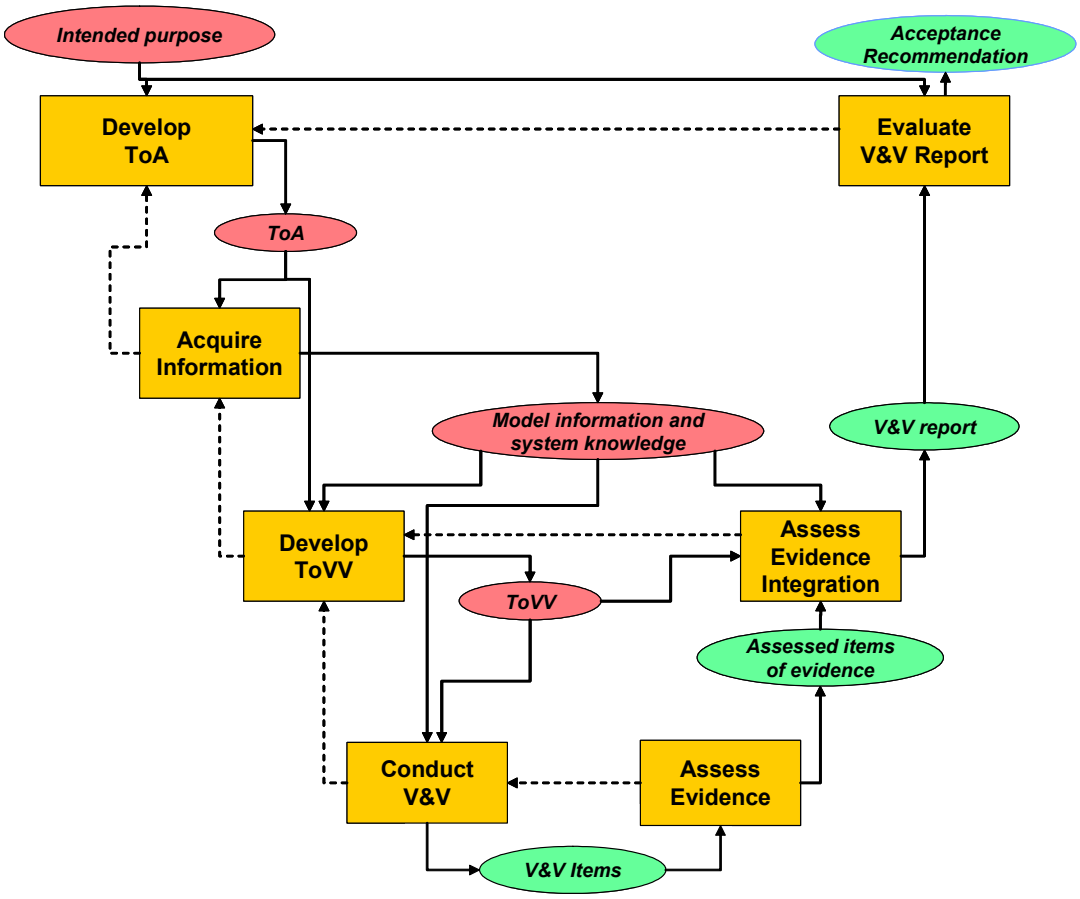
FOI-R--1379--SE
December 2004
ISSN 1650-1942
**Metodrapport**

Choong-ho Yi, Dirk Brade

# Generisk VV&A process: En VV&A handbok från THALES JP 11.20




Systemteknik
172 90 Stockholm

TOTALFÖRSVARETS FORSKNINGSINSTITUT
Systemteknik
172 90 Stockholm

FOI-R--1379--SE
December 2004
ISSN 1650-1942
**Metodrapport**

Choong-ho Yi, Dirk Brade

# Generisk VV&A process: En VV&A handbok från THALES JP 11.20

| Utgivare | Rapportnummer, ISRN | Klassificering |
|---|---|---|
| Totalförsvarets Forskningsinstitut - FOI | FOI-R--1379--SE | Metodrapport |
| Systemteknik | **Forskningsområde** | |
| 172 90 Stockholm | 2. Operationsanalys, modellering och simulering | |
| | **Månad, år** | **Projektnummer** |
| | December 2004 | E6032 |
| | **Delområde** | |
| | 21 Modellering och simulering | |
| | **Delområde 2** | |



| Författare/redaktör | | Projektledare |
|---|---|---|
| Choong-ho Yi | FOI | Choong-ho Yi |
| Dirk Brade | KTH | **Godkänd av** Monica Dahlén |
| | | **Uppdragsgivare/kundbeteckning** FM |
| | | **Tekniskt och/eller vetenskapligt ansvarig** Choong-ho Yi |


**Rapportens titel**

Generis VV&A process: En VV&A handbok från THALES JP 11.20


**Sammanfattning (högst 200 ord)**

Under mars 2003 – augusti 2004 har FOI-projektet ”VV&A (verifiering, validering och ackreditering) av simuleringsmodeller” medverkat i ett europeiskt samarbetsprojekt THALES JP 11.20 (förkortat till JP 11.20) ”Common validation, verification and accreditation framework for simulation”, inom CEPA (Common European Priority Area), WEAG (Western European Armaments Group). Syftet med JP 11.20 var att ta fram en teknisk plattform för en gemensam europeisk VV&A metodologi. Rapporten presenterar resultatet från detta samarbete genom att bifoga en av de två JP 11.20 huvudrapporterna i sin helhet som bilaga.

I den bifogade JP 11.20 rapporten presenteras en generisk VV&A process som en ”User’s Manual”, en lättförståelig handbok för användare, d.v.s. de som utför VV&A arbeten. Processen består av ett flöde aktiviteter och produkter från dessa aktiviteter, och tillhandahåller en vägledning för användare vid planering och genomförande av VV&A enligt JP 11.20 metodologin.



**Nyckelord**

VV&A (verifiering, validering och ackreditering) process, VV&A handbook, gemensam europeisk VV&A metodologi, THALES JP 11.20, CEPA, WEAG

| Övriga bibliografiska uppgifter | Språk Svenska |
|---|---|
| **ISSN** 1650-1942 | **Antal sidor:** 79 s. |
| **Distribution enligt missiv** | **Pris: Enligt prislista** |

| Issuing organization | Report number, ISRN | Report type |
|---|---|---|
| FOI – Swedish Defence Research Agency<br>Systems Technology<br>SE-172 90 Stockholm | FOI-R--1379--SE | Methodology report |
| | **Programme Areas**<br>2. Operational Research, Modelling and Simulation | |
| | **Month year**<br>December 2004 | **Project no.**<br>E6032 |
| | **Subcategories**<br>21 Modelling and Simulation | |
| | **Subcategories 2** | |



| Author/s (editor/s) | | Project manager |
|---|---|---|
| Choong-ho Yi | FOI | Choong-ho Yi |
| Dirk Brade | KTH | **Approved by**<br>Monica Dahlén |
| | | **Sponsoring agency**<br>Swedish Armed Forces |
| | | **Scientifically and technically responsible**<br>Choong-ho Yi |


**Report title (In translation)**

Generic VV&A process: THALES JP 11.20 VV&A User's manual


**Abstract (not more than 200 words)**

During March 2003 to August 2004 the FOI-project ”VV&A (verification, validation and accreditation) of simulation modells” has participated in an European cooperation THALES JP 11.20, “Common validation, verification and accreditation framework for simulation”, within CEPA (Common European Priority Area), WEAG (Western European Armaments Group). The purpose of JP 11.20 was to develop a technical platform for a common European VV&A methodology. This report presents the result from the cooperation by enclosing one of the two JP 11.20 main reports in it whole as an appendix.

In the enclosed JP 11.20 report a generic VV&A process is presented as a User’s Manual which is intended to be easy to understand for the user, i.e. those who conduct VV&A. The process consists of a flow of activities and products that are generated from the activities, and enables the user to plan and conduct a VV&A endeavour according to the JP 11.20 methodology.



**Keywords**

VV&A (verification, validation and accreditation) process, VV&A User’s Manual, Common European VV&A methodology, THALES JP 11.20, CEPA, WEAG

| Further bibliographic information | Language Swedish |
|---|---|
| **ISSN** 1650-1942 | **Pages** 79 p. |
| | **Price acc. to pricelist** |

# Innehållsförteckning

# 1. Bakgrund

Det är en genomgående trend att världens försvarsmakter alltmer använder sig av modellering och simulering (M&S) som stöd för sin verksamhet. Exempelvis ersätts övningar och prov alltmer av simuleringar p.g.a. ekonomiska skäl. Vidare kan man med hjälp av simuleringsmodeller undersöka mer komplicerade förlopp som annars inte kunnat studeras p.g.a. risker för människor och miljö. Dessutom skapar den tekniska utvecklingen inom M&S nya möjligheter hela tiden. Med tanke på dessa möjligheter som M&S kan erbjuda och de ekonomiska fördelarna med M&S i kombination med minskande försvarsbudget, kan man dra slutsatsen att vikten av M&S kommer att fortsätta att öka.

Men, hur vet vi då att de modeller som används inom FM är trovärdiga för sina syften? Användning av modeller som inte är trovärdiga kan leda till katastrofala konsekvenser. Exempelvis kan soldaterna mista livet i strid om de tränats ”fel” i en simulerad miljö som inte är lämplig för det aktuella syftet. VV&A (verifiering, validering och ackreditering) är en process som syftar till att belysa trovärdigheten (korrekt och lämplig) för en simuleringsmodell för ett givet syfte. Verifiering är en process som avgör om en modell utvecklas korrekt enligt specifikationer; validering är en process som avgör om en modell är lämplig för ett syfte; ackreditering är ett officiellt bemyndigande att en modell får användas i ett visst syfte.

Under mars 2003 – augusti 2004 har FOI-projektet ”VV&A av simuleringsmodeller” medverkat i ett europeiskt samarbetsprojekt THALES JP 11.20 (förkortat till JP 11.20) inom WEAG:s CEPA (Common European Priority Area) med namnet ”Common validation, verification and accreditation framework for simulation”. Rapporten presenterar resultatet från detta samarbete genom att bifoga en av de två JP 11.20 huvudrapporterna i sin helhet som bilaga. Den bifogade huvudrapporten grundar sig på och integrerar resultaten från de över trettio tekniska rapporter som producerats inom JP 11.20, och bedöms vara mest lämplig för presentation av JP 11.20 resultatet. Som inledning för denna bifogade rapport ger vi en kort beskrivning av JP 11.20 i allmänhet (syfte, konsortiet och innehåll), den del (arbetspaket) av JP 11.20 som FOI var ansvarig för, och ett annat arbetspaket som de huvudrapporterna tillhör.

# 2. THALES JP 11.20

Syftet med JP 11.20 är att *ta fram en teknisk plattform för en gemensam europeisk VV&A metodologi* [1]. Idag finns det ingen sådan gemensam VV&A metodologi. VV&A-arbeten utförs olika i olika organisationer och olika länder med egna metoder, processer, och policies. Följaktligen blir det svårt för en utomstående, d.v.s. annan organisation eller annat land än det som utvecklat modellen själv, att bedöma modellens trovärdighet. Det försvårar eller kanske till och med förhindrar samarbeten där VV&A ingår. Detta i sin tur utgör ett stort hinder för ackreditering och återanvändning av modeller.

## 2.1 Konsortiet

Fem länder har deltagit i detta samarbete: Frankrike (ONERA leder projektet), Italien (DATAMAT), Nederländerna (TNO), Danmark (UNI-C) och Sverige (FOI och FMV). JP 11.20 består av sju arbetspaket (WP, Work Package), och varje WP har tilldelats en ansvarig organisation [1]:

- WP0 Management (ONERA)

- WP1 Problem Analysis (FOI)
- WP2 Definition of Acceptance Target (TNO)
- WP3 VV&A Techniques (UNI-C)
- WP4 Optimisation of Testing (DATAMAT)
- WP5 Outline for a VV&A Methodological Framework (ONERA)
- WP6 Evaluation and Dissemination of the Results (DATAMAT)

Bland dessa arbetspaket presenterar vi nedan något mer detaljerat WP1 som FOI var ansvarig för och WP5 som är en integrering av resultaten från alla tekniska WP inom JP 11.20.

### 2.2 WP1 Problem Analysis

FOI ledde WP1, och detta arbetspaket består av fyra delar (Work Element, WE):

- WE1100: VV&A Global Taxonomy
- WE1200: VV&A Criteria Definition
- WE1300: VV&A Levels Definition
- WE1400: VV&A General Process

Syftet med WE1100 var att tillhandahålla i) en strukturerad och tydlig terminologi för internt bruk inom konsortiet, ii) taxonomier för att klassificera olika V&V-ansatser, och iii) en ordlista av de viktigaste och mest förekommande termerna för att bl.a. nå en konsortie-gemensam förståelse [2]. Exempelvis har ett V&V taxonomiträd föreslagits som ett hjälpmedel för VV&A beträffande ackrediteringstyper (ackreditering av en modell eller ackreditering av simuleringsresultat) och tillgänglig information om modellen och verkliga systemet. Trädet tar också hänsyn till att det krävs att modell, data och experiment (experimental frame) är korrekta och lämpliga för att få trovärdigt resultat från simuleringar.

VV&A handlar om utvärdering (och bedömning). För en objektiv utvärdering krävs väldefinierade kriterier. I WE1200 presenterades ett antal sådana kriterier som kan användas under V&V arbeten beträffande korrekthet och lämplighet av en modell, data och experiment [3]. Dessa kriterier kan användas även efter V&V processen för att utvärdera de utförda V&V aktiviteterna själva.

Även om en modell eller simuleringsresultat gått igenom en V&V process, kvarstår alltid viss osäkerhet beträffande deras korrekthet och lämplighet. Denna osäkerhet är beroende av hur rigoröst V&V arbeten har utförts. WE1300 definierade några V&V nivåer som kan användas som en rimlig indikation av kvarvarande osäkerhet [4].

En generisk VV&A process har introducerats i WE1400 genom att integrera resultaten huvudsakligen från andra WEs inom WP1 [5]. Denna process är tänkt att ge vägledning under planering och genomförande av VV&A aktiviteterna.

### 2.3 WP5 Outline for a VV&A Methodological Framework

Syftet med WP5 är att integrera resultaten från alla WP på ett konsistent och sammanhängande sätt till ett ”VV&A Methodological Framework”. WP5 besår av två WEs:

- WE5100: VV&A Methodological Guidelines
- WE5200: VV&A Process Specification

WE5200 beskriver detta metodramverk som en ”User’s Manual”, en lättförståelig handbok för användare som är V&V agenter, d.v.s. de som utför V&V arbeten. Se bilagan. Handboken, d.v.s. WE5200, beskriver bl.a. en generisk VV&A process som består av ett flöde aktiviteter och produkter från dessa aktiviteter. WE1400 från WP1 utgjorde en väsentlig grund för denna WE.

WE5100 är en ”Reference Manual” innehållande djupare diskussioner (definitioner, förklaringar) kring de begrepp och metoder som används i handboken [6]. Avsikten är att V&V agenter kan vända sig till denna manual om han/hon behöver mer information vid användning av handboken.

Några utmärkande drag av JP 11.20 metodramverket är:

- Att åtskilja olika frågeställningar från olika perspektiv, t.ex. från verkligheten, problemformuleringen, och modelleringen
- Generisk VV&A process
- Att stor vikt läggs på definition av V&V krav
- Att ge stöd för utvärdering och ackrediteringsbeslut

Förutom att WE5200 är en integrering av resultat från alla tekniska WP, ger denna WE en överblick över området på ett lättförståeligt sätt. Av dessa skäl har vi valt WE5200 för att presentera slutresultatet från JP 11.20.

## 3. Fortsättning

Resultaten från JP 11.20 är lovande. Men dessa resultat måste fördjupas och förfinas för att nå en internationell gemensam VV&A metodologi, vilket inte var möjligt under den korta projekttiden. Planering för ett treårigt fortsättningsprojekt pågår. Förhoppningen är att detta kan starta under våren 2005. Alla länder i JP 11.20 konsortiet har visat intresse för fortsatt medverkan. I dagsläget är det inte bestämt vilken MoU (Memorandum of Understanding), t.ex. THALES eller EUROPA, som kommer att gälla för fortsättningen.

## 4. Websites

Några av de tekniska rapporterna från JP 11.20, t.ex. huvudrapporterna WE5100 och WE5200 och rapporterna WE1100 – WE1400 från WP1, kommer att läggas upp på JP 11.20s hemsida www.jp1120-revva.com på Internet under december 2004. Här kan det också nämnas att FOI och FMV har en gemensam VV&A hemsida på Internet, www.vva.foi.se.

## Referenser


[1] Programme Definition Document, WEAG THALES JP 11.20, 2003.
[2] VV&A Global Taxonomy, WEAG THALES JP 11.20 Project Report D1101, 2004.
[3] VV&A Criteria Definition, WEAG THALES JP 11.20 Project Report D1201, 2004.
[4] VV&A Levels Definition, WEAG THALES JP 11.20 Project Report D1301, 2004.
[5] VV&A General Process, WEAG THALES JP 11.20 Project Report D1401, 2004.
[6] VV&A Methodological Guidelines: Reference Manual, WEAG THALES JP 11.20 Project Report D5103, 2004.

## Bilaga

**VV&A Process Specification (PROSPEC): “User’s Manual”, v1.3, WEAG THALES JP 11.20 Project Report D5201**

# THALES JP 11.20

## Common Validation, Verification and Accreditation Framework for Simulation
## REVVA

# VV&A Process Specification (PROSPEC)
# "User's Manual"
# v1.3

**Work Element:** WE 5200
**Prepared by:** ONERA/FOI/FMV/TNO

**Distribution:** CERT, DATAMAT, FMV, FOI, TNO, UNI-C, ALEKTO
**Approved for Release to:** JP 11.20 Consortium
JP 11.20 Management Group

| | Document Number | |
|---|---|---|
| V&V&A THALES JP11.20 | THALES JP11.20-WE5200-PROSPEC-D5201 | |
| | Version Date: 2004 December 15 | Number of Pages: 67 |

# EXECUTIVE SUMMARY

This document guides the reader through the REVVA Generic Process proposed by the THALES JP11.20 Technical Working Group ("REVVA"). It harmonises and interconnects the concepts and ideas that were proposed in earlier project reports, and enables the reader to plan and conduct a VV&A endeavour according to the REVVA methodology.

For this purpose, chapter 2 gives an introduction to the methodology, and reviews and summarises its underlying concepts, which are introduced in more detail in [METHGU2 2004]. Organisational aspects such as choosing appropriate actors to play roles involved in the VV&A endeavour are addressed in Chapter 3. Chapter 4 outlines the REVVA Generic Process, describes the flow of activities and products, and describes the products generated during VV&A. Finally, Chapter 5 provides guidance on how to perform the activities proposed for the individual phases, while Chapter 6 points out the benefits of experience capitalisation using repositories.

This report must be read in concert with [METHGU2 2004]. For background information please refer to the other technical project reports summarised in Chapter 7.

# REPORT DOCUMENTATION PAGE

| | |
|---|---|
| 1. Report number as used by Company responsible for document<br>THALES JP 11.20-WE5200-PROSPEC-D5201 | |
| 2. Name and Address of Company responsible for document<br>ONERA | |
| 3. WEU Contract number<br>N/A | |
| 4. Name and Address of Contracting Authority<br>WEAO Research Cell, Rue de la Régence, 4,<br>B-1000 Bruxelles | |
| 5. Name and Address of SLIE<br>René Jacquart, CERT, ONERA | |
| 6. Report Classification and Caveats on use<br>WEU Unclassified | 7. Date of first issue<br>Mai 30th, 2004 |
| 8. Document title<br>Users Manual (PROSPEC) | |
| 9. Author(s)/Company<br>Dirk Brade (FOI/FMV), Jeroen Voogd (TNO) | |
| 10. Descriptors, Key words<br>JP 11.20, WE5200, User's Guide | |
| 11. Abstract (Must be WEU Unclassified)<br>See EXECUTIVE SUMMARY | |

# CHANGE RECORD

| VERSION | DESCRIPTION |
|---|---|
| **1.0** | *Initial release* |
| **1.1** | *MG comments v1 integrated* |
| **1.2** | *MG comments v2 integrated* |
| **1.3** | *Clarification on residual uncertainty, convincing force, and probative force, pages 12 and 13* |

| **Version** | **First Issue** | **1** | **2** | **3** | **4** | **5** | **6** |
|---|---|---|---|---|---|---|---|
| **Date** | | 2004 Nov 25 | 2004 Dec 10 | 2004 Dec 15 | | | |
| **Written by** | Jeroen Voogd, Dirk Brade | Dirk Brade | Dirk Brade | Dirk Brade | | | |
| **Checked by** | Fredrik jonsson, Håkan Lågerström, René Jacquart | Choong-ho Yi | Choong-ho Yi | Choong-ho Yi | | | |
| **Approved by** | | | | | | | |

# TABLE OF CONTENTS

# 1 INTRODUCTION

This document is intended to guide the reader through planning and implementing verification, validation, and accreditation (VV&A) of a Modelling and Simulation (M&S) product, such as a simulation model or simulation results. It does neither discuss the background and context of the presented concepts, nor does it point out the variety and complexity of any V&V effort, but rather summarises, compiles, and harmonises the results, which have been completed during the earlier project stages and are documented in detail in the associated technical reports (see section 5). This document guides those acting as Acceptance Leader and V&V Leader through preparing, planning, conducting, and evaluating a VV&A endeavour, and focuses on the assessment of an M&S product. Those who take most benefit from this document being used are the customers, i.e., who acquire the M&S product for a particular intended purpose. After reading this document, one should have a clear idea of the activities and products in the REVVA Generic Process and of the dependencies between its enabling concepts, and should have developed an understanding of the essential elements of the WEAG THALES JP11.20 project.

## 1.1 Work Element Overview

The purpose of WE5200 was to develop a report on the methodology, which – in essence – describes how to apply it in practice. This report constitutes the “User’s Manual” of the methodology, and describes the process followed by the actors in their various roles. During previous study various options have been considered, which now consolidate into one comprehensive approach.

## 1.2 Abbreviations

| | |
|---|---|
| AC | Acceptability Criterion |
| EF | Experimental Frame |
| IoE | Item of Evidence |
| JP | Joint Programme |

| | |
|---|---|
| M&S | Modelling and Simulation |
| MoE | Measure of Effectiveness |
| REVVA | THALES JP11.20 nickname: “Reference for VV&A” |
| SEF | Simulation Experimental Frame |
| SEM | Simulation Executable Model |
| SME | Subject Matter Expert |
| SoI | System of Interest |
| ToA | Target of Acceptance |
| ToVV | Target of Verification and Validation |
| V&V | Verification and Validation |
| VV&A | Verification, Validation, and Accreditation |
| WEAG | Western European Armament Group |

## 1.3 Glossary

The glossary used is provided as appendix of [METHGU2 2004].

# 2 PROCESS FOUNDATIONS

As elaborated in more depth in [METHGU2 2004], the methodology proposes

- an organisation,
- a process,
- a set of products, and
- some hints, recommendations, and guiding principles.

In this section the reader gets an overview over the building blocks of the methodology and the dependencies between the contents described in the project reports referenced in section 7. Those reports explain the background of the individual research issues and motivate the achieved results. To harmonise the results here, slight modifications were done, but the essence of the concepts as documented in the individual work element reports did not change.

## 2.1 Underlying Concepts and Terminology

As indicated by several definitions [Schmidt 1985; Shannon 1975] of *computer-based simulation*, an *executable quantitative symbolic model* is exercised with *initialisation and runtime input data* within a well defined *experimental frame (EF)*, to allow the observation of model behaviour over time. The experimental frame must be designed in such a manner that it allows achieving a set of *objectives*, which originate from the *context* of simulation. To clearly demark the experimental frame of a simulation experiment from the experimental frame of a real experiment with an existing system, it will be referred to as *simulation experimental frame* (SEF) in the following. Within the SEF the control parameters with associated value ranges constitute the “input” over time to the simulation model, and all “output” from the model is recorded as values over time of the goal parameters. The observed and appropriately post-processed model behaviour yields the simulation results. This concept is visualised in Figure 1.

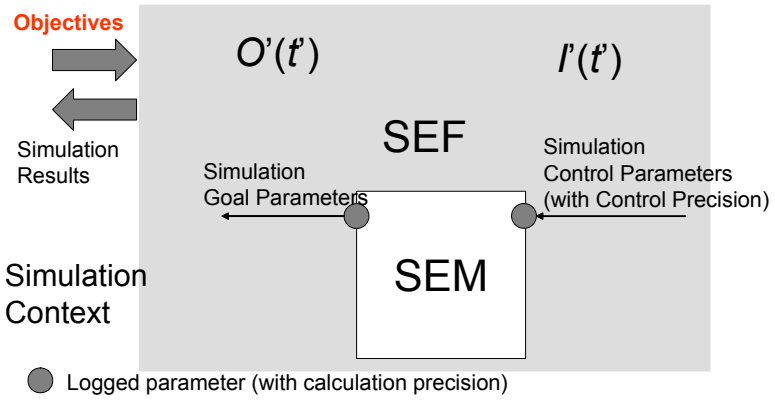



**Figure 1: Dependencies between the contextual objectives, the simulation experimental frame, the Simulation Executable Model, and the derived simulation results**

Whether the observed model behaviour approximates the behaviour of the associated System of Interest (represented as, e.g., system data, theoretic models, postulates, or expert opinion) depends on

- the similarity of chosen experimental frames of both Simulation Executable Model (SEM) and the System of Interest (SoI), including the identified influences on the behaviour of the System of Interest (*initialisation data*, *runtime input data*, and *embedded data*);

- the conducted idealisation and abstraction of the System of Interest (model).

Ultimately, the conclusions drawn from experimentation with the System of Interest (if it was available) and simulation under the same contextual objectives must not deviate in such a manner that unacceptable harm is caused. In the context of this document, the concept of **validation** answers the question of whether it is impossible to distinguish the model and system in the experimental frame of interest [Zeigler 2000]. It addresses the question whether the M&S product captures the system behaviour to the extent demanded by the objectives of the study (and not further). For this purpose, we distinguish the property of validity (which is inherent to the SEM, but – at least currently – cannot be proven), and the process of validation, i.e., of making plausible that a SEM has this desired property.

- Validity: The property of a simulation model to have, within a specific experimental frame, a behaviour which is indistinguishable from the behaviour of the System of Interest[1].

- Validation: The process which is used to construct, under a set of time, cost, skills, and organisational constraints a justified belief about model validity.

To technically support validation, model **verification** deals with the demonstration that a model is *correctly* represented and was transformed correctly from one representation form into another, according to all transformation and representation rules, requirements, and constraints [Brade 2004]. We also distinguish the inherent property of correctness from the process of demonstrating correctness.

- Correctness: The property of a simulation model to comply with formal rules and bodies of reference information for its content and representation, and for the transformation into another representation.

- Verification: The process which is used to construct, under a set of time, cost, skills, and organisational constraints a justified belief about model correctness.

This separation stresses the current situation that V&V cannot guarantee absolute correctness and validity: In practice, today there is neither absolute verification nor absolute validation of a non-trivial model, thus, there always remains some *residual uncertainty* concerning its factual validity and correctness. When founding a V&V approach on *Items of Evidence*, which are made available during the process of V&V and are supposed to substantiate a set of Acceptability Criteria, the overall residual uncertainty concerning the simulation model's fitness for purpose is introduced by the uncertainty associated with the strength of the individual items of evidence (expressed

[1] Please note that "indistinguishableness" and "equality" are not identical concepts. Two items, which are indistinguishable, are not necessarily equal – if one decides to look at them more closely, the items may become distinguishable, when the existing inequality is revealed.

as their *probative forces*), and the uncertainty associated with the necessity and sufficiency as well of the referenced items of evidence (expressed as *convincing forces* of the individual arguments), as of the Acceptability Criteria themselves. The degree of residual uncertainty depends on the convincing force of the approach taken to the demonstration of correctness and validity, and the probative force of the individual Items of Evidence[2]. Even in demonstrating that a model or simulation results are invalid or incorrect there sometimes is uncertainty.

Table 1 visualises the uncertainty associated with the decision to accept or reject a SEM for a particular intended purpose in form of type I and type II error.

**Table 1: Factual validity vs. perceived validity**

| **Unknown fact** / **Perception (Action)** | **Factually valid** | **Factually invalid** |
|---|---|---|
| **Perceived as valid (and accepted)** | Ok | Type II Error; $\beta$ |
| **Perceived as invalid (and rejected)** | Type I Error "False Alarm"; $\alpha$ | Ok |

To clarify the concepts of residual uncertainty associated with verification and validation, we introduce the terms *factual correctness* and *factual validity*. Both V&V do not influence the factual correctness and factual validity of an M&S product, but only allow perceiving them through the activities within the process (*perceived correctness* and *perceived validity*). The more thoroughly the V&V is conducted, the better the perceived correctness and validity approximate the factual correctness and validity, and the lower becomes the upper bound for the type I or the type II error, respectively. With an estimate of these upper bounds available, the acceptance for use (acceptance decision)

[2] For the rigorous formalization of residual uncertainty, convincing force, probative force, and their dependency quantitative metrics such as probability or possibility are required, which are not yet included in the methodology. The approach taken here is based on linguistic variables and a verbal expression of dependencies (see sections 5.7). It fill be subjected to further formalization during the REVVA follow-on program.

can be based on the 2-tuplet “perceived validity and correctness” and the “residual uncertainty associated with the perception”. However, the current state of the art does not allow the reasonably objective measurement or estimation of the uncertainty associated with V&V efforts, thus, all levels of probative force, convincing force, and residual uncertainty used in this document (see section 5.5 and 5.6) are qualitative and often highly subjective. They are intended to aid decision making, but due to their subjectivity always require trust in the competence and honesty of the people involved.

In the defence VV&A community, **accreditation** is defined as “the official certification that a model, simulation, or federation of models and simulations is acceptable for use for a specific purpose” [DMSO 1996]. However, this definition is not consistent with the use of the term “accreditation” in other domains, where accreditation is not associated with products, but organisations [Rae, Robert, and Hausen 1995]. Also the official authoritative accreditation procedures vary from nation to nation. To avoid conflicts, the concept of **acceptability** for the intended purpose is introduced, assuming that acceptability is an indispensable prerequisite for accreditation or certification, whatever it is called.

Acceptance or rejection here is judged based on *Acceptability Criteria*. Similar to the software domain, different types of Acceptability Criteria are distinguished, with the most common classification into functional Acceptability Criteria (e.g., “it must be possible to print the document”) vs. non-functional Acceptability Criteria (e.g., “the software must run under MyOS”). For executable models (simulation models implemented in software, which can be executed on a computer) functional Acceptability Criteria can be further distinguished as related to the representation of the System of Interest and others. Due to their importance, here Acceptability Criteria addressing the behaviour of the simulation model and its indistinguishableness from the behaviour of the SoI are explicitly referred to as *Validation Criteria* (which are a distinguished subset of Acceptability Criteria). It further is assumed that all software quality related issues are covered by the appropriate software quality standards, and that the examination of “traditional” software Acceptability Criteria is covered by the appropriate test procedures. In the following, validation criteria exclusively address the M&S product’s

validity and correctness with respect to its chosen representation of the System of Interest; non-functional requirements and functional requirements not addressing the validity or correctness of the chosen representation of the real world are considered to be Acceptability Criteria.

In accordance with [Davis 1992], two types of acceptance are distinguished for JP11.20:

- **Model acceptance**, which needs to be associated with a range of intended uses (*domain of validity*) and can only be provisional;
- **Simulation results acceptance**, which includes the assessment of the used input data and the experimental frame for the execution of the experiments.

## 2.2 Scope of the Method

M&S are no goals by themselves, but must be understood as supporting element in a higher level task, and they are based on enabling technologies, which is visualised in Figure 2.

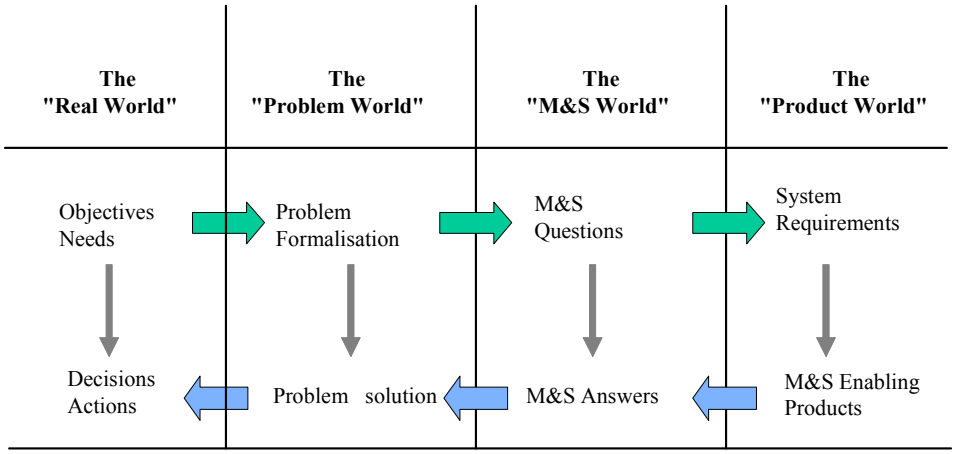



**Figure 2: A Perspective on the problem addressed**

The distinction of “worlds” helps to separate concerns: In the “Real World” a high level need may be to preserve the free democratic governments of Europe from military invasion, resulting in the action to maintain military forces, which are prepared for their mission. In the “Problem World” the task is located to keep those forces prepared, with

an appropriate training program as problem solution. How this training program can be supported by means of simulation is answered within the “M&S World”, providing as answers forces, which are fit to react on simulated threads. The technical equipment required to place the forces in the simulated combat situations is created within the “Product World”.

Figure 2 shows that M&S has a limited scope – and then a limited impact and responsibility – on decision making. It helps to clarify, as well, the distinction between M&S uses (the “M&S World”) and potential M&S developments, when they are needed (“Product World”). In section 3.2.2 the four worlds will help us to identify the actors which are appropriate to play particular roles.

V&V can be implemented most efficiently during the development of the M&S product, but the current situation in defence model development and use encourages the specification of approaches to post-developmental V&V (in the following: post-hoc V&V). Within JP11.20 due to pragmatic reasons, the focus lies on the post-hoc V&V. Managerial issues like budgeting, scheduling, and allocation of human resources have an important impact on success or failure of a V&V endeavour. This document concentrates on the technical aspects of V&V, but addresses managerial issues when necessary.

## 2.3 M&S Product-Oriented VV&A

The REVVA methodology is not bound to a particular modelling paradigm and independent from the chosen model development process or current state of model development. It exclusively concentrates on available intermediate, final, or supplemental M&S products, such as the Simulation Executable Model, or model documentation. Because it is most likely that especially for legacy M&S products there is no more process information available anyway, and because conclusions for product quality based on process assessment are unreliable, the M&S development process followed is not assessed during VV&A. For new M&S products this leaves maximum freedom to the developers concerning their internal procedures; however, they will be confronted with detailed documentation requirements, to ensure that all the information

that is required to perform efficient V&V and to reduce the residual uncertainty associated with the use of the M&S product is available.

# 3 ORGANISATION

Such as the development of a model and the execution of simulation experiments require numerous activities, V&V require the skill and knowledge from several areas, usually involving quite a few individuals. As an extreme, one single person could be responsible for the completion of all tasks during model development and V&V, but typically, different teams develop the model and perform V&V. In the following, a *role* will be characterised by the skills required to accomplish a particular task or set of tasks, and the responsibilities that are taken. There also needs to be distinguished between groups with different interests, including those who are going to acquire a simulation model or simulation results (and are likely to pay for it), and those who deliver the requested M&S product. These interest groups are called *parties* in the following. *Actors*, who have the required skill and are entitled to take the requested responsibilities play the roles, are members of parties.

You find this section at this early location in the document to enable you already now to find your party and role in the V&V endeavour. Because the REVVA Generic Process presented in section 4 is exhaustively referred to, we recommend that you fly over this section first and return for details after having read section 4.

## 3.1 Parties

A party is assumed to be an organisation or organisational unit. A typical situation given complex analysis problems is that a simulation model is developed and used for the computation of simulation results by the same party (e.g., thread simulation in a defence analysis department). The post-processed simulation results then are passed on to their military superiors, who use them for, e.g., decision making. Contrarily, for training applications the Simulation Executable Model provided by one party (e.g., an industrial aircraft provider who also builds air combat training simulators) is passed on to another party (e.g., a military pilots training centre), which wants to take benefit from executing the model. In both cases exists a “customer-supplier relationship”:

- *Customer*: A customer is an organisation or organisational unit which plans to use or is using an M&S product (such as a SEM, simulation results, or data) developed by another party. Customers are, for example, decision makers in the defence materiel acquisition administration, or training project leaders (but not the trainee). The customer may or may not have the appropriate actors to play the Acceptance Leader, the V&V Leader, and the V&V Executioners (see section 3.2).

- *Supplier*: The supplier is an organisation or organisational unit which provides the M&S product. This product may be simulation results, if the supplier was tasked to perform a simulation study, or an executable model, if the supplier was tasked to provide a model, which can be executed by others. The customer may or may not have the appropriate actors to play the V&V Leader and the V&V Executioners (see section 3.2).

As soon as a Customer-Supplier relationship includes contractual agreements and the flow of money, it is likely that there arises a conflict of interest between those parties. A relationship of trust between the customer and the supplier is desirable, but it must be always kept in mind that the supplier is trying to sell something to the customer, with all its implications. Thus, we introduce the

- *3rd Party VV&A Agent:* The 3rd Party VV&A Agent is an organisation or organisational unit external of the customer and the supplier. Its degree of independence is assessed based on managerial, technical, and financial factors. Within this organisation are appropriate actors available to play the roles of the Acceptance Leader, the V&V Leader, and the V&V Executioners (see section 3.2).

- *Acceptance Authority*: The Acceptance Authority is an organisation or organisational unit external of both the customer and the supplier, accredited to officially accept M&S products and trusted by the customer. Its degree of independence is assessed based on managerial, technical, and financial factors.

Within this organisation are appropriate actors available to copy the activities performed by the Acceptance Leader, the V&V Leader and V&V Executioners of any other party (see section 3.2).

All these parties may provide the actors to play the roles below.

## 3.2 The Roles

The assignment of tasks to persons or individuals (management of human resources) should be based on an agreement of the parties involved. In the following, roles interacting with and responsibilities within the VV&A process are identified. Each role is outlined by

- the required knowledge and skill to complete the assigned tasks,
- the authority given and responsibility taken in the process introduced in section 4, and
- its interaction with other roles.

A role does not determine, whether it is played by one actor, or shared by several actors, which even might come from different parties. However, particular roles require a sufficient distance between the individuals or teams performing them, while others are likely to be played by the same, single individual (see section 3.2.2).

### 3.2.1 VV&A Core Roles

VV&A core roles are directly involved in the VV&A endeavour by using, planning, conducting, evaluating, or assessing the substantial VV&A work.

- The *Contextual User*, who comes from the Problem World, defines the contextual objective as depicted in Figure 1. To achieve these objectives, the Contextual User uses M&S, i.e., substitutes a series of real experiments by simulation. To

become actively involved in the development of the Acceptability Criteria, this role should be able to express the contextual objectives precisely in language of the application domain. The Contextual User is mainly active during the first and the last phase of the REVVA Generic Process, expressing needs during phase 1, “Develop ToA”, and preparing or making the acceptance decision during phase 7, “Evaluate V&V Report”. It is assumed that the Contextual User always is in the customer party.

- The *Acceptance Leader* is a user representative (trusted by the Contextual User), who is responsible for the assessment of the M&S product and builds the bridge between the Problem World and the M&S World. The actor(s) in this role should be knowledgeable about requirements engineering, needs a deep understanding of the intended purpose, and supports the customer in transforming it into objectively assessable Acceptability Criteria in form of a “Target of Acceptance” (ToA, see section 4). The role also finally judges the success or failure of the V&V effort. The V&V Leader reports to the Acceptance Leader, who is mainly active during the early and the late phases of the VV&A process, leading phase 1, “Develop ToA”, reviewing during phase 3, “Develop ToVV”, and leading during phases 6, “Assess evidence integration” and 7, “Evaluate V&V Report”. The actor who plays the Acceptance Leader should be trusted by the customer.

- The *V&V Leader* knows approaches to V&V, techniques, and tools. This role is responsible for developing an appropriate V&V approach to substantiate the Acceptability Criteria with the information about SoI and SEM available. This role identifies the required Items of Evidence to demonstrate or disprove that an M&S product is correct and valid with respect to the Acceptability Criteria. The V&V Leader supervises the V&V activities performed by the V&V Executioners, and reports to the Acceptance Leader. He is active during nearly all phases of the V&V process, supporting during phase 1, “Develop ToA”, leading during phases 2, “Acquire information” and 3, “Develop ToVV”, supervising during phase 4, “Conduct V&V”, assessing during phase 5, “Assess Items of Evidence”, and

reporting during phase 6, “Assess evidence integration”. The decision, which party the V&V Leader comes from, must be made carefully and deliberately. If it is played by an actor from the supplier side, communication and information exchange between the M&S developers and the V&V Leader would be simplified, but then it must be assumed that the V&V results can be influenced by interests of the supplier party. The Contextual User can assume that the V&V approach will be more critical, if planned by somebody from its own (customer) party, but legacy solutions also might bias this activity. The highest degree of objectivity is achieved, if the actor comes from an independent 3rd Party VV&A Agent, but on the expense of an increased communication overhead.

- The *V&V Executioners* is a composite of roles; it consists of a number of actors playing several roles that actually implement the analysis and test activities required to provide the Items of Evidence specified by the V&V Leader.

    - *Operators*: Those who control the M&S product during its evaluation and are responsible for executing the simulation experiments. An operator should be familiar with both the application domain of the model and the model itself. Note, the one who, e.g., controls the opponent’s forces in a combat simulation is considered to be an operator, while those under test or training are not. During V&V, operators are active in phase 4, “Conduct V&V”.

    - The *System Analyst* and the *Subject Matter Expert* (SME) provide knowledge about the System of Interest, in addition to the knowledge that can be taken from books and other sources. System analysts and subject matter experts involved in V&V mainly become active in the phase 2, “Acquire information”, and 4, “Conduct V&V”, where they (as reviewers) are an invaluable source of knowledge about the System of Interest used for comparison purposes. Although the required skills and imposed responsibilities for both model development and V&V are similar, the roles should be played by different actors to ensure a sufficient degree of

independence between the knowledge used for modelling and the knowledge used for evaluation of the model.

- The *M&S Expert* is knowledgeable about different approaches to modelling a System of Interest and the possible solution methods for the model, including simulation paradigms, methods, and tools. During V&V this role acquires information during phase 2, “Acquire Information”, supports in phase 3, “Develop ToVV”, and acts in 4, “Conduct V&V”, by identifying critical aspects of a model designed and implemented according to a particular paradigm or method, and implemented using a particular tool.

- *Software Engineers* and *Programming Experts* are professionals skilled in designing and implementing software to compute the behaviour of the model. During V&V their support is needed in phase 4, “Conduct V&V”, to trace any detected problems to their origin in the source code (if they originate from “programming bugs”), and to set up tests.

### 3.2.2 Affected Roles

Affected roles take advantage of the REVVA methodology, but are not directly involved into the technical REVVA Generic Process. Often they are decision makers outside of the process, are responsible for the smooth organisational flow of the VV&A effort, and control the flow of information among all parties involved. To precisely define skills and responsibilities of these roles is beyond the scope of this document.

- The *M&S Promoter* sees an advantage of having (usually inferior) people within his organisation use a simulation model or simulation results, and desires to benefit indirectly from the consequences of using M&S products. The responsibility of this role is to overcome administrative obstacles and to identify a potential M&S Sponsor. The M&S Promoter is likely to be involved in the acceptance decision, and may be in a hierarchical organisation such as

encountered in military forces the one who ultimately accepts the risk of using the M&S product. Note that M&S Promoter and Contextual User can be the same actors, but this is not necessarily the case.

- The *M&S Sponsor* creates the financial foundation for the development and VV&A of the M&S product. The actor of this role is member of the customer party. It is likely that the same individual or group of individuals also plays the technical role of the Contextual User.

- The *M&S Project Manager* organises and controls a particular use or series of uses of an executable model (i.e., model selection, experiment design, and experiment evaluation). If there is no appropriate model available, the responsibilities also include the initiation of M&S product development. Besides the knowledge and experience required for the organisational aspects, knowledge in M&S is advisable. To support concurrent V&V, the simulation project manager assures that communication between the involved parties takes place in a cultivated atmosphere, and provides the access to the required information about the model. He is advised to carefully review the V&V results, as they are likely to support the identification of project risks.

- A *VV&A Project Manager* is required, if V&V is not integral part of model development, and organises the managerial aspects of the V&V endeavour, when it is managerially separated from model development. This role is responsible for choosing an appropriate VV&A process model (such as proposed in section 4), to create a detailed schedule for the VV&A effort, and to choose individuals to act in the roles. It is likely that the same individual or group of individuals also plays the technical role of the Acceptance Leader and/or the V&V Leader (see below).

## 3.3 Choosing Actors

An appropriate cast must be found for the roles. Whether an actor or group of actors is appropriate depends on organisational aspects, including the desired degree of

independence and required transfer of information, and on its educational background and experience.

### 3.3.1 Parties from which the actors may come from

The assumption here is that if technical activities should be shared, responsibilities on both the customer side and the supplier side have to be clearly identified to prevent conflicts of interests. Roles may be played by:

- actors, who are working under the responsibility of the customer independently from the supplier to assess the M&S products, including the V&V products generated by M&S developers,
- actors, who are working under the responsibility of the supplier (i.e., are dependent on the supplier) and are implementing V&V activities, applying "sound software engineering principles" as in intrinsic part of M&S project and contracts
- actors, who are members of a 3rd party (such as the 3rd Party VV&A Agent or the Acceptance Authority), therefore independent from both the customer and the supplier, and assumed to be only minimally biased.

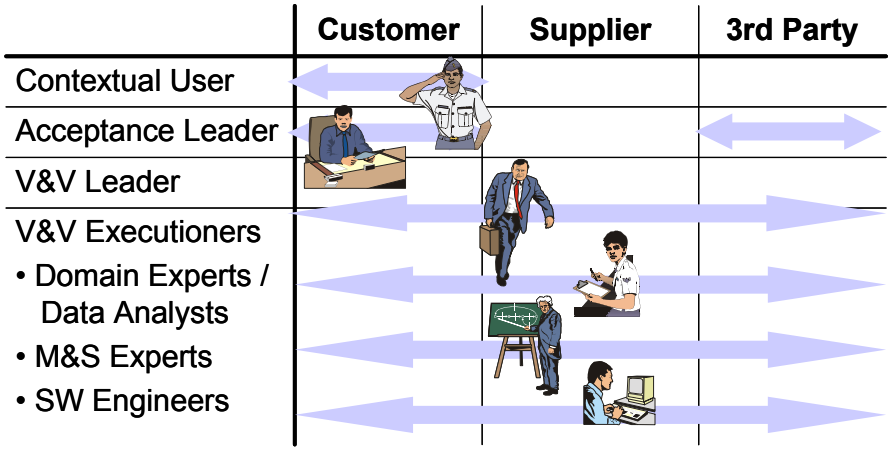



**Figure 3: Parties of the actors**

The organisation of the overall VV&A activity is a decision of the customer, mainly based on the evaluation of the risks and impacts of the intended purpose of M&S use. As a constant, it is assumed that the Contextual User always is played by a member of the customer party. The Acceptance Leader is always situated on the Contextual User's side, i.e., he either is a member of the customer party or of a 3rd party trusted by the customer. The parties from which the V&V Leader and the V&V Executioners come from are chosen depending on the V&V project's needs and constraints. As a rule of thumb, having actors from the supplier party play them creates the least administrative overhead and minimises direct cost (quick communication channels, reduced problems concerning the protection of intellectual property), while actors chosen from the 3rd party will be most expensive (learning period), but least biased and most likely to enforce the delivery of quality reports.

This choice of actors should not be monolithic, because some parts or components of M&S applications might be critical, when others have a little impact on the operational use and simulation results. The possibility of organising and sharing the V&V among the different parties should be considered for each component of the relevant whole-part decomposition, depending on its levels of impact.

This perspective of sharing technical V&V activities, but distinguishing responsibilities leads to the identification of three major levels of V&V organisation, also shown in Table 2:

- Dependent V&V (DV&V): The V&V is conducted by the M&S supplier according to the customer's V&V requirements (i.e., the actors for V&V Leader and V&V Executioners are members of the supplier party), and accepted "as is" by the customer.

- Independent Assessment (IA): The V&V work is conducted by the M&S supplier, but is assessed by an independent Acceptance Leader (3rd Party) trusted by the customer,

- Independent V&V (IV&V): V&V activities are planned and conducted independently from both the supplier and the customer by the 3rd Party VV&A Agent.

Table 2 gives an overview over cost-effective assignment of actors to roles, considering independence from the customer's perspective.

**Table 2: Actors, Roles, and Independence**

| | Acceptance Leader | V&V Leader | V&V Executioners |
|---|---|---|---|
| **DV&V** | Not explicitly assigned | Supplier | Supplier |
| **IA** | Customer or 3rd Party VV&A Agent | Supplier | Supplier |
| **IV&V** | Customer or 3rd Party VV&A Agent | 3rd Party VV&A Agent | 3rd Party VV&A Agent |

### 3.3.2 The Actors Backgrounds

An actor or group of actors needs to have the required education, skill, and experience to satisfactory play the assigned role. To be most effective and efficient, they must be able to exist and move within the "four worlds", as sketched in Figure 4. The actor of the Contextual User is located in the "Problem World", and is able to phrase the problem and to develop an idea of the problem solution. The Acceptance Leader creates the bridge between the "Problem World" and the "M&S World", by knowing the terminology of both worlds, being able to understand the Contextual User and to aid him when he needs to express his needs in terms people in the "M&S World" can understand unmistakably. The System Analyst or SME lives in the M&S World close to the border to the "Problem World", and therefore is able to abstract and idealise the SoI in a problem-oriented manner. At home in the "M&S World" is also the M&S Expert, who can create a computable description of the System Analyst's perception of the SoI. This then serves for the HW/SW Expert, who lives in the "Product World", as specification to build the

machine, which can perform the required computations. The V&V Leader needs the capability to move through the “M&S World” and the “Product World”, and should have a basic understanding of the “Problem World”.

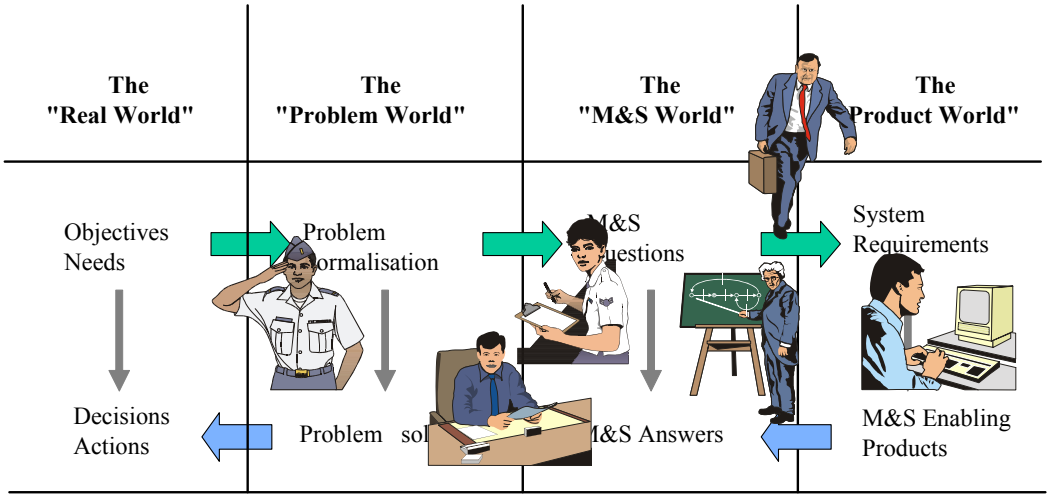



**Figure 4: Choosing Actors**

# 4 PROCEDURE “WHAT TO”

This section contains the description of the REVVA Generic Process, some of its background, and the documentation of its phases and products. Please recall that in “REVVA” the “A“ stands for “acceptance”, not accreditation – if you are heading for accreditation, you might look for appropriate acceptance related products of the REVVA Generic Process, which you then can feed into your official accreditation procedure.

## 4.1 VV&A Planning

The responsibility for planning the organisational aspects of the V&V project is assigned to the V&V Project Manager, or – if V&V is integral part of model development – to the M&S Project Manager. Over the whole VV&A endeavour, the VV&A plan will be available in different levels of detail. In the very beginning it documents the decision to follow the REVVA Generic Process, identifies the individuals that become active as actors in the different phases, schedules the phases over time, and allocates resources. The V&V plan will be refined as V&V progresses. As soon as the ToA is defined (phase 1 of the REVVA Generic Process, see section 4.2), priorities expressed by allocation of time and money can be assigned to the individual Acceptability Criteria. With the ToVV available (phase 3), the scheduling of activities can be refined in more detail, and the acquisition or development of missing tools can be planned. Depending on the experience with the simulated SoI and M&S V&V in general, and on the knowledge about SEM and SoI, the V&V plan will need more or less adaptation during its implementation. Thus, there is no explicit VV&A planning phase included within the REVVA Generic Process, and it is assumed that the activity to adjust the managerial aspects of the V&V plan whenever necessary is directly coupled to the technical refinement steered by the process.

## 4.2 Flow of Phases and Products

Although V&V activities should – whenever possible – be performed concurrently with model development to facilitate the early detection and correction of incorrectness or invalidity, the order of major steps taken in an M&S product-oriented VV&A effort and the flow of VV&A products can be described as a separate stand-alone generic VV&A

process. The progress along the VV&A process and its iterations is controlled by the availability of information about both the SoI and the SEM (including intermediate developmental M&S products and the final M&S products) that shall be evaluated during the V&V effort.

The REVVA Generic Process as depicted in Figure 5 supports product-oriented VV&A during or after model development (e.g., as required for reuse for another related intended purpose), and can be used as guidance for planning a VV&A effort. The "V-Form" for the process representation was deliberately chosen, mirroring the preparation for V&V and the execution of the V&V activities on the left trunk ("\") of the "V", against the evaluation and the integration of the V&V results for the purpose of assessment on its right trunk ("/").

All activities that are performed during V&V are organised in phases. Due to, e.g., progress in SoI development or SEM development, new knowledge gained during V&V, the unavailability of expected and scheduled information, or unfeasibility of a particular test, iterations among the phases are necessary. As a consequence, products that originate from the revisited phases need to be strictly version controlled. The following process is a high-level description of VV&A, which is rather intended to be refined to meet a particular project need, than tailored.

## 4.3 Description of Phases and Products

Each phase description contains a summary of activities, lists the input and output products, and points out the involved roles and their type of involvement. The REVVA Generic Process is no waterfall process, but iterative, which means that especially those products close to the bottom of the "V" become available in several versions.

### 4.3.1 Develop ToA

Based on the intended purpose of model use, a detailed set of Acceptability Criteria is developed in such a manner that passing the Acceptability Criteria implies fitness for purpose. It is helpful to distinguish between different types of requirements (which include, e.g., availability of particular intermediate products, reuse of pre-selected federates, operation on a particular computer platform and operating system, but also

statements concerning SEM behaviour), focusing on those, which exclusively address the model's correctness and validity. All Acceptability Criteria and the rationale for their derivation are recorded as the "Target of Acceptance" (ToA). For the development of the ToA, the chosen model development process and the system development process should be taken into account, as often Acceptability Criteria do not only address properties of the final product itself, but also those of intermediate products that will be created as defined in the development processes. This phase is executed by the contextual user under the guidance of the Acceptance Leader. For the formulation of expressive, relevant, and precise Acceptability Criteria intensive communication between the Contextual User and the Acceptance Leader is crucial.

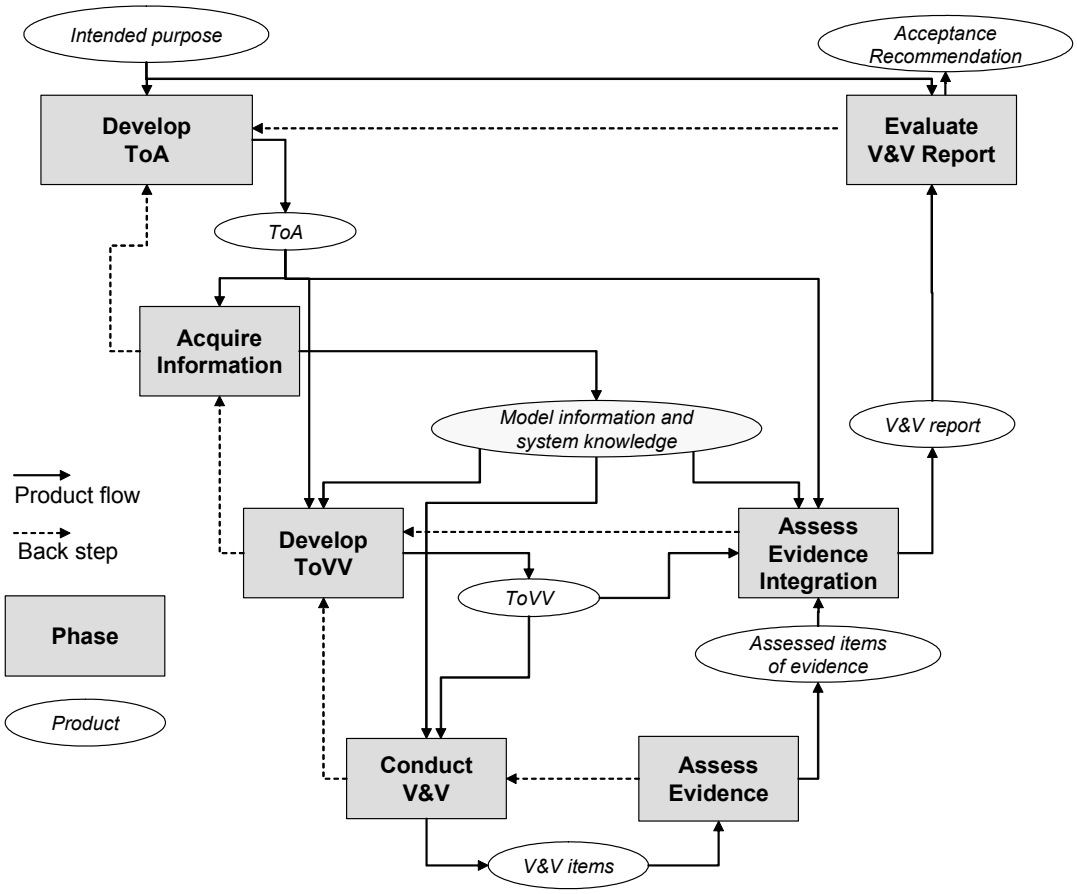



**Figure 5: The REVVA Generic Process**

Acceptability Criteria should be prioritised. For simulation-based endeavours with a low impact on real world decisions or actions, some superficial indicators that the Acceptability Criteria are passed may be sufficient, while safety critical aspects might

require an unmistakable proof. The required effort may vary significantly among the individual Acceptability Criteria. With the “four worlds” concept introduced in Figure 2 in mind, the derivation of Acceptability Criteria is performed from the problem perspective for the “M&S World”.

*Target of Acceptance (product):* The Target of Acceptance (ToA) is the result of Phase 1, “Develop ToA”. It contains a precise specification of the Acceptability Criteria that the model needs to meet to be considered acceptable for a particular, well-defined intended use, answering the question “What exactly needs to be assessed?”, and the rationale for their derivation from the intended purpose. On top of a refinement hierarchy stands the vague intended purpose, which is refined into a set of sub-purposes, which again can be decomposed, until Acceptability Criteria related to the M&S product’s correctness and validity can be derived directly from the lowest sub-purposes (“lower criteria” here means criteria that are lower ranked in the hierarchy). Good Acceptability Criteria have the technical precision to be determinable (or at least repeatedly assessable). How to document the ToA is described in section 5.1. All Acceptability Criteria that originate form the application domain of the M&S product should be understandable for the targeted user. All Acceptability Criteria that address M&S related issues should be comprehensive for M&S experts. The set of Acceptability Criteria does not imply any methods or techniques how to assess them. They and their rationale documented in the ToA serve as input for the V&V phases 2 “Acquire Information”, 3 “Develop ToVV”, and 6, “Assess Evidence Integration”. The ToA identifies the relevant aspects from Figure 2’s “Problem World” which need to be transferred into the “M&S World”. If it is decided to have a 3rd Party VV&A Agent to do the V&V, the ToA defined the technical objective of the contract with the 3rd Party VV&A Agent.

### 4.3.2 Acquire Information

To prepare the planning of a V&V effort, one needs to get an overview over the data, information, and the knowledge about the System of Interest and the Simulation Executable Model that is available or is likely to become available. Under consideration of the intended purpose of model use and the detailed Acceptability Criteria (documented in the ToA), knowledge about the System of Interest, its structure and

behaviour, its subsystems and their structure and behaviour, or related systems is collected and filed. (In related work this body of real world knowledge is referred to as “referent”.) Sources of information about the System of Interest include expert opinion and measured data from test labs, test ranges, or real operation of the System of Interest itself, a prototype, or a related system. The set of sources from which this information is acquired, which is used for validation, should not be identical with the set of information sources used for conceptual modelling, but additional information sources should be exploited to avoid the duplication of potential modelling errors. All acquired information about the System of Interest is placed in a common repository.

Also the degree of insight into the model heavily impacts the design of a V&V approach. All relevant existing (post-development V&V) or expectable (concurrent V&V) information about the model should be identified and made available for the following V&V activities. Sources for this information are the simulation conceptual model, model documentation, interviews with the model developers, design documents, the formal model, program code, and computed model behaviour. All acquired information about the model is placed in a common repository.

*Model information and system knowledge (product):* This information will be used as foundation of the approach to demonstrate the model’s correctness and validity. The product identifies all sources of information and knowledge and all bodies of information and knowledge that are available or will become available during the V&V effort. The model information includes all documentation of the model, its design documents, its code, the ideas behind it (modeller’s perception of the real world), and behaviour data of the model and its sub-models. The conceptual model plays an extremely important role for gaining knowledge about the model. Information about the System of Interest includes results of system analysis, data measured at the System of Interest and behaviour observations concerning the system itself and its subsystems, its design documents, knowledge gained from the physical decomposition of the system and the examination of its components, expert opinion, and examination results from a related or similar system. (This information is often referred to as so called “referent”.)

The acquired information and knowledge about both the M&S product and the System of Interest is ideally stored in (an) appropriate remotely and securely accessible data base(s), because it needs to be accessible for several roles during the whole V&V endeavour.

### 4.3.3 Develop ToVV

With the ToA and an overview over the available data, information, and knowledge about the System of Interest and the Simulation Executable Model, the Target of V&V (ToVV), which documents the approach taken to the substantiation of the Acceptability Criteria, is developed. For each Acceptability Criterion a rationale is developed, which points out how with the information at hand and the available technical means it can be demonstrated that the Acceptability Criterion is passed or failed. To substantiate that the Acceptability Criterion is met becomes a *V&V Objective*. Developing the ToVV usually includes the decomposition of a V&V Objective into more easily assessable V&V sub-objectives. The V&V Leader develops the ToVV, consulted by those who will actually implement it later (V&V Executioners). When developing the ToVV, not only the available information about model and System of Interest, but also all constraints given by budget, deadlines, tool availability, and human resources need to be taken into account. There should be a clear separation between the contents of the ToVV, which will be assessed for its convincing force and approved by the Acceptance Leader, and the managerial aspects required for a smooth flow (documented in detail in the V&V plan). The ToVV directs the execution of V&V and will be used for the integration of the evidence. The development of the ToVV takes place in the “M&S World” of Figure 2, focusing on M&S related issues, under consideration of the constraints imposed by the “Problem World”, resulting in required activities in the “Product World”.

*Target of Verification and Validation (product):* The Target of Verification and Validation (ToVV) originates from the “Develop ToVV” phase. With the “What to demonstrate” given by the ToA, the ToVV elaborates on the “how to demonstrate it”. It identifies the Items of Evidence required to substantiate the Acceptability Criteria contained in the ToA, and documents the rationale for the necessity and sufficiency of these Items of Evidence. Those Acceptability Criteria that cannot directly be assessed are further

decomposed and documented according to the documentation rules for the ToVV (see section 5.3). The rationale for this decomposition includes the information about the model and the knowledge about the System of Interest available and the justification, why passing the lower V&V sub-objectives also implies passing the Acceptability Criterion from which they were derived. Besides the required information within the Items of Evidence, the ToVV also identifies their individual desired probative forces needed to consider them “strong enough”. In some cases the identification of the techniques and tools that can be most efficiently used for the demonstration may be included. Under consideration of the constraints imposed from the “Problem World” (Figure 2), the ToVV concentrates in issues within the “M&S World”.

### 4.3.4 Conduct V&V

The V&V Executioners plan and execute the V&V to provide the V&V items required by the ToVV. They acquire evidence or create it by implementing the appropriate analyses or tests. The team reports back to the V&V Leader, if, due to, e.g., missing or insufficient information about the model, missing knowledge about the System of Interest, or unavailability of the required tools, a particular required V&V Item cannot be acquired, or if an elementary V&V objective is demonstrated to be failed. Each test result, analysis report, or proof outcome is documented as V&V Item.

For the execution of V&V, close cooperation with those responsible for quality assurance and software testing is advisable. Rather than duplicating the tests and evaluations conducted by those individuals, their test suites may be extended by test cases or complete test scenarios, which aim rather on the demonstration of the correctness and validity of the underlying model, than on the functionality of the enabling hard- and software. The V&V Executioners must be aware of many potential errors, including non-reproducibility of experimental conditions, sampling errors, and bad allocation of effort. The conduction of V&V activities takes place in both the “M&S World” (e.g., development of test cases and oracles) and the “Product World” (e.g., tracing from the formal model to the executable model) of Figure 2.

*V&V items (product):* As result of the "Conduct V&V" Phase, the V&V items constitute the "atomic building blocks" of V&V. A V&V Item consists of some piece of information about the Simulation Executable Model, the evaluation objective, reference information, an evaluation technique, and the evaluation result. For validation, the reference information consists of knowledge about the System of Interest. For verification, the reference information consists of, e.g., representation rules, model information in a different representation form, or formalism.

To be acceptable as an Item of Evidence (see following phase), the summary of information given in the V&V Item must be complete and comprehensible, and needs to facilitate the repetition, duplication, or reproduction of the activities that resulted in the documented outcome. V&V Items have different probative forces, depending on the method or technique used for their creation, and the reference information or knowledge used. Note that V&V Items are not necessarily created by the V&V conductors, but that they also can be leveraged from, e.g., testing, previous model applications, or model integration testing. The V&V Items build together with the ToA and the ToVV the foundation for the assembly of the V&V report.

The V&V Items are likely to be created by numerous individuals from different organisations, and need to be accessed by authorised individuals for review or assessment. It is advisable to store the V&V Items in an appropriate remotely and securely accessible data base.

### 4.3.5 Assess Evidence

The key issue of this phase is to accept the individual V&V items as Items of Evidence. or to reject them for strengthening or improvement. The V&V Leader reviews each V&V Item gathered, created, or otherwise provided, consults his V&V Executioners, and assesses the probative force of each V&V Item, which becomes an Item of Evidence, if accepted. If the V&V Leader considers the probative force of a V&V Item as unacceptably low, the V&V Item needs to be strengthened or discarded. Otherwise, the (accepted) Item of Evidence is added to the evidence pool, which its perceived probative force annotated. The Items of Evidence serve as the atomic building blocks on which the acceptance decision later is based. Note that all information required for

the assessment of the V&V Item as Item of Evidence must come with the Item of Evidence itself. Looking up reference information during this phase should rather be the exception, than the rule.

*Items of Evidence (product):* The Items of Evidence document the individual executions of single V&V techniques and their outcomes, as conducted or acquired by the V&V Executioners, with annotations of their individual probative forces. The assessed Item of Evidence includes (in addition to the information contained in the V&V item from which the Item of Evidence originates) the V&V Leader's assessment of the V&V result, the acceptance or rejection of each Item of Evidence, and a judgement of its probative force.

### 4.3.6 Assess Evidence Integration

The key issue of this phase is to build and accept or reject the rationale of supporting the Acceptability Criteria with the available Items of Evidence. The V&V Leader assembles and integrates the approved Items of Evidence according to the most recent version of the ToVV. Under reconsideration of the ToA, the Acceptance Leader reviews the assembly of the evidence and judges how well the evidence substantiates that the Acceptability Criteria are passed (convincing force). From the Items of Evidence he learns, which aspects of the model have been examined in detail, while their associated probative force gives an estimate of their reliability as estimated by the V&V Leader. The current version of the ToVV contains the way how the Items of Evidence are supposed to substantiate the claims that the Acceptability Criteria are met. If the available evidence leaves gaps or loopholes for the substantiation of the Acceptability Criteria, the ToVV needs to be adjusted and the additional V&V activities conducted to provide the missing Items of Evidence. When the V&V activities are considered to be finished and no further iterations through the cycle "Develop ToVV- Conduct V&V – Assess Evidence" are made, the V&V Leader compiles the V&V report from the ToA, the final, accepted version of the ToVV, and the evidence available. Note that here it is assumed that the Acceptance Leader trusts the V&V Leader concerning the assessment of the probative force of the individual Items of Evidence. If there is no relationship of trust between the Acceptance Leader and the V&V Leader, then the

Acceptance Leader might need to re-assess the Items of Evidence and their probative forces, too.

*V&V report (product):* The gathered or otherwise created Items of Evidence assembled and integrated by the V&V Leader to substantiate the Acceptability Criteria in the ToA according to the most recent version of the ToVV, build the substance of the V&V report. The V&V report links the rationale why the referenced Items of Evidence substantiate the claim that the Acceptability Criteria are passed with the Items of Evidence. If Items of Evidence for the substantiation of particular Acceptability Criterion are missing or are too weak, or if disproving evidence was created, this is also recorded in the V&V report. The completed V&V report is passed on to the Acceptance Leader for assessment.

### 4.3.7 Evaluate V&V Report

Finally it is up to the Acceptance Leader to support the Contextual User in the evaluation of the V&V report. Based on the probative force of the evidence, the convincing force of the ToVV, and the selection of Acceptability Criteria as motivated in the ToA (all documented in the V&V report), the Acceptance Leader estimates the residual uncertainty associated with the statement that the M&S product actually is fit for its intended purpose. When the ToA was created it was implicitly assumed that meeting the Acceptability Criteria indicates fitness for purpose of the M&S product. This needs to be confirmed under consideration of the assessed ToVV, the assessed Items of Evidence available, and their probative force. If the residual uncertainty is considered to be too high, either the intended use must be modified, or the V&V effort partially repeated with an extended ToA. The formally authorized decision maker is encouraged to base the acceptance decision on the acceptance recommendation.

*Acceptance Recommendation (product):* The final recommendation whether to accept or reject the M&S product for its intended use, considering the uncertainty that is left even after V&V was successfully conducted, is documented in form of the acceptance recommendation. The acceptance recommendation confirms that the acceptability for the intended purpose is demonstrated by the Items of Evidence gathered to

substantiate the Acceptability Criteria. If any required evidence is missing or too weak, or if a previously undetected weakness in the Acceptability Criteria was detected, the acceptance recommendation states this and indicates the consequences for the M&S product's use. The acceptance recommendation and the V&V report lay the foundation or the final acceptance decision by the Contextual User or the formally authorized instance.

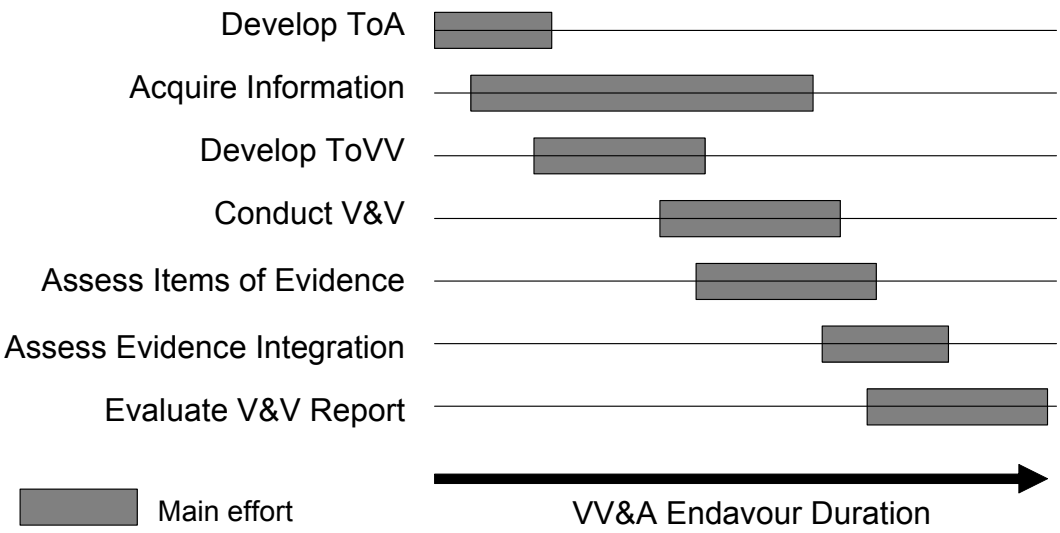



**Figure 6: Sketch of the temporal overlap among the phases of the REVVA Generic Process**

## 4.4 Temporal and Causal Dependencies

The REVVA Generic Process is rather an iterative process than a waterfall process. The expectable overlap among the phases during the whole duration of the VV&A endeavour is sketched in Figure 6. It shows clearly that the development of the ToA, which holds the Acceptability Criteria and the rationale of their derivation, should be concluded early in the overall VV&A endeavour. The acquisition of information can start as soon as the first parts of the ToA are available, but activities of this phase are likely to continue for quite a while, if either the model or the SoI are under development. (For concurrent V&V, due to the ongoing increase of knowledge about the model and the SoI, changes in the ToVV and the need for stronger evidence are likely to occur.) With the ToA nearly completely stabilised, based on the information available so far (or

scheduled to become available) the ToVV can be developed. Vice versa, information acquisition will be influenced by the information needs expressed in the ToVV. V&V activities can be conducted, when the ToVV is sufficiently developed, but infeasibility of some planned V&V activities might require changes in the ToVV. Small iteration circles between the assessment of the V&V items and their production are essential for high efficiency. When the integration of evidence is assessed, the need for slight changes of the ToVV might occur. The evaluation of the V&V report can start with the first chapters becoming available.

## 4.5 Roles Allocation in the Process

The responsibilities of the individual roles within the phases of the REVVA Generic Process have been described in section 3.2. Figure 7 visualises who becomes active in which phase.

- In the phase “Develop ToA” close cooperation between the Contextual User and the Acceptance Leader is crucial. The Acceptance Leader helps the Contextual User to precisely identify and document the Acceptability Criteria.

- The activities in the phase “Acquire Information” is lead by the V&V Leader, who identifies types of data, information, and knowledge promising to be useful to substantiate that the Acceptability Criteria defined in the ToA are passed (or failed). To collect this information, support of the Contextual User and the V&V Executioners is required. Not depicted in Figure 7 is the necessity to maintain a cultivated atmosphere concerning interaction with the model developers, and, if applicable, the SoI developers.

- To “Develop ToVV” is the task of the V&V Leader, who is advised to involve and consult the V&V Executioners concerning the practicability of the chosen approach. It is also reasonable (and might be contractually required) to get an early approval of the first version of the ToVV from the Acceptance Leader.

- The phase “Conduct V&V” rests in the hands of the V&V Executioners.

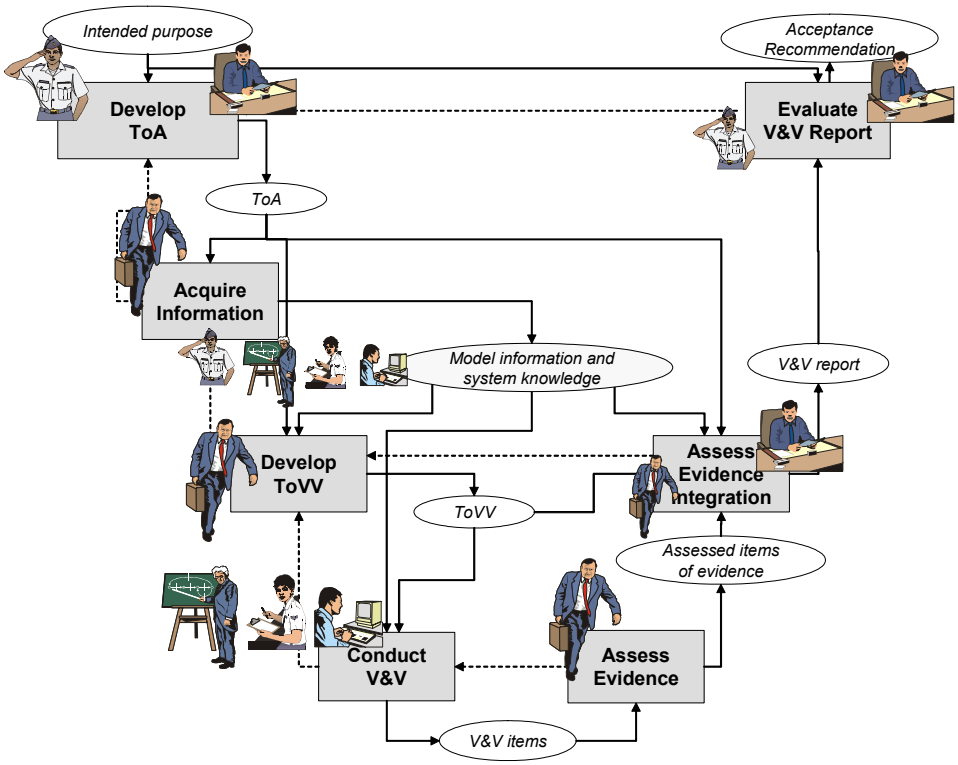



**Figure 7: Roles assignment by phase**

- The V&V Leader assesses the V&V items during the “Assess Evidence” phase as “Items of Evidence” (IoE). Note that on the Customer’s discretion also approval by the Acceptance Leader might be required (not depicted).

- During “Assess Evidence Integration” the V&V Leader performs the integration the Items of Evidence according to the most recent version of the ToVV. Under re-consideration of the ToA, the Acceptance Leader determines whether the integrated evidence substantiates that the Acceptability Criteria are passed or failed, and documents his confidence for each individual Acceptability Criterion separately.

- Reassuring with the Contextual User under reconsideration of the intended purpose, the Acceptance Leader “Evaluates the V&V Report” and gives an

acceptance recommendation to those authorized to formally accept the M&S product.

An Overview over the responsibilities of the VV&A main roles in the REVVA Generic Process is given in Table 3. The REVVA Generic Process is introduced in section 4.

**Table 3: Roles and Responsibilities in the REVVA Generic Process**

| | **Acceptance Leader** | **V&V Leader** | **V&V Executioners** |
|---|---|---|---|
| **Develop ToA** | Perform | Observe | - |
| **Acquire Information** | - | Lead | Perform |
| **Develop ToVV** | Observe* | Perform | Support |
| **Conduct V&V** | - | Lead | Perform |
| **Assess Evidence** | Observe* | Perform | Report |
| **Assess Evidence Integration** | Perform | Consult | - |
| **Evaluate V&V Report** | Perform | - | - |

* “Double-check” on customer’s discretion

# 5 PROCEDURE "HOW TO"

This section gives more guidance on how to proceed within the individual phases outlined in section 4.

## 5.1 Develop ToA

During the first phase of the REVVA Generic Process, the vague intended purpose is broken down into clearly defined Acceptability Criteria. These are recorded in the form of the ToA. It broadly has three steps:

- Identify Acceptability Criteria,
- Make Acceptability Criteria "measurable",
- Assign levels of impact.

A summation of this section is given in the following box as a quick guide for practitioners; references to paragraphs where the items are elaborated upon are given.

Identify Acceptability Criteria. Use hints in paragraph 5.1.1.

- determine intended purpose for the M&S product,
- derive sub-objectives from the intended purpose where necessary,
- derive Acceptability Criteria from sub-objectives,
- derive sub-criteria from Acceptability Criteria where necessary and give decomposition arguments. Limit the derivation to the use of the M&S product.

Make Acceptability Criteria "measurable". Use hints in paragraph 5.1.2.

- determine how to measure the criteria for each node in the ToA. This is expressed as a Measure of Effectiveness (MoE),
- determine constraints on the MoE that indicate when the effectiveness is sufficient.

Assign levels of impact. Use hints in paragraph 5.1.3.

- estimate worst case impact of M&S product use,
- derive from this worst case impact the level of impact for all nodes in the ToA.

### 5.1.1 Identify Acceptability Criteria

A good starting point for derivation of the Acceptability Criteria is given by the requirements used in the development process of the M&S asset. If these requirements are not available not sufficiently precise or if the M&S asset needs to be validated for a different purpose than it was developed for, the criteria need to be refined or derived anew. The determination of the Acceptability Criteria is a top-down activity, which starts with questioning the Contextual User about details of the intended purpose, and continues with the consequent development of objectives and sub-objectives, as appropriate. Finally, when the sub-objectives hierarchy is developed, from the lowest sub-objectives directly the Acceptability Criteria are derived, which constitute the leaves of the ToA and address the desired correctness and validity properties of the M&S product. The upper half of Figure 8 illustrates the concept of objectives decomposition. The lower half of the figure depicts the successive development of a V&V, which will be addressed during phase 3 “Develop ToVV”, see section 5.3.

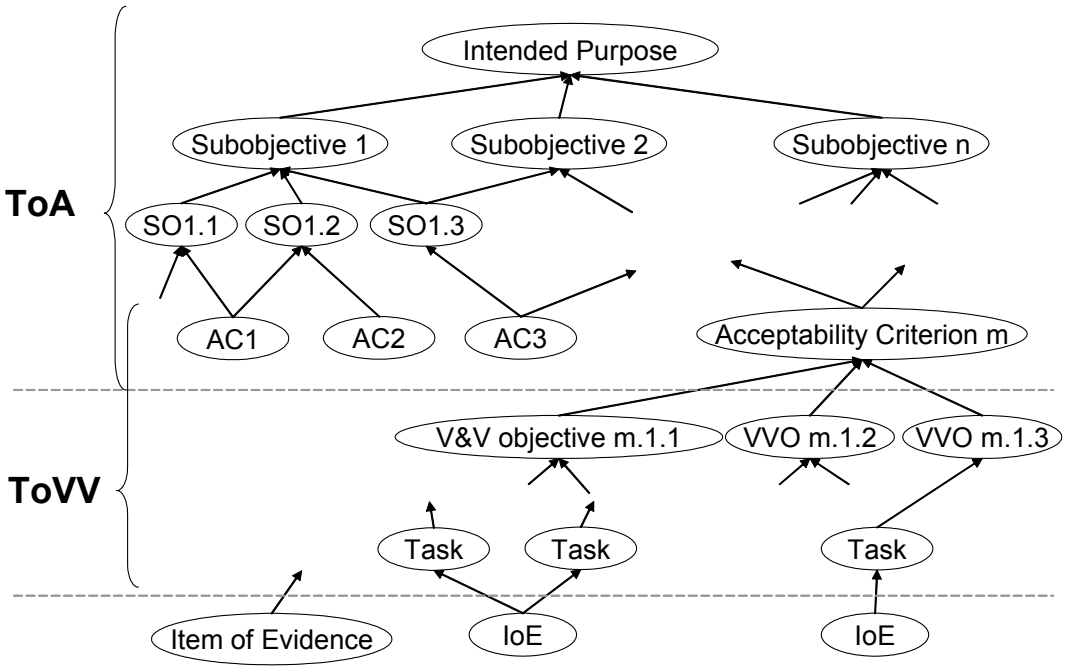



**Figure 8: ToA and ToVV**

The Target of Acceptance (ToA) documents the hierarchical decomposition of the vaguely stated intended purpose into testable Acceptability Criteria, and the rationale for the chosen decomposition. The risk associated with the intended purpose of model use is analysed under explicit consideration of the ToA, to facilitate the assignment of priorities to the Acceptability Criteria in the ToA. The structure and information contained in the ToA is used for assessing correctness and validity of the M&S asset under consideration.

First, the Acceptability Criteria must be related to validity aspects. Other requirements may be important for the overall success of the M&S asset, but cannot be used in the evaluation of the correctness and validity. Figure 9 gives an overview of the typical requirements used in development and which the ToA is concerned with.

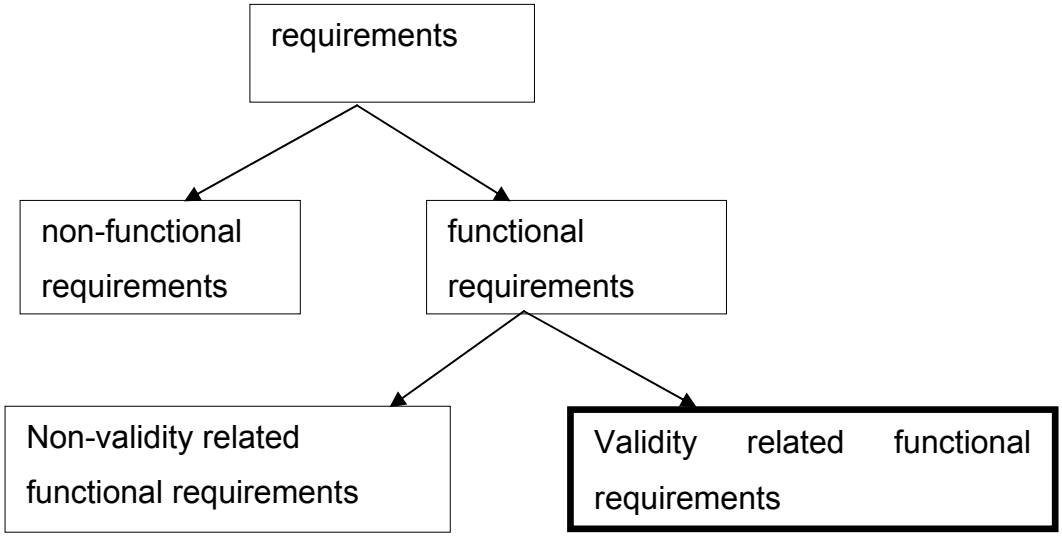



**Figure 9: Taxonomy of requirements (simplified) indicating the type of requirements of interest for this work.**

The basic structure of the ToA stems from the way the objectives are elicited. It starts with the intended purpose from the customer perspective, which means that the focus is on the impact of the M&S asset within the operational context. The objectives are - when necessary - refined into a set of sub-objectives. These sub-objectives can again be split into several still more detailed objectives, which finally are directly translated

into Acceptability Criteria addressing the M&S product. Therefore the basic structure of the ToA is a tree. In practise, however, it is not necessarily a true tree structure, but rather a directed acyclic graph. It can happen that two requirements both need the same or somehow interactive sub-requirements. The refining of the requirements is called the decomposition of the ToA.

The objectives that are refined must have an additional piece of information associated with them. This extra information is the argumentation on why the sub objectives together constitute their parent objective. This argumentation is named the "Decomposition argument". This argumentation is the glue in the hierarchical structure of the ToA. The objectives that are not further refined do not need this argument.

There exists a wide range of requirements engineering techniques [Kotonya and Sommerville 1998]. The most straightforward technique is interviews. This technique often lacks formality and is more appropriate for obtaining the global view and general requirements. However, carefully prepared and structured interview techniques or brainstorm sessions reduce this problem. Literature [Lane and Alluisi 1992] and practice shows that using a closed interview approach (a fixed question/answer set) in combination with some knowledge/expert system on the application and problem domain, it is possible to utilise these techniques for specifying requirements.

High-level criteria for assessment of correctness and validity for an executable model include:

- I/O behaviour goodness of fit: Similarity of behaviour of model and system under comparable experimental conditions within the scope and limitations of the Experimental Frame,

- Sub-models I/O behaviour goodness of fit: Similarity of behaviour of the sub-models and their associated real subsystems under comparable experimental conditions,

- Correctness and validity of the underlying conceptual model: Internal consistency of the conceptual model, and consistency with available knowledge about the System

of Interest. Appropriateness of the chosen abstraction and idealisation within the context of the intended use. Ability of the model's input parameters to reflect the relevant influences on the behaviour of the System of Interest, and the ability of the model's output parameters to approximate the attributes of interest of the System of Interest (goal parameters),

- Correctness and validity of the underlying formal model: Internal consistency of the formal model, consistency with the modelling formalism, consistency with the conceptual model and completeness of coverage of the conceptual model. Appropriateness of the chosen heuristics and algorithms,

- Correctness and validity of embedded data: Consistency between encoded data items and recorded measurements or otherwise generated real world data. Ability of the "hard-wired" data items to quantitatively express the system attributes of relevance.

The following criteria for the assessment of a simulation experimental frame - following the identification proposed in [METHGU2 2004] - should be taken into account:

- Expressiveness of the experiment: Degree to which the (simulation) experiment is expected to provide new relevant information,

- Consideration of the modelling method: Considered compensation of deviations between real experiment and simulation experiment due to the model's stochastic or deterministic properties,

- Representation of external influences: Coverage of relevant environmental conditions, which impact the outcome of the associated real experiment, and validity of input representation that will be provided to the model. Ability of data values to represent the environmental influences quantitatively with sufficient accuracy,

- Representation of (real) experiment goal parameters: Ability to project the required model output parameters to the relevant attributes of the System of Interest.

### 5.1.2 Make Acceptability Criteria measurable

The Acceptability Criteria must be made measurable in the sense that it must be made clear that the results for sub-criteria together determine whether a criterion is passed or not. The derived Acceptability Criteria are in general stated in the form of sentences containing vague terms such as “must match the System of Interest”. This specifies neither what must be measured nor which values are allowed. A Measure of Effectiveness (MoE) is the measure that is used for judging the passing or failing of the criterion. It is a function that can be measured - possibly using the results from sub-criteria - and produces a result. Examples of MoE are:

- the value of a variable,
- the error in an output variable,
- a situation in a scenario that should or should not occur,
- the judgement by an expert on a difficult to quantify property, e.g. the “feel” of a vehicle simulator equipped with a motion base.

The MoEs must have a number of properties, see [Sproles 2000]:

- “MoEs represent the viewpoint of the customer,
- MoEs assist in making the right choice by indicating ‘how well’ a solution meets the stakeholders need,
- MoEs should be able to be quantified in some manner,
- MoEs specify neither performance nor constraints, i.e. it only specifies what shall be used for measurements,
- MoEs are independent of the chosen method of solving the problem.”

Since the MoE are typically functions that result in a value, the limits of this value between which Acceptability Criterion is considered to pass need to be specified. These are the constraints. Typical examples of constraints are:

- The results of the measure applied to the M&S product and the System of Interest (whose result are documented in the referent) must differ less than a specified percentage,
- The result of the measure should be less than or more than a specified value,
- The result of the measure should be “true”.

### 5.1.3 Impact

The risk associated with the use of the M&S product is considered to be the driver for V&V. Impact domains have been identified by [Muessing, Laack and Wrobleski 1997] and [Mugridge 1999], which may serve as criteria for the estimation of the worst case impact. The potential consequences of the use of erroneous M&S product are, for reasons of pragmatics, qualified in four distinct classes for all identified impact domains, ranging from “negligible”, to “marginal”, “critical”, and “catastrophic”. Examples (which also serve as guidance) for impact classification are given in Table 4.

For each Acceptability Criterion in the ToA, the impact of using the M&S product for its intended purpose – despite of failing this Acceptability Criterion – deliberately or unknowingly would have needs to be determined. An impact estimate is associated with each leaf of the ToA. The more critical the impact of the individual Acceptability Criterion is, the lower should be the residual uncertainty associated with its assessment, i.e. the more convincing should be the branch of the ToVV associated with its substantiation, and the more probative should be the referenced evidence (see phase 3). Managerial aspects including cost, time, and availability of (human) resources need to be resolved for each leaf node individually such that the total monetary and time budget and other restrictions are respected. In this text the focus is restricted to the VV&A related steps.

For most of these issues fitting hooks can be found from this work to steps in the practitioner's favourite management process.

## Table 4: Impact domains and severity categories from [Mugridge 1997]

| Severity Category | | CATASTROPHIC | CRITICAL | MARGINAL | NEGLIGIBLE |
|---|---|---|---|---|---|
| Ref | Impact Domain | | | | |
| 1 | **Personal Safety** | Death | Severe injury | Minor injury | Less than minor injury |
| 2 | **Occupational Illness** | Severe and broad scale | Severe or broad scale | Minor and small scale | Minor or small scale |
| 3 | **System Damage** | Loss of system | Major system damage | Minor system damage | Less than minor system damage |
| 4 | **Environmental Impact** | Severe environmental damage (eg. Chernobyl) | Major environmental damage (eg. Most land blight) | Minor environmental damage (eg. pollution of a stream) | Trivial environmental damage (eg. minor spillage with no long term effects) |
| 5 | **Operator Workload** | Operator cannot continue to operate system | Severe reduction in the ability of operator to operate system | Major reduction in the ability of operator to operate system | Minor reduction in the ability of operator to operate system |
| 6 | **Financial Loss** | Above £1m | £250k to £1m | £10k to £250k | Less than £10k |
| 7 | **Security Breach** | Top Secret | Secret | Confidential | Restricted |
| 8 | **Reliability** | Total loss of functional capability | Severe reduction in functional capability | Significant reduction in functional capability | Slight reduction in functional capability |
| 9 | **Project Schedule** | Slip impacts on overall defence capability | Slip impacts on other projects (eg. life extension of existing system) | Slip results in major internal schedule reorganisation | Schedules republished |
| 10 | **Mission Impact** | Mission loss (operational) | Severe mission degradation (operational) | Slight mission degradation (operational) Mission loss (training) | Mission delayed (operational) Mission degraded (training) |
| 11 | **Criminal Liability** | Custodial sentence imposed | Large fine imposed (£5k plus) | Small fine imposed (up to £5k) | Conditional discharge etc. |
| 12 | **Civil Liability** | Multiple, large civil suits (£10k plus) | Single, large civil suit (£10k plus) | Multiple, small civil suits (up to £10k) | Single, small civil suit (up to £10k) |
| 13 | **Maintenance Burden** | Projected servicing schedules severely adversely affected | Unscheduled maintenance predictions severely adversely affected | Projected servicing schedules slightly adversely affected | Unscheduled maintenance predictions slightly adversely affected |
| 14 | **Political Impact** | Government falls | Minister resigns | Commons debate/National Press aware | Parliamentary Question/Local Press aware |
| 15 | **Delivered System Performance** | Design does not meet requirement in critical areas - leading to a failure to accept system | Design does not meet requirement in non-critical areas - leading to major modification programme | Impact on operating procedures | Some trivial deficiencies |

## 5.2 Acquire Information

To prepare the development of a ToVV, an overview over the intended purpose related available model information and knowledge about the System of Interest needs to be achieved. The REVVA approach is product-oriented, which implies that the information that needs to be collected mainly addresses the model and the System of Interest themselves, and does not directly include the modelling process followed by the supplier or the System of Interest development process (assuming that the System of Interest is a human-build system). However, if the model or the System of Interest is still under development, the process-oriented information becomes relevant for identifying the expectable, but not yet available model information and knowledge about the System of Interest, and the point in time when they are likely to be available. It is also likely that during the development of the ToVV the information and knowledge requirements become more clear, resulting in the need for additional information. Thus, besides acquiring the information and knowledge itself, it also is of extreme importance to identify source of information and knowledge, which can be consulted as soon as it becomes necessary.

### 5.2.1 Model Information

The information about the model that is or will become available significantly influences the selection (planning and execution) of applicable V&V activities. In the following, model information is classified by

1. the representation form of the model,
2. the sub-model layers, for which insight is provided,
3. the behaviour information.

Ad (1): Representation forms of a model

- The *Conceptual Model* is an abstracted and idealised description of the System of Interest, including the decomposition of the System of Interest into interacting subsystems, and the representation of properties of interest in the form of

attributes and their change over time. It is communicative, i.e., written in the language of the model's application domain, which is understood by experts of the domain.

- The *Formal Model* is the formal description of the Conceptual Model, compliant with a well-defined modelling formalism. It expresses the Conceptual Model quantitatively, unambiguously, and platform-independently, and thereby prepares several methods for its computational solution.

- The *Executable Model* technically implements the Formal Model and provides the additional information required for executing and operating the model on a computer or in a network of computers.

Ad (2): Sub-model layers include

- the overall model layer,

- layers of sub-models, which represent subsystems of the System of Interest, within a sub-models hierarchy,

- down to the atomic sub-model (objects), which are not further decomposed in this particular model.

Ad (3): Behaviour information can be distinguished as

- the symbolic description of model behaviour (e.g., functions, state transition graphs), often addressed as "static" model information, and

- interpreted behaviour data, often addressed as "dynamic model information".

Rarely this information will be directly available. It can be extracted from the model documentation, interviews with the modellers and programmers, design documents, trace file analysis, simulation experiments, program source code, or reverse engineering. A lack of model information in combination with insufficient time or budget to reconstruct this information can render a complete VV&A endeavour useless. Thus, it

is important to identify as many sources of model information as possible during the early VV&A steps.

### 5.2.2 Information about the System of Interest

Especially during validation, the model or simulation results are assessed giving consideration to the state of the art knowledge of reality. The less is known about the System of Interest, the less confidence in the validity of the model can be established by comparing the symbolic model or simulation results to the System of Interest. The less empirically measured quantitative data is available (which for example is the case for large scale combat simulations [Davis 1992]), the more important qualitative behaviour descriptions and knowledge of the structure of the System of Interest become.
The extent and quality of the real world knowledge, information, and data depends on the 'exploitability' of the System of Interest. Classification of knowledge of a System of Interest can be done as proposed in [Zeigler, Praehofer & Kim 2000], reproduced in Table 5.

**Table 5: Levels of system knowledge**

| Level | Name | What we know at this level |
|---|---|---|
| 0 | Source | What variables to measure and how to observe them |
| 1 | Data | Data collected from a source system |
| 2 | Generative | Means to generate data in a data system |
| 3 | Structure | Components at a lower level coupled together to form a generative system |

The sources of information are identified by the scientific research and technical experimentation associated with the problem domain. Data eventually can be collected during test and observation of the System of Interest or a related system, although some system adversary properties (which are also likely to be the reason that the system is actually modelled and simulated) may complicate data collection. Generative means usually require at least the identification of stochastic dependencies and are

known by SME. Structural knowledge is gained by, e.g., physical decomposition of the System of Interest.

### 5.2.3 Repository

To enhance reuse, as much information about the System of Interest and the model as technically and legally possible, should be stored in an appropriate repository (see section 6).

## 5.3 Develop ToVV

The leaf nodes of the ToA are likely to be interpretable by the customer, but not measurable by the V&V Executioners. Then there is a need for expanding the ToA. The leaf nodes of the ToA are to be decomposed into still smaller and smaller acceptability sub-criteria, until the V&V Executioners decides that all the new leaf nodes are directly measurable given the M&S asset, the real world knowledge, and the V&V techniques and tools. This expansion of the ToA is called Target of Verification and Validation (ToVV).

Where in the ToA the criteria are customer-oriented and deal with the use of the M&S asset, the ToVV criteria deal with the inner workings of the asset and the approach how to substantiate that the Acceptability Criteria are met. This level of detail is usually beyond the scope of the customer. Building the ToVV is largely similar to building the ToA, but then on a lower level and with other experts instead of the customer, taking into account which knowledge and information about both the model and the System of Interest is available.

With the well-defined hierarchy of Acceptability Criteria, an overview over the available model information, the available system knowledge, a choice of V&V techniques, and a suite of tools at hand, a strategy how to substantiate the claim that each individual Acceptability Criterion is passed can be developed. After completion of the ToVV the leaf-nodes indicate V&V tasks that need to be performed and the evidence that needs to be acquired, as sketched in Figure 8. Usually the V&V Leader is in charge of

developing the ToVV, but parts of it could also be developed by the Acceptance Leader or the customer, if they have a clear idea of how they want the examination of an Acceptability Criterion to be performed.

Thus, developing the ToVV involves the broad steps:

- Expand the ToA
- Technique selection

A summation of this section is given in the following box as a quick guide for practitioners; references to paragraphs where the items are elaborated upon are given.

Build ToVV structure. Use hints in section 5.1.

- copy the Acceptability Criteria from the ToA and place them on the highest level of the ToVV,
- derive sub-criteria from these Acceptability Criteria where necessary and give decomposition arguments. Continue until all criteria are directly measurable and appropriate knowledge and information about the model and the System of Interest is available to do so. The measures used are now likely to be Measures of Performance (MoP), which indicates that they involve system details, as opposed to MoE dealing with the systems outside.

Determine V&V tasks for each leaf in the ToVV.

- V&V technique selection including the convincing force and probative force needed (use hints in paragraph 5.3.1),
- handle managerial issues (use hints in paragraph 5.3.2).

### 5.3.1 Technique Selection

A variety of V&V techniques is known. The selection of a techniques mainly depends on whether the V&V executioner is capable of implementing the technique within time and budget (skill, tools), whether all the input to conduct the techniques is available (prepared in phase 2), and whether the techniques promise to produce sufficiently meaningful and reliable results (convincing force and probative force, prepared in phase 1).

The application of the technique can result in a direct substantiation but it may also be the case that a series of techniques must be applied (indirect substantiation).

#### 5.3.1.1 Direct Substantiation

Ideally, the Acceptability Criterion requires the M&S product to meet a directly measurable property, e.g., the determinable distribution of a particular goal parameter must fit the known distribution of its measurable real world counterpart on a given significance level. In this case, the Acceptability Criterion can be substantiated by direct implementation of an appropriate statistical test.

#### 5.3.1.2 Indirect Substantiation

More likely is that many Acceptability Criteria cannot directly be substantiated by the implementation of a single V&V technique. The tree presented in Figure 9 aids the identification of aspects of the M&S product that should be subjected to V&V, and shows both their demarcation and dependencies. It shows the relationships within a subset of relevant V&V aspects, and points out that, e.g., V&V of the model and V&V of the experimental frame support the V&V of the simulation results, or, another e.g., that for V&V of the symbolic model specification also its embedded data should be subjected to V&V. Nevertheless it also shows that there actually is a difference between the V&V of input data models and V&V of runtime input data, which must not be ignored.

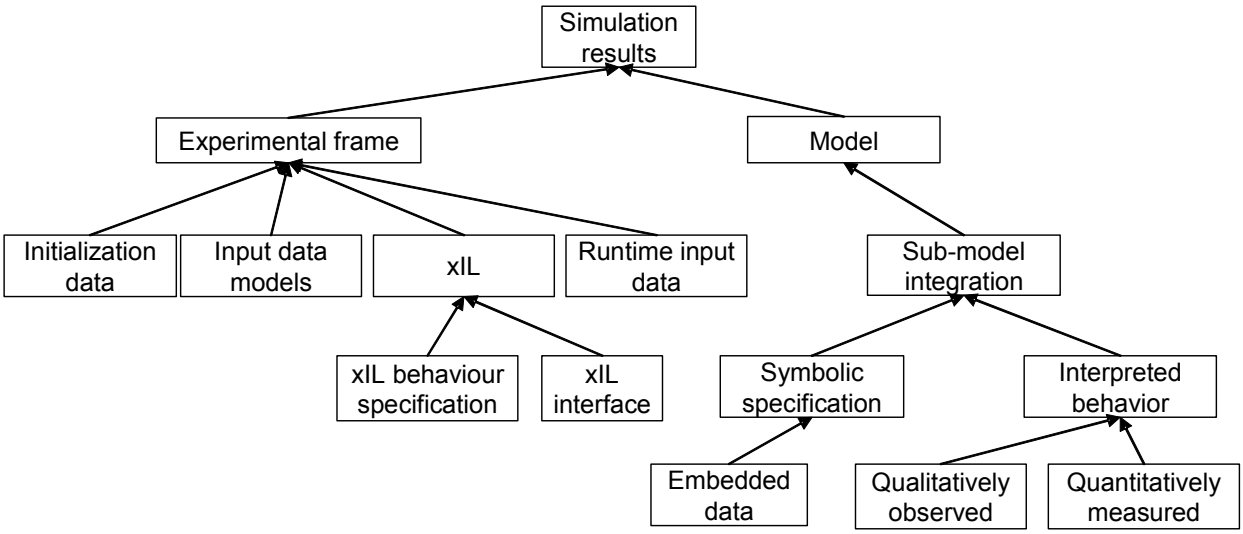



**Figure 10: V&V taxonomy**

Using the dependencies between the M&S assets depicted in Figure 10 and the knowledge about the dependencies in the model information space, V&V approaches for the indirect substantiation of the Acceptability Criteria can be developed. Note, the taxonomy tree does not explicitly address errors that originate from incompatibility of the items associated with individual distinct nodes – the assessment of the simulation results does not only require the assessment of the used model and the used experimental frame, but also whether the model is valid within the experimental frame.

### 5.3.2 Managerial Issues

The development of a V&V approach is accompanied by the administrative and organisational aspects of V&V planning. A schedule must be created, and roles should be assigned to actors (allocation of human resources), which extends the “V&V plan”. The issues sketched in section 2.3 should be considered then. Many aspects of these managerial issues can also be stored in the ToVV leaf nodes.

## 5.4 Conduct V&V

The Items of Evidence are acquired as required by the ToVV. The V&V Executioners chooses the appropriate techniques and tools (if not predefined in the ToVV) to produce the evidence with the required probative force. As a result of conducting V&V activities,

the “atomic building blocks” of the V&V effort, the *Items of Evidence*, are created. An Item of Evidence is exhaustively documented, and its creation must be transparent and repeatable. It can be reproduced, whenever required.

It must be assumed that the implementation of the V&V activities according to the ToVV may run into a dead-end road in some places, which requires the adaptation of the ToVV to the new situation. Changes in the ToVV should be done very carefully, as a major degradation is likely to results in an unacceptably weak V&V effort.

## 5.5 Assess Items of Evidence

The ToVV identifies Items of Evidence, which must provide a “sufficiently strong substantiation” for the claim that the Acceptability Criteria are passed. Although each well-defined Acceptability Criterion is analysable or testable, depending on the reliability of the analysis or test itself, a deficiency of the M&S product may remain undetected. The *probative force* of each individual Item of Evidence is assessed based on the repeatability of the associated V&V activity. This assessment should immediately follow the creation or acquisition of the Item of Evidence, to keep delays due to maybe necessary iterations and improved repetition of V&V activities as short as possible. Here it is assumed that the V&V Leader assesses the Items of Evidence produced by the V&V Executioners. However, all Items of Evidence will also be fully accessible for the Acceptance Leader, allowing an extra, but probably not as prompt assessment, if required.

Criteria for the assessment of the probative force of an Item of Evidence add objectivity to this currently subjective procedure. The probative force of an Item of Evidence is considered to be high, if the V&V result is reproducible, independently from its subjective elements (human beings). It is considered to be low, if it strongly depends on its subjective elements and its various results depend on the different individuals involved. If the repeatability cannot be directly observed, the following factors, which are likely to influence repeatability, should be considered:

- Technique objectivity: Degree of impact of potential human bias. Ranges from direct, undirected human judgement (worst) to formulation and formalisation of constraints or desirable properties, which are used to assess the model indirectly (best).

- Evaluated model information: Degree of insight into the model. Reliability, amount, and level of detail of evaluated model documentation, including structure and behaviour descriptions in their different representation forms, ranging from an opaque model (worst) to detailed in-depth insight into the model, its submodels, the different representation forms, model description and computed model behaviour (best).

- Evaluated model data and test case design: Density and breadth of simulation data evaluated, access to model “internal” data (e.g., state parameter values, internal interactions), and maturity of the taken approach to simulation experiment set-up for testing purposes. Ranges from limited sets of random behaviour samples (worst) to the ability for exhaustive white-box testing (best).

- Consulted body of system knowledge and comparison data: Variety of sets comparison data, their origin, age, and measurement method, ranging from pure postulation (worst) to exhaustive system knowledge supported by sound theory and exhaustive data (best).

It is the V&V Leader’s responsibility to assess the probative force of each individual Item of Evidence as low, medium, or high, based on the above criteria. If the assessed probative force of the Item of Evidence is lower than required by the ToVV, it needs to be strengthened by repeated, improved execution of the V&V activity, or the ToVV needs to be adjusted in a manner, that additional, supporting evidence can be provided. If this is not the case, the Acceptability Criterion supported by the part of the ToVV and the particular Item of Evidence must not be considered as substantiated, or should at least be marked as questionable.

## 5.6 Assess Evidence Integration

A single Item of Evidence will usually not allow the conclusion that a particular Acceptability Criterion is passed, but several Items of Evidence are assembled according to the (most probably adjusted) ToVV. The evidence is assembled by the V&V Leader, and the convincing force of the assembly is assessed by the Acceptance Leader.

The convincing force of sub-criteria hierarchy as documented in the ToVV is an expression of the preciseness and coverage of the Acceptability Criteria, as defined in the ToA. It ranges from “fragmentarily addressed” (worst) to “completely covered” (best). An Acceptability Criterion is considered to be completely covered, if the rationale for the derivation of directly succeeding acceptability sub-criteria makes clear that meeting the sub-criteria automatically implies meeting the parent criterion, too. However, there may be gaps between the selected acceptability sub-criteria, left there unknowingly or deliberately (time, budget). The larger those gaps are the less convincing is the ToVV. Lacking a matured formal framework for precisely determining the convincing force of decomposition (and re-composition), the assessment of the convincing force of each step of evidence assembly is at the Acceptance Leader’s discretion.

## 5.7 Evaluate V&V Report

When no disproving evidence has been acquired or created, the affirmative evidence is considered to be “strong enough”, and the strategy according to which the affirmative evidence is assembled to substantiate the claim that the Acceptability Criteria are met is considered to be “sufficiently convincing”, then the M&S product is perceived as correct with respect to all relevant specifications and constraints, and as valid for its intended purpose, as represented by the ToA. However, as the evidence may be erroneous and its assembly according to the ToVV may contain gaps, there is always some residual uncertainty concerning this perception. To prepare a responsible acceptance or rejection decision, an upper bound for this residual uncertainty is determined.

Criteria for the determination of the residual uncertainty include:

- Coverage of the intended purpose of the M&S product by the specified Acceptability Criteria: Degree to which risks associated with the use of the M&S product are mitigated by demonstrating that the defined Acceptability Criteria are passed. New insight gained during V&V implementation may reveal previously undetected gaps among the Acceptability Criteria.

- Coverage of the specified Acceptability Criteria: Share and relevance of Acceptability Criteria that have been addressed during V&V. Time and budget constraints, or technical unfeasibility, may force the V&V Leader to omit the substantiation of some Acceptability Criteria.

The level of residual uncertainty quantifies the uncertainty associated with the perception of correctness and validity of the M&S product. The M&S product needs to be perceived as correct and valid (i.e., acceptable) to be rated on the residual uncertainty scale. The level of residual uncertainty is defined in dependence of the convincing force of the decomposition of Acceptability Criteria (ToVV) and the probative forces of the Items of Evidence, according to the following (intuitive) regulations:

- The higher the convincing force of the ToVV and the probative forces of the Items of Evidence, the lower becomes the residual uncertainty.

- High probative force is the indispensable prerequisite for a low residual uncertainty. This postulation is based on the assumption that for a lower residual uncertainty every single Acceptability Criterion defined in the ToA needs at least to be subjected to examination. If the coverage of Acceptability Criteria is incomplete, a significant residual uncertainty associated with the perceived correctness and validity of the M&S product with respect to the *complete* ToA remains.

- The minimum residual uncertainty is only achieved by a combination of maximum convincing force and maximum probative force. All Acceptability Criteria need to

be covered, and the perception of the correctness and validity needs to be close to proof.

The level of residual uncertainty needs to be identified for each Acceptability Criterion and each relevant set of Acceptability Criteria individually. While for particular Acceptability Criteria a high degree of uncertainty is acceptable (criteria which may be failed without serious consequences), for others only very low uncertainty is acceptable (criteria whose failure will have serious impact).

The semantics of each single level are given by the description of the influences of the M&S product on the real world that are considered to be responsibly acceptable. The levels *high*, *medium*, *low*, and *very low* are distinguished.

- The residual uncertainty is **high:** Although there is some indication that the Acceptability Criteria are met, the M&S product must not play a relevant role in the context of its intended purpose, when there is any non-negligible credible worst case impact. Otherwise there must be other information sources consulted to compensate incorrect or invalid M&S influences. High residual uncertainty is the consequence of a low convincing force of the ToVV, regardless of the achieved probative forces.

- The residual uncertainty is **medium:** The M&S product must not play a relevant role in the context of its intended purpose, when there is any critical or catastrophic credible worst case impact. Otherwise there must be other information sources available to compensate incorrect or invalid M&S influences. Medium residual uncertainty is the consequence of a medium convincing force, regardless of the achieved probative forces.

- The residual uncertainty is **low:** The M&S product may play a relevant role in the context of its intended purpose, if the credible worst case impact is not catastrophic. Low residual uncertainty is the consequence of a high convincing force, supported by Items of Evidence of medium probative force.

- The residual uncertainty is **very low:** The M&S product may play a relevant role in the context of its intended purpose, even if the credible worst case impact is catastrophic. Low residual uncertainty is the consequence of a high convincing force, supported by Items of Evidence of high probative force.

The achieved level of uncertainty, regardless how low, must never imply that the results of the use of the M&S product can be blindly transferred to the real world.

# 6 EXPERIENCE CAPITALISATION

An M&S product is likely to be used several times during its life, with or without modification, for closely or less closely related intended purposes, with data originating from different sources, or in any other somehow changing configuration. From the history of use important information for the new intended use can be retrieved, and especially verification results can be reused, as long as a particular subset of Acceptability Criteria remains unchanged. For V&V often a large amount of comparison information is required, and the additional information about an (executable) model collected during a previous V&V effort might be extremely valuable during following V&V activities. The large amount of information needs to be organised and managed, which encourages the use of a repository. An M&S V&V repository can hold the following information:

- M&S product identification: A key to access a particular M&S product.

- Supplemental model information: All information that provides additional insight into the (executable) model or the simulation results, as, e.g., the conceptual model or the formal model should be included in the repository.

- Intended purpose and V&V information: The multiple uses of the M&S product are recorded, and those V&V activities, evidence, and other information items associated with this use clearly identified. If the M&S product is modified for the intended use, also configuration control must be managed.

- Read/write access: It needs to be distinguished by role, who may read and who may write which parts of the repository. As supported by every modern version management system, multiple parallel write access must be regulated.

- Security and privacy: Military models often contain or access confidential data, which must not be revealed to any unauthorised person. Commercial interests of model developers also require responsible storage of sensitive data provided by

them. For each M&S product and their associated roles access rules need to be defined which make sure that no security and privacy requirements are violated.

- Configuration control: The M&S product is likely to be used in different configurations, and each configuration might hold information that is also of interest for the other configurations.

- Non-monotonic management of data/knowledge. VV&A leads intrinsically to invalidation of previous beliefs on results, models and Experimental Frame. The knowledge gained in this case must be captured and documented.

The above list is a first brain storming of issues that should be addressed when planning an M&S V&V repository and is incomplete and needs to be continued. The contents, structure and architecture of such a repository need to be carefully planned and designed, considering the needs of all roles involved.

# 7 REFERENCES

Brade, D. 2004. A Generalized Process for the Verification and Validation of Models and Simulation Results. Dissertation, Fakultät für Informatik, Universität der Bundeswehr München.

CRIT. 2004. THALES JP11.20 Report JP1120-WE1200-D1201-CRIT-V1.0.

Davis, P.K. 1992. Generalizing Concepts and Methods of Verification, Validation, and Accreditation (VV&A) for Military Simulations. Report published by RAND, Santa Monica, CA. ISBN 0-8330-1298-3.

Kotonya G., I. Sommerville,' Requirements engineering: Processes and Techniques'. John Wiley and Sons: Chichester, 1998.

Lane N.E., E.A. Alluisi, 'Fidelity and Validity in Distributed Interactive Simulation: Questions and Answers'. IDA Document D-1066. Institute for Defense Analysis: Alexandria, November 1992.

LEVELS. 2004. THALES JP11.20 Report JP1120-WE1300-D1301-LEVELS-V1.0.

METHGU2. 2004. THALES JP11.20 Report JP1120-WE5100-D5101-METHGU2-V1.1.

Muessing, P., D. R. Laack, and J. J. Wrobleski, Jr. 1997. Optimizing the Selection of VV&A Activities: A Risk/Benefit Approach, Proceedings of the 1997 Summer Computer Simulation Conference, Society for Computer Simulation International.

Mugridge, C. 1999. Verification, Validation and Accreditation of Models and Simulations Used for Test and Evaluation – a Risk/Benefit Based Approach. Internal Report, Technical Development Group, Defense Evaluation and Research Agency UK.

OPTIM. 2004. THALES JP11.20 Report JP1120-WE4200-D4201-OPTIM-V1.0.

PROC. 2004. THALES JP11.20 Report JP1120-WE1400-D1401-PROC-V1.0.

QUALIF. 2004. THALES JP11.20 Report JP1120-WE4300-D4301-QUALIF-V1.0.

Rae, A., P. Robert, and H.-L. Hausen. 1995. Software Evaluation for Certification: Principals, Practice, and Legal Liability. McGnaw-Hill, London, UK

Schmidt, B. Fachberichte Simulation. 1985. Systemanalyse und Modellbau – Grundlagen der Simulationstechnik. Springer-Verlag.

Shannon, R.E. 1975. Systems simulation and the art of science. Prentice Hall, Eaglewood Cliffs, N.J.

Sproles. 2000. Coming to Grips with Measures of Effectiveness. The Journal of the International Council on Systems Engineering, Vol. 3, No.1.

TAXO. 2004. THALES JP11.20 Report JP1120-WE1100-D1101-TAXO-V1.0.

TECH2. 2004. THALES JP11.20 Report JP1120-WE3100-D3102-TECH2-V2.0.

TOAGUID. 2004. THALES JP11.20 Report JP1120-WE2200-D2201-TOAGUID-V1.0.

XPCAP. 2004. THALES JP11.20 Report JP1120-WE4100-D4101-XPCAP-V1.1.

Zeigler, B.P., H. Praehofer, and T.G. Kim. 2000. Theory of Modeling and Simulation. Second Edition. Academic Press. ISBN: 0-12-778455-1