Ahead*. A paper by NTE(TTD).