Vision Paper: Integrating VV&A Methods and Cost-Effectiveness Analysis in the Acquisition Process for Training Simulation Solutions

Wim Huiskamp
Jeroen Voogd
TNO, P.O. Box 96864, 2509 JG, The Hague,
THE NETHERLANDS

Hans Korteling
TNO, P.O. Box 23, 3769 ZG, Soesterberg,
THE NETHERLANDS

ABSTRACT

Simulation is an important technology that enables NATO and its member nations to train their soldiers. The benefits of simulation-based training include saving of time, money, and even lives, when training for unsafe scenarios. Simulation also facilitates joint and combined training. Moreover, simulation-based training is capable of expanding the limits of live training, thus facilitating larger exercises. The acquisition of valid and cost-effective training simulation solutions is crucial to the mission-readiness of our armed forces, in particular when available funding and resources are limited.

This paper presents our vision of the defence acquisition processes for simulation systems by integration with verification, validation and accreditation (VV&A) methods and cost-effectiveness analysis (CEA), in order to assure that a valid and cost-effective training simulation solution is acquired. We envision supporting the whole acquisition process (e.g., stating the need, training needs analysis, requirements analysis, evaluation and selection, and acceptance of assets), taking all other cost aspects of the training simulation solution lifecycle into account (e.g., deployment, maintenance, re-use, retirement).

All three of the processes mentioned above include multiple activities and tasks, and they require large amounts of information. We propose to combine activities where possible and to ensure that inputs and outputs match. For example, a good set of requirements based on the user needs is input for the acquisition process, but also important as one of the starting points of VV&A, as well as for CEA.

The authors hope to further develop the ideas presented in this vision paper. This future work should be performed in close cooperation with ministry of defence (MoD) procurement organizations and should also engage with the wider NATO research community, in particular the SAS panel and the NMSG. The presentation of the initial concept to the SAS-095 task-group on Cost-benefit Analysis of Military Training is seen as a first step in that direction.

1.0 INTRODUCTION

Modelling and simulation is an important technology that enables NATO and its member nations to perform training, analysis, and concept development, as well as to test and experiment. Some particular benefits on the training side include saving time, money, and even lives, when training for unsafe scenarios. Simulation-based training is not necessarily constrained by available training ranges, thus facilitating larger exercises. Simulation also facilitates joint and combined training by enabling distributed networked training events, which saves travel time and replaces live units by role players or computer generated forces. Distributed simulation is in fact rapidly becoming a necessary prerequisite for collective mission training.

The increasing use of training simulation systems has revealed that the expected cost-effectiveness improvements are not always achieved (Farmer et al., 1999; Emmerik & Korteling, in press; Janssens, 2012). This may be caused by several overarching reasons (Emmerik & Korteling, 2012):

- Specification: lacking or insufficient training-needs analysis, resulting in unclear specifications or problems translating user-needs into technical requirements.
- Didactical: lacking understanding of the limitations of the training system for the intended use, or insufficient use of the added training possibilities of the system.
- Integration or embedding in the organization: lacking vision regarding the acquisition and use of training simulation systems, lacking governance on the operational use, infrastructure, maintenance support, etc.

This paper presents an overall vision for the defence acquisition processes for simulation systems by closely integrating it with verification, validation, and accreditation (VV&A) methods and cost-effectiveness analysis (CEA). The objective is the acquisition of valid and cost-effective training simulations.

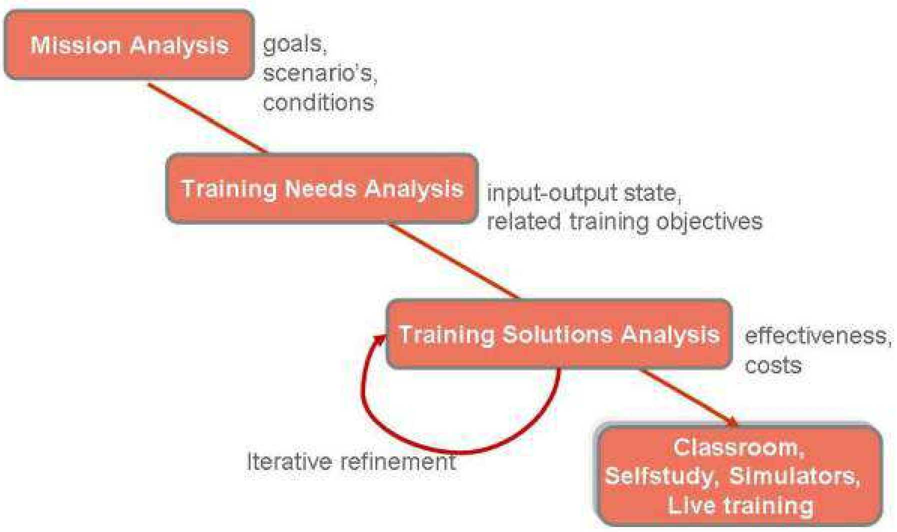

Figure 1: Selecting training simulation solutions

Our vision consists of supporting the whole acquisition process (e.g., stating the need, training needs analysis, requirements analysis, solution evaluation and selection, and acceptance of assets), taking all other cost aspects of the training simulation solution lifecycle into account (e.g., deployment, maintenance, re-use, retirement).

2.0 CURRENT ACQUISITION PRACTICES

In this section, we describe the major common and current processes and procedures involved in the acquisition of expensive defence-related training simulations.

2.1 TYPICAL ACQUISITION PROCESS

The Netherlands MoD uses a procurement process known as DMP (Defence Materiel Process). This process (NL MoD, 2007) is similar to what is used in other nations and will be described here as a generic example. Several phases may be distinguished:

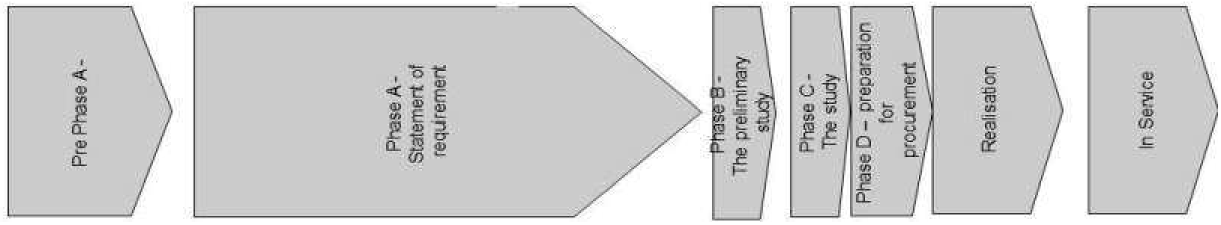

Figure 2: Dutch procurement process phases

- DMP Pre-Phase: Start of process
  Once a capability gap is observed in some area (in our case a training gap), the new requirement will be part of the integral defence planning process (IDPP), and a DMP-process is initiated. Section 2.2 describes the training needs analysis, which should typically be executed before or during this phase.
- DMP Phase-A: Setting the requirements
  The requirements, which are necessary to achieve the objectives of the defence organization, are derived transparently from policy and plans.
- DMP Phase-B: Preliminary study
  Phase B is concerned with translating the requirement into functional and, where possible, technical requirements that the product must satisfy. A market review is carried out and product alternatives and risks are investigated.
- DMP Phase-C: Study
  In Phase C, the general requirements from the previous phase are worked out in more detail, and a shortlist is drawn up of the most eligible alternatives. They are then assessed in terms of the more elaborate requirements.
- DMP Phase-D: Preparing the procurement
  In Phase D, a particular product and the supplier are selected. In general, at the conclusion of this phase and after a line of credit has been approved, a contract is signed.
- Realization Phase: Evaluation
  Evaluations have two separate evaluation milestones: a project evaluation immediately following the conclusion of the project and a usage evaluation after the product has been in use for some time.

A large number of aspects are common to all of the phases and are elaborated and detailed further with each successive phase. Aspects include the specification of requirements, analysis of consequences, formulation of the procurement strategy and time schedule, and risk management.

The acquisition of simulation systems is a complex process requiring extensive experience, knowledge, and skill in order to specify, design, develop, and integrate systems into a solution that meets operational, functional, security, and technical requirements. Acquisition and development of simulation systems is a multifaceted problem with many stakeholders. It ranges from understanding operational needs to technical exchange of data to lifecycle costing assessments. One of the big challenges in realizing the full potential of simulation for defence purposes is selecting the best solution or mix of solutions (schoolhouse, simulation, live training) for a specific training need. The MoD needs training solutions that have been validated for the intended purpose (fit for purpose) and that are also cost-effective, in order to meet the shrinking budgets and limited resources available for our armed forces.

The acquisition of training simulation systems currently follows the generic process described above. In order to improve the cost-effectiveness of the training simulation solution, a number of specific activities should be carried out and these must be mapped onto the different phases in the acquisition process.

2.2 Training Needs Analysis (TNA)

Assuming that a training gap must be addressed, the DMP pre-phase is initiated. The first step is to carefully analyse and specify the mission, tasks, and training needs that have to be met (Figure 1). This is the so-called training needs analysis (TNA). Most DMP processes for advanced training systems include a TNA, although the level of detail and sophistication may vary substantially (Cohn et al., 2009).

A typical TNA always starts with a *mission analysis* (e.g., Cohn et al., 2009; Farmer et al., 1999; Korteling et al., 2001; Korteling et al., 2011). This involves a global description of the context of a task or function to be trained for. Here a task is defined as a goal-directed sequence of activities allocated to a person, which can be described at various levels. The mission analysis includes a description of the goals of the task(s) and how these goals should be achieved, along with the functions of systems involved and the relevant scenarios, circumstances, or environmental conditions (physical, mental, and tactical aspects) that are encountered in task performance.

Next, the *training analysis* specifies the training requirements in terms of (a set of) related learning or training objectives (e.g. Cohn et al., 2009; Farmer et al., 1999). The training objectives are determined on the basis of the discrepancy between the required knowledge, skills, and attitudes (i.e., competencies to be deduced from the task or function to be trained in the context of the mission) and the current competence-level of the trainees (target group). Training objectives have to be specified in terms of existing and required competence levels of the trainees, type of competencies that have to be trained for, and the required task performance to demonstrate that the training objectives have been met.

The final step of the TNA is the *task analysis* (e.g. Cohn et al., 2009; Farmer et al., 1999). From the training goals identified in the training analysis, the task analysis decomposes a task or function into (preferably non-overlapping) elements (tasks or subtasks). Of these (sub-) tasks the most salient features are concisely described in, for example, terms of input, operation(s) to be performed, output, critical conditions, and critical task-elements. In addition, a brief global description of the subtask may be given. The number of subtasks should be kept as low as possible. A non-redundant task analysis (without repetitions) provides the most efficient representation of all task and context variables that are involved in the task.

With the task and training analysis, the TNA finally results in a (clear) description of the education and training requirements for the trainee that have to be fulfilled by the training system. The degree to which this aim is successfully accomplished—with proper arguments and to a sufficient level of detail—varies significantly between different DMP instances.

2.3 VERIFICATION AND VALIDATION

The optimal VV&A method depends on the individual needs and constraints of an organization or application domain. However, common principles and best practices are clearly recognizable, and this was the key driver behind the development of the Generic Methodology for Verification and Validation (GM-VV), which is currently being standardized as a recommended practise by SISO (SISO, 2012; SISO, to be published; Roza et al., 2012). The main purpose of GM-VV is to provide a general baseline and guidance for VV&A of M&S that is applicable and tailorable to the individual VV&A needs of a wide variety of M&S technologies and application domains. This is the reason we use the GM-VV in this work for mapping VV&A and CEA onto an acquisition process.

The GM-VV is an abstract framework that consists of three parts (the conceptual framework, the implementation framework, and the tailoring framework) that build upon existing VV&A methods and practices. The GM-VV conceptual framework provides essential VV&A terminology, semantics, concepts, and principles. The framework facilitates communication, understanding, and implementation of VV&A across and between different M&S contexts. In contrast to many views on V&V—namely, that it starts with the M&S requirements and ends with the developed M&S asset—this framework is premised on the idea that models and simulations are always developed and employed to fulfil the specific needs of their stakeholders (e.g., trainers, decision makers).

The GM-VV implementation framework consists of the interrelated products, processes, and organization. The product dimension contains information-based VV&A products that can have multiple instances, and representational and documentation formats. These VV&A products are produced by the processes, activities, and tasks defined by the process dimension. They can be executed recursively, concurrently, and iteratively. The roles defined in the organizational dimension are involved in the execution in one or more of the VV&A processes, activities, and tasks.

The GM-VV tailoring framework provides ways to tailor the implementation framework for each individual M&S organization, project, or application domain. The tailoring allows for modification of the building blocks in the GM-VV product, process, and organizational dimensions to satisfy the specific VV&A requirements and constraints in the M&S environment in which the GM-VV is applied. During the execution of the V&V work, risk-based tailoring is used to find the optimum cost-benefit ratio (e.g., distributing project resources based on M&S use risk).

GM-VV Argumentation Structure

The objective of a V&V effort is to develop evidence upon which an acceptance recommendation is based. This V&V objective is articulated as an acceptance goal. This high-level goal should be translated into a set of concrete and assessable acceptability criteria for the M&S system or result(s). Relevant and convincing evidence should then be collected or generated to assess the satisfaction of these criteria. When it is convincingly demonstrated to what extent the M&S system or result(s) does or does not satisfy all these acceptability criteria, a claim can be made on whether or not the M&S system or result(s) is acceptable for its intended use (i.e., acceptance claim) (**Figure 3**).

Developing an acceptance recommendation usually involves the identification of many interdependent acceptability criteria. Collecting the appropriate evidence is not always straightforward, and it is not always possible due to various practical constraints. Besides the usual resource constraints, in this paper we deal with situations where the training system might not be available yet, or where one has to rely on claims from various training systems suppliers.

The GM-VV argumentation structure approach is well suited for the work proposed in this paper. It provides a systematic approach to deriving relevant criteria. These criteria may be related to the training system at hand (utility, fidelity, correctness) and to the processes with which a training system is developed. Not all criteria are equally important; this fact translates into different weights for the branches in the argumentation structure.

The argumentation structure approach is flexible enough to be used in various ways. One approach is to have criteria for all assets in a complete training program (classroom, simulation, live). Another approach is to look at the prospective training system and develop criteria on how well the future users have looked at the effective use of the prospective system in relation to the other training systems and material.

If the argumentation structure has been constructed for a specific training purpose, it can be applied to all candidate systems identified in the DMP-B and C phases. This will show for each system where it will work well and where it is lacking. Once a specific training simulation has been chosen, and it becomes available, the experimental frame can be executed (again, but now more thoroughly), and the second part of the argumentation structure, the claim network, can be build. This leads to an acceptance recommendation in which all the strong and weak points of the system can be traced back to what this means for the purpose.

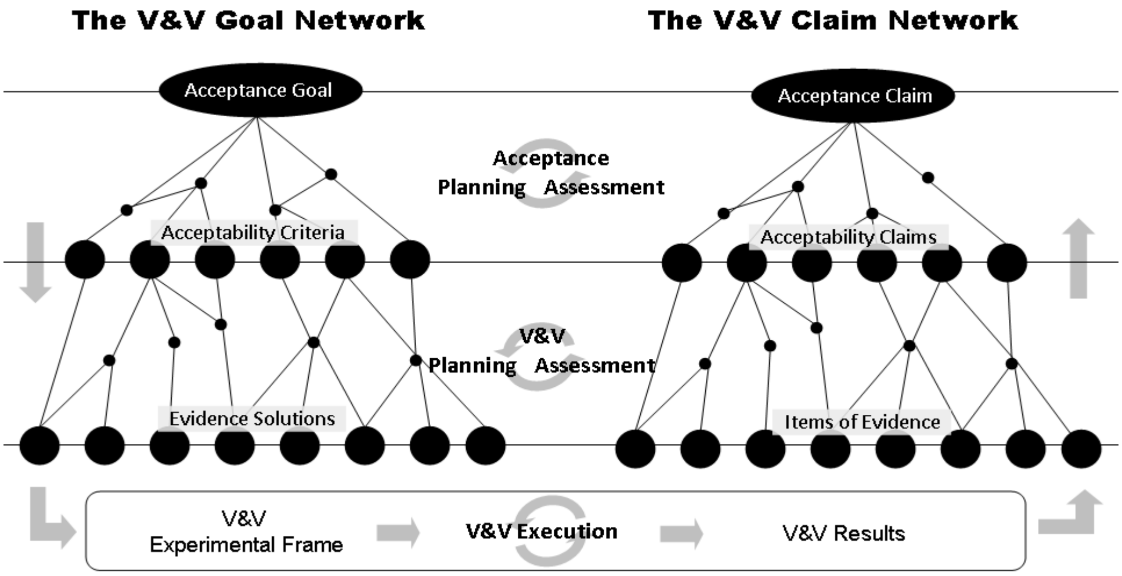

Figure 3: Argumentation structure consisting of a Goal (left) and a Claim (right) Network

The GM-VV has been successfully applied to a variety of M&S applications and domains: serious games, simulation-based experimentation, methodologies involving the use of M&S assets, data such as synthetic terrain databases, etc. A V&V Centre of Excellence, Q-tility (Q-tility, 2012) has been established to provide V&V services based on the GM-VV.

2.4 GLOBAL COST ANALYSIS

Present DMP activities usually focus on *cost aspects* of simulation. Cost generally means “That which must be given or surrendered in order to acquire, produce, accomplish or maintain something” (NATO SAS TG-069, 2009). The procurer’s task is typically to obtain a training system that fulfils the training needs and system requirements (Section 2.1) at the lowest cost possible and within a certain budgetary limit that is often set in advance. In many procurement processes, cost is mainly quantified as the economic price of the system itself. Other cost factors, such as labour, calendar time, or resulting capability may not be sufficiently taken into account (NATO SAS TG-069, 2009). Cost should be relevant (e.g., attributed or refer) to either the simulation infrastructure, or to the objective system(s) developed. Capital investment and usage (i.e., operational) costs should also be included, but are often based on rough estimates due to lack of historical data from existing systems or because of unknown characteristics of the new system.

A common definition for cost and its role within the lifecycle includes cost factors and cost drivers:

- a cost factor can be either (1) a cost-estimating relationship (CER), in which cost is directly proportional to a single independent variable; or (2) a brief arithmetic expression wherein cost is determined by the application of a particular factor, such as the percentage of labour or material.
- a cost driver may be interpreted as either (1) an element or additive component of cost; or (2) a determinative affecter of cost, represented in some cases as a multiplicative weighting applied to a

relevant additive cost component. Often, the term cost driver is reserved to denote the most significant causal factor in cost (NATO SAS TG-069, 2009).

In line with the basic principles of cost- and lifecycle analysis, our vision aims to support the whole acquisition process (e.g., stating the need, training needs analysis, requirements analysis, evaluation and selection, and acceptance of assets), taking all other cost aspects of the simulation solution lifecycle into account (e.g., deployment, maintenance, re-use, retirement). The proposed support vision calls for an integration of verification, validation, and accreditation (VV&A) methods and cost-effectiveness analysis (CEA) to make sure that a valid and cost-effective training simulation solution is acquired.

3.0 INTEGRATION OF PROCESSES

A tight integration of the three processes discussed above (acquisition, CEA, V&V) would mean that all process steps, roles, and information flows would be tuned in order to obtain the most efficient integrated process, resulting in the most cost-effective training-simulation solution. This is something that we aim for in the near future, but for now this comprehensive integration has not yet been established. Instead, here we focus on a somewhat higher level of integration. This section provides some practical guidelines and steps to be taken for the specification and acquisition of training simulators and instructional games. This is done from a combined didactical and cost-effectiveness point of view, aiming at optimal transfer of training at minimal cost. The guidelines and principles presented here are based on information from handbooks and studies that have been carried out before on the requirements analysis and specification of synthetic training environments (Cohn, 2009, Farmer et al., 1999, Korteling et al., 2001; Korteling et al., 2011; Milham et al., 2009; Stanney, 2009; Verstegen, 2004; Young, 2001; Young, 2004). The guidelines for V&V are derived from Voogd et al. (2009) and SISO (2012). The important elements which form the backbone of the acquisition process are discussed below. For each process the most important CEA and V&V aspects are indicated, and ideas are presented for their integration.

3.1 Pre-DMP Analyses

After the TNA (see Section 2.1), the feasibility and prospects of simulation for specific tasks to be trained for relative to other education and training alternatives should be investigated (Farmer et al., 1999; Korteling et al., 2001; Korteling et al., 2011). This activity may be seen as training solution analysis (TSA). The alternatives to be considered include acquisition of a new system, adaptation of existing systems, optimal mix of live, virtual, and classroom training. When global comparisons between alternatives have to be made, a global cost-utility analysis may be a good solution (see Section 3.1.2).

In this phase, V&V should begin. The argumentation structure (AS) (**Figure 3**) is initiated, and the following DMP stepped criteria are added. Criteria on effectiveness, efficiency, and risk are added to the AS. Criteria are not restricted to those directly related to the purpose (training) or the M&S system. One can also add criteria related to the quality of (sub-) processes and roles. For example: Was the TNA performed correctly (right process, right experts, etc.)? Was an up-to-date listing of already available M&S assets available and used?

3.1.1 Global cost-utility analysis

For simulation-based training, the first step of a cost-effectiveness analysis involves the investigation of the feasibility and prospects for (different forms of) simulation for the specific tasks to be trained for relative to other alternatives. When global comparisons between alternatives have to be made, as is prescribed by DMP Phases B and C, a global cost-utility analysis may be a good solution. Such an analysis includes weighted measures of expected training values and foreseeable training cost.

3.1.2 Utility Analysis

The utility analysis considers the expected training value of alternative solutions in terms of general utility components. It helps to choose among potential alternatives. The following general prerequisites of a utility analysis can be respected:

- A number (at least two) of alternative possibilities must be defined by using the same set of criteria;
- The criteria must be distinctive, non-overlapping, and exhaustive for the decision problem;
- The criteria can be weighted according to their relative importance and, thus, may be combined into one utility factor.

The utility of training a certain task or function by simulation is its usefulness for the training organization (i.e., defence). It can be globally assessed by the following major components:

- Value of the provided training for the organization. That is, how important is this training for fulfilling the typical mission and preserving the continuity of the division, branch, or system;
- Importance of the addressed training difficulties or bottlenecks like safety, limited training ranges, etc., and the possibilities for overcoming these limitations by using simulation technology;
- Availability of training resources, taking into account environmental restrictions, training, and exercise logistics, instructors, time constraints, etc.;
- The attractiveness of a simulator. Simulation provides opportunities to perform tasks in a realistic context, which is also attractive and stimulating;
- Instructional games can be played outside a professional context, outside working hours. If workers perceive that a simulator is attractive, this can yield more training hours without additional organizational cost.

3.1.3 Cost analysis

The cost analysis considers the cost reduction potential of alternative training simulators, which is mainly determined by simulation technology and the personnel required (students and instructors). For every alternative training solution, at least the following cost categories should be estimated per year:

- Number of students
- Instructor cost and student salaries
- Cost of scenario/lessons development (3-D databases, maps, etc.)
- Hardware cost and software cost (development and/or license cost, updates). Infrastructure cost (network, extra electricity facilities, building maintenance, etc.)
- Infrastructure operating cost (electricity, heating)
- Simulator maintenance and technical support (personnel) costs
- Cost of accommodation
- Travel cost
- Cost of documentation

3.1.4 Decision on the Basis of the Cost-Utility Analysis

The final step of this global cost-utility analysis is to compare the total cost of every alternative with their utility values, and decide which alternatives have a high utility to cost ratio and, thus, can be put on a shortlist for further analysis (Figure 4). In that case, the mission, tasks, and training needs have to be further mapped onto the selected solution(s) (or mix of solutions) to specify the most cost-effective training solution on a more detailed level.

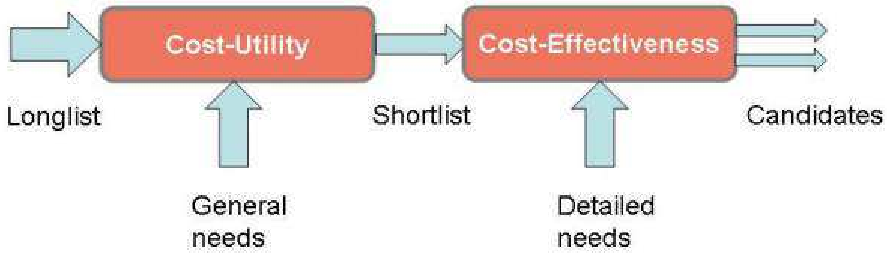

Figure 4: Downselecting Alternatives

3.2 DMP A and B Phases: Global Functional Specifications

In this step, the functional requirements of the training simulator are described, aiming at maximal cost-effectiveness.

Functional specifications are defined as simulator characteristics on an intermediate level; for example, the human–machine interface components required and/or the field of view of the out-of-the-window views rendered by the display system. This should be done without exactly specifying how (by what technical means or components) these requirements should be attained (Korteling et al., 2011). The TNA has resulted in a list of required competencies and competence levels (criteria for task behaviour and simulation capabilities). The analysis at competence level should aim at the description of the minimal set of tasks or subtasks in relation to relevant training scenarios to be simulated, which are necessary for training all relevant competencies. For each (sub-) task identified in the TNA a set of functional specifications (global simulator characteristics) has to be provided. In order to specify a cost-effective simulator in this way, a number of cost-related factors have to be considered. Following are some examples:

- Trainee level: initial training usually requires less fidelity, or realism, and thus less cost, than the training of experienced trainees.
- Type of task: it is usually more difficult (expensive) to achieve physical fidelity in simulators for perceptual motor tasks than for cognitive tasks. In addition, procedural tasks are relatively easier (and cheaper) to simulate with high fidelity.
- Part-task training: selectively focusing on those task variables that can be easily trained for on a simulator and with high training effectiveness may be crucial for cost-effective simulation.
- Level of fidelity: in practice, the desired level of fidelity should be based on a cost-benefit trade-off analysis. Achieving extra fidelity involves costs that should not exceed the benefit of higher transfer and/or efficiency of training.

Full-fidelity simulation is, in many cases, not required. The methods described in GM-VV should be applied in this phase to decide on the fidelity required to meet a certain training need. This leads to technical criteria and (quick-and-dirty) methods for their measurement. The following V&V activities can be performed here:

- Derive criteria from the high-level needs and put them in the argumentation structure
- Check each alternative on a coarse-grained level to see if there are showstoppers (i.e., an M&S system clearly fails an important criterion).
- Derive and check quality criteria related to the Phase-B processes and roles.

3.3 DMP C Phase: More Detailed Specifications

3.3.1 Technological analysis and technical specifications

In the functional specifications laid out in the previous step, the global simulator characteristics and capabilities needed to train the required competences are specified, but not what type of display system or what specific product would be best. This step explains how to specify these technical requirements. This is done in three parts: First, identify simulator subsystems and related technologies. Second, investigate which technologies and products are on the market and investigate their (technical) performance characteristics. Finally, identify the cost of technologies and products and the possibilities for cost-savings. This might include considering co-operations with other branches of the defence organization or with partner nations on scenario or database developments, etc.

A technological inventory shows the major training platforms, engines, and subsystems that are available, their performance characteristics, and their lifecycle costs. Note that this inventory may be re-used in future acquisition projects or could be shared with partner nations.

Where needed, the high-level criteria in the argumentation structure can be refined and more specific tests may be defined to uncover showstoppers for some alternatives. Again, process- and role-oriented quality criteria may be added and checked.

3.3.2 Cost-effectiveness trade-off analysis and training system specification

The final step is the specification of an optimal simulator configuration from a cost-effectiveness point of view. This is the result of a trade-off between the minimal necessary simulator components and reasonably high training benefits. In order to acquire such a simulator, the following activities have to be undertaken:

- For using simulation to train each subtask, estimate the costs of the technological requirements.
- Select those subtasks which require simulator-based training because they cannot be trained in a cost-effective way without simulators (high cost, dangerous, environmental restrictions).
- Select those subtasks that may benefit substantially from the potential advantages of simulation (e.g., attractiveness, flexible scenario generation, authentic training, and guided-discovery learning); and/or select those subtasks which require a considerable amount of conventional training (which should be known if the task domain is not completely new).
- Eliminate those subtasks that fulfil relatively minor training needs and require complex technological components (high costs); and/or eliminate those subtasks that lead to complex (expensive) technological requirements, which can be trained effectively by conventional means, such as exercises, classroom settings, books, videos, etc.

Based on the remaining subtasks a cost-effective simulator can be (globally) specified. This can be done by aggregating the remaining subtasks and fitting them into a coherent simulator. In order to produce a complete simulator, it may be necessary to include simulator characteristics or components to train subtasks that were not selected initially.

In the V&V argumentation structure, the coverage of the various elements of a complete training system design, including one of the alternatives M&S-based training systems, can be indicated (**Figure 5**). From this coverage, it becomes clear what parts of the purpose are not (fully) supported. The costs may also be estimated for adaptations, such that all purposes are fully supported. This can be done for all alternative training systems (Figure 6), in order to compare them.

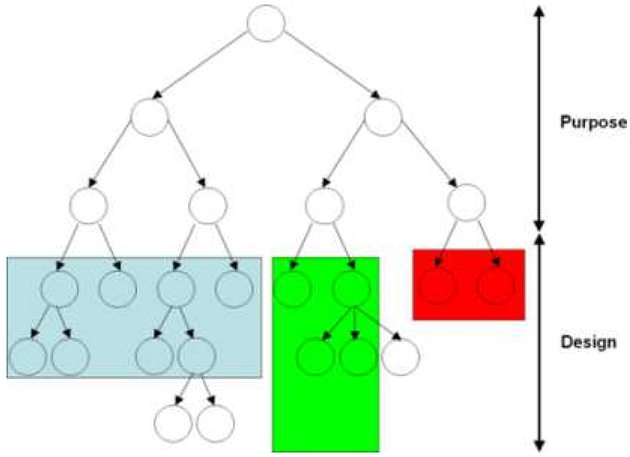

Figure 5: Argumentation structure coverage of a complete training system design.

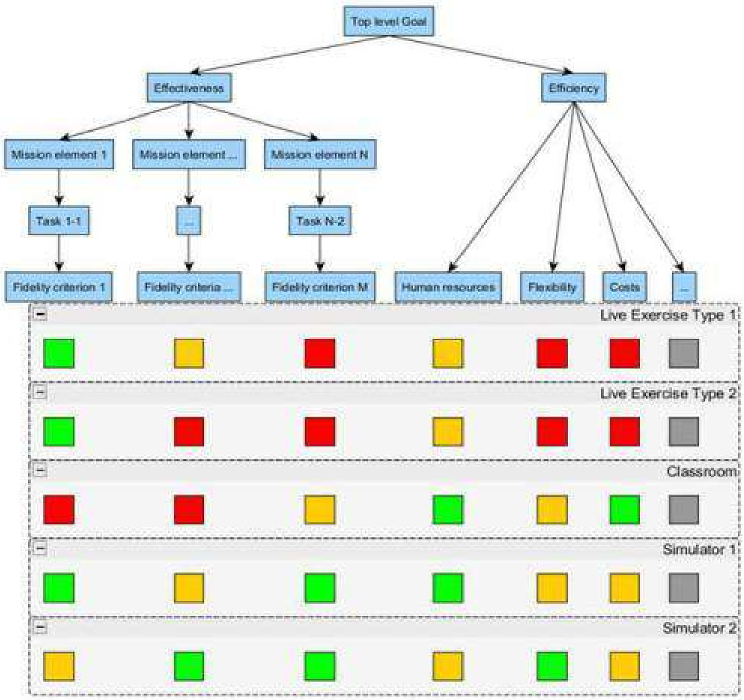

Figure 6: Visually presenting the comparison of alternatives based on effectiveness and efficiency

3.4 DMP D Phase: Selection

The utility, validity, and correctness criteria discussed above are derived and used in the process of choosing the most effective M&S asset from a number of alternatives or candidates. In general, the effectiveness values required are determined in the first few DMP phases, and they are used in the phase where simulation assets are chosen.

The problem is that this overall effectiveness is in general not always attainable because of overall limitations. The available budget is a clear example of a limit that forces choices in components that drive the overall effectiveness down. Asset availability time, needed user expertise and, especially in military applications, security issues, can also become showstoppers and place constraints on the acceptable solutions.

This might mean that some aspects of a simulation will be below the determined effectiveness limit, while other aspects score above this limit. In that case, the aspects that fall below standard will prevent the customer’s overall purpose from being met. The simulation, however, might still be suited for some parts of the customer's purpose. The forced choices during the development of the simulation can be made such that the best possible effectiveness can be reached. All choices where the utility falls below the effectiveness must be recorded and communicated back to the customer and users as limits on the original purpose.

Optimal effectiveness is obtained when all relevant factors are taken into account and the negative impact on effectiveness of forced choices for components that score below minimum effectiveness are minimized. The impact of the influencing factors may differ. This impact must be derived from the customer's purpose and the contribution to that purpose. The estimated risk of using a component that does not score above the effectiveness value is an important issue to take into account.

The sections above discuss the ways in which effectiveness is influenced in practice and how effectiveness criteria can be established. In order to decide which assets to use in a simulation, it is necessary to know how many resources are involved in changing (upgrading and possibly downgrading) these assets. Then, as described in Section 3.3, each alternative plus the costs associated with changes and the risks of non-effectiveness must be taken into account, such that the most optimal combination of assets and changes can be determined.

4.0 EARLY CONCLUSIONS AND ROAD AHEAD

4.1 Early conclusions

Acquiring cost-effective simulation systems for training applications is a challenge that must be addressed because of increasing demands for joint and collective training in times of shrinking budgets and limited available resources. The acquisition processes currently used for military systems seem more fitting for hardware needs that are specified in great detail and leave little room for negotiation. In the case of training simulation systems, however, there is often a range of possible solutions that allow utility trade-offs, which lead to trade-offs on effectiveness, efficiency (costs), and risk, which are translated into choices for fidelity and correctness.

The selection process should start with a detailed TNA. Validation methodology can then be applied to analyze how the training needs are to be weighted with respect to each other. Training effectiveness estimations will help assessing the degree to which a certain training simulation system meets the client’s needs. The combined outcome of these analyses will then allow the selection of the most effective solution.

It would be counter-productive to expect a "one size fits all" solution in the near future, but this paper proposes an approach that could be taken to improve the situation. The discussions should help refine a conceptual model of how to integrate TNA, VV&A, and cost-effectiveness analysis into the acquisition process. This is a starting point for developing and evaluating formal processes and technical solutions in support of the process.

4.2 Road ahead

The integrated overall process, consisting of the acquisition, VV&A, and CEA processes should also be effective and efficient. The three different perspectives can each be described as processes consisting of activities, which are executed by roles, and which result in certain products. The products must be defined, such that the information contained in them that is used by more than one process, can be re-used effectively and efficiently. Based on this analysis, an overlap in activities and roles may be identified and removed.

A different combination of activities is required for each situation. At the beginning of the acquisition, an overall process must be defined for the situation at hand that integrates all three processes as effectively and efficiently as possible.

A number of questions remain to be answered regarding CEA. What are the major cost drivers, for example? How can effectiveness of training be estimated and compared for different forms of training?[1] How can historical data be used? What are the opportunities and consequences of using Commercial Off The Shelve (COTS)?

The authors hope to further develop these ideas in close cooperation with MoD procurement organizations. A proposed study on this topic is currently under review. The study will also engage with the wider NATO RTO community, in particular SAS panel and NMSG, to contribute to this investigation. The addressed challenges are obviously not limited to the military domain. Many businesses and organizations are facing similar issues. In this respect, the ongoing research within the Netherlands on ''Topsector Creative Technologies'' may lead to some useful new ideas and civil-military co-operations.

REFERENCES

[1] Cohn, J. (2009). Building Virtual Environment Training Systems for Success. In: D. Schmorrow, J. Cohn, D. Nicholson (Eds). *The PSI Handbook of Virtual Environments for Training and Education.* Vol, pp 193-207.

[2] Emmerik, M.L., Korteling, J.E. (in press). *Optimal use of synthetic training systems for the Defence*. Report TNO-DV (to be published), Soesterberg: TNO Behavioral and Societal Sciences

[3] Farmer, E. Rooij, J. van; Riemersma, J., Jorna, P., Maraal, J. (1999). *Handbook of simulator-based training*, Brookfield, USA: Ashgate; Cambridge, GB: University Press.

[4] Janssens, J. P. (2012). *Vision Modeling and Simulation CLAS*. (in Dutch) Netherlands Ministry of Defence Amersfoort: Simulation Centre - OTC Manoevre.

[5] Korteling,J.E., Baeyer, A. von, Helsdingen, A.S., (2001). *Handbook of low-cost simulation for military training*, Grenoble, France: Corys TESS.

[1] See Chapter Korteling, Oprins and Kallen

[6] Korteling,J.E, Helsdingen, A.S., Sluimer, R.R., Emmerik, M.L., Kappé, B. (2011). *Transfer of Gaming: transfer of training in serious gaming*. Report TNO-DV B142, Soesterberg TNO Behavioural and Societal Sciences.

[7] Milham, L., Bell Carrol, M., Stanney, K. and Becker, W. (2009). Training system requirements analysis. In: D. Schmorrow, J. Cohn, D. Nicholson (Eds) *The PSI Handbook of Virtual Environments for Training and Education*. Vol 1, pp 165-192.

[8] NATO SAS Task Group-069 (2009). *Code of practice for life cycle costing* . Paris: RTO NATO

[9] Q-tility, 2012, Q-tility Verification and Validation Solutions, http://www.q-tility.nl

[10] Roza, M, Voogd, J.M., Sebalj, D. (2012) The Generic Methodology for Verification and Validation to support acceptance of models, simulations and data, Journal of Defense Modeling and Simulation: Applications, Methodology, Technology 1–19, DOI: 10.1177/1548512912459688

[11] SISO (2012) SISO Guide: GM-VV Vol. 1: Introduction and Overview, SISO-GUIDE-001.1-2012, SISO

[12] SISO (to be published) SISO Guide: GM-VV Vol. 2: Implementation Guide, SISO-GUIDE-001.2-2012, to be published, SISO

[13] Stanney, K.(2009). Requirements Analysis: section perspecive. In: D. Schmorrow, J. Cohn, D. Nicholson (Eds) *The PSI Handbook of Virtual Environments for Training and Education*. Vol 1, pp 115-130.

[14] Verstegen, D.M.L. (2004*). Handbook GOLM need specification method*. (in Dutch) Report 04-TM-A024, Soesterberg: TNO Human Factors.

[15] Voogd, J.M., Lemmers, A., Gerretsen, A. Roza, M., Karelse, R. (2009). Effective Use of Simulation Means in Collective Mission Simulation, *Proceedings IITSEC 2009*, Paper 9198.

[16] Young, R.R. (2004). *The Requirements Engineering Handbook*. Boston: Artech House

[17] Young R.R.(2001). *Effective requirements practices*. Boston: Addison-Wesley.

[18] NL MoD (2007) *Overview of the Defence Materiel Process*, Netherlands Ministry of Defence.

[19] Cohn, J., Kay, S., Milham, L., Bell Carroll, M., Jones, D., Sullivan, J., & Darken, R. (2009). Training effectivenesss evaluation: from theory to practice. In: D. Schmorrow, J. Cohn, D. Nicholson (Eds) *The PSI Handbook of Virtual Environments for Training and Education*. Vo l3, pp 157-172.

AUTHOR BIOGRAPHIES

WIM HUISKAMP is Chief Scientist Modelling, Simulation and Gaming in the M&S department at TNO Defence, Security and Safety in the Netherlands. Wim leads TNO's research programme on Live, Virtual and Constructive Simulation, which is carried out on behalf of the Dutch MOD. Wim is a member of the NATO Modelling and Simulation Group (NMSG), serving as its vice-chair since 2012. He has also acted as member and chairman in several NMSG Technical Working groups. He is currently co-chair of MSG-080, Chairman of the NMSG M&S Standards Subgroup (MS3) and he is the liaison of the NMSG to the Simulation Interoperability Standards Organization SISO.

J.E. (HANS) KORTELING received his B.Sc. and M.Sc. degrees in psychology from the University of Amsterdam (1985, cum laude), where his main interests included visual perception and psychophysiology. Until 1998 he worked as a research psychologist in the Traffic Behavior Department and in the Skilled Performance Department and as research group leader in the Vehicle Control Group of TNO Human Factors, the Netherlands. In this period Hans received his PhD. with a thesis on Multiple-task Performance and Aging (1994, cum laude). Since the beginning of 1998, Hans was in charge of the Simulation Research Group of TNO Human Factors which focused on the specification and validation of research- and training simulators. Since 2002 he has been working as program manager of Training Research for the Armed Forces. From 2005 Hans is also senior scientist of the Training & Performance Innovations Department of TNO Human Factors. At present his primary research interests include, simulator validation research, serious gaming, cognitive neuroscience, and (cognitive) aging / learning for life.

JEROEN VOOGD is a member of the scientific staff in the Defence, Security and Safety Division at TNO. He holds a Ph.D. (1998) in Computational Physics from the University of Amsterdam in the field of modelling and simulating of biophysical systems on parallel and distributed computing platforms. A recurring theme in his work of the last years is the quality of simulations. This includes VV&A of simulator assets, as well as quality assurance within TNO. He is also one of the Technical Representatives of Q-tility: an organisation specialised in verification and validation solutions for models, simulations and serious-games.