Draft of Causal Explanations

Document information	
Project Title	EMFASE
Project Number	E.02.32
Project Manager	University of Trento
Deliverable Name	Draft of Causal Explanations
Deliverable ID	D3.1
Edition	00.01.00
Template Version	03.00.00
Task contributors	
University of Trento; SINTEF; Deep Blue; University of Southampton	

Please complete the advanced properties of the document

Abstract

The objective of EMFASE WP3 is to provide causal explanations of the phenomena observed in the empirical evaluations. The purpose of such explanations is to provide a better understanding of the underlying mechanisms of (the application of) risk assessment methods and thus to support the development of risk assessment method selection guidelines. In this deliverable we present an initial theory of causal explanation that it is based on existing theories, but it has been specialized for security risk assessment methods based on the results of the empirical studies conducted in the project.

Authoring & Approval

Prepared By - *Authors of the document.*		
Name & Company	Position & Title	Date
Federica Paci / UNITN - UoS	WP3 Leader	24/02/2015
Katsiaryna Labunets/UNITN	Project Contributor	23/02/2015
Martina De Gramatica / UNITN	Project Contributor	24/02/2015
Bjørnar Solhaug / SINTEF	WP1 Leader	28/01/2015
Martina Ragosta/DBL	Project Contributor	18/02/2015

Reviewed By - *Reviewers internal to the project.*		
Name & Company	Position & Title	Date
Elisa Chiarani / UNITN	Project Manager	25/02/2015
Martina Ragosta/DBL	Project Contributor	25/02/2015
Bjørnar Solhaug / SINTEF	WP1 Leader	25/02/2015

Reviewed By - *Other SESAR projects, Airspace Users, staff association, military, Industrial Support, other organisations.*		
Name & Company	Position & Title	Date
<Name / Company>	<Position / Title>	<DD/MM/YYYY>

Approved for submission to the SJU By - *Representatives of the company involved in the project.*		
Name & Company	Position & Title	Date
Fabio Massacci / UNITN	Project Coordinator	<24/03/2015>

Rejected By - *Representatives of the company involved in the project.*		
Name & Company	Position & Title	Date
<Name / Company>	<Position / Title>	<DD/MM/YYYY>

Rational for rejection
None.

Document History

Edition	Date	Status	Author	Justification
00.00.01	24/11/2014	Working Document	Federica Paci/UNITN/UoS	New Document
00.00.02	16/12/2014	Working Document	Bjørnar Solhaug / SINTEF	Draft of Section 2.2
00.00.03	17/12/2014	Working Document	Martina De Gramatica/UNITN	Draft Section 3
00.00.04	18/12/2014	Working Document	Martina Ragosta/DBL	Draft Section 2.1
00.00.05	18/01/2015	Working Document	Katsiaryna Katsyarina	Revision Section 3

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

			Labunets/UNITN	
00.00.06	28/01/2015	Working Document	Bjørnar Solhaug / SINTEF	Finalization of Section 2.2
00.00.07	18/02/2015	Working Document	Martina Ragosta/DBL	Finalization of Section 2.1
00.00.08	20/02/2015	Working Document	Katsiaryna Katsyarina Labunets/UNITN	Revision Section 4 and 5
00.00.09	22/02/2015	Working Document	Federica Paci/UoS	Added Abstract, Executive Summary and Introduction
00.00.10	23/02/2015	Working Document	Katsiaryna Labunets/UNITN	Revision Section 4 and 5
00.00.11	22/02/2015	Final Document	Federica Paci/UoS	Finalization of the document
00.00.12	25/02/2015	Final Document	Elisa Chiarani/UNITN	Quality check
00.00.13	25/02/2015	Final Document	Bjørnar Solhaug / SINTEF	Internal Review
00.00.14	26/02/2015	Final Document	Martina Ragosta/DBL	Internal Review
00.00.15	26/02/2015	Final Document	Katsiaryna Labunets/UNITN	Address Internal Reviewers Comments
00.00.16	27/02/2015	Final Document	Federica Paci/UoS	Address Internal Reviewers Comments. Final draft for PO
00.01.00	24/03/2015	Final Document	Fabio Massacci	Approved for submission

Intellectual Property Rights (foreground)

This deliverable consists of foreground owned by one or several Members or their Affiliates.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

Table of Contents

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

List of tables

List of figures

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

Executive summary

The objective of EMFASE WP3 is to provide causal explanations of the phenomena observed in the empirical studies. The purpose of such explanations is to provide a better understanding of the underlying mechanisms of (the application of) risk assessment methods and thus to support the development of risk assessment method selection guidelines. The causal explanations will be built upon existing theories, but they will be specialized for security risk assessment methods and refined based on the empirical results of the project.

This document provides a survey of theories from different fields to identify candidates for explaining and exploring the mechanisms of security risk assessment. In addition, the document shows how candidate theories are applied to the results of the first round of empirical studies that we conducted within EMFASE and it provides an initial version of causal explanations for the evaluation results.

More specifically, this document makes the following contributions.

- An overview of state of the art on theories explaining why a method is successful in achieving is intended objectives;
- An initial interpretation of the results of the empirical studies conducted within EMFASE based on the Method Evaluation Model selected as candidate theory;
- An initial theory of causal explanations built upon the Method Evaluation Model and the results of the empirical studies. The theory shows how different features of risk assessment methods have different impact on the actual effectiveness and perception of these methods;
- An explanation of the results based on the initial theory of causal explanations.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

1 Introduction

1.1 Purpose of the document

The objective of EMFASE WP3 is to provide causal explanations of the phenomena observed in the empirical studies. The purpose of such explanations is to provide a better understanding of the underlying mechanisms of (the application of) risk assessment methods and thus to support the development of risk assessment method selection guidelines.

This document provides a survey of theories from different fields to identify candidate theories for explaining and exploring the mechanisms underlying a security risk assessment process. The document justifies why the Method Evaluation Model (MEM) has been selected as candidate theory. In addition, the document shows how the MEM model is applied to the results of the first round of empirical studies that we conducted within EMFASE. It also provides an initial theory of causal explanations for the evaluation results built on top of the Method Evaluation model and the results of the empirical studies.

The document is structured as follows. In Section 2, we give an overview of the state of the art on theories explaining why a method is successful in achieving its intended objectives and we justify why we have selected the Method Evaluation Model as candidate theory. In Section 3, we evaluate the results of the first round of empirical studies based on the Method Evaluation Model and we show that the model cannot explain our results. In Section 4, we propose an initial theory of causal explanations built upon the Method Evaluation Model and the results of the empirical studies. In Section 5, we conclude the documents by showing how the proposed theory of causal explanations provides an explanation to our experimental results.

1.2 Intended readership

D3.1 is mainly an internal working document for EMFASE. Thus, intended readers of this document are primarily the EMFASE project partners and the EUROCONTROL. This document is to be used by the members of the project EMFASE as it provides an initial theory of causal explanation for the results of the first round of empirical studies conducted within EMFASE.

In particular, the content of the document will be used as input/feedback to the activities of WP2 to define which are other possible empirical studies we need to conduct to validate the theory of causal explanations.

Other potential readers are generally all stakeholders within the ATM domain that need to take security into account in an operational area. More specifically, the document is of interest to all SESAR JU projects within the transversal areas of WP16 that are related to security management and risk assessment. For these stakeholders the document gives insight into some of ATM security risk assessment methods that could be relevant to apply or investigate further.

1.3 Inputs from other projects

The document does not make use of input from other projects, but the content is related to both SESAR 16.02.03 and SESAR 16.06.02. References to these projects are given in the relevant sections.

1.4 Acronyms and Terminology

Term	Definition
ATM	Air Traffic Management
E-ATMS	European Air Traffic Management System
SESAR	Single European Sky ATM Research Programme

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

Term	Definition
SJU	SESAR Joint Undertaking (Agency of the European Commission)
SJU Work Programme	The programme which addresses all activities of the SESAR Joint Undertaking Agency.
SESAR Programme	The programme which defines the Research and Development activities and Projects for the SJU.
TAM	Technology Acceptance Model
MEM	Method Evaluation Model
PEOU	Perceived Ease to Use
PU	Perceived Usefulness
ITU	Intention to Use

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

2 Theories for Method Evaluation – State of the Art

In this section we provide an overview of the theories used to compare methods success in the Information Systems and Software Engineering communities.

2.1 General Method Evaluation Theories

General design research in Information Systems and Software Engineering tends to emphasize the development of new methods while addressing the evaluation and comparison of existing methods in only a limited fashion [1][2].

The problem of "how to evaluate" Information Systems-related methods is a challenging issue (e.g.[11][3]). There are inherent problems evaluating any methodology or design technique since there is typically no theory, no hypotheses, no experimental design and no data analysis to which traditional evaluation criteria can be applied.

Similarly, Software Engineering researchers have traditionally been very poor at making evaluation theories explicit [4]. Many of the empirical studies conducted over the past few decades fail to relate the collected data to an underlying theory. As a consequence, results are hard to interpret, and studies cannot be compared.

The two most diffuse theories to evaluate and compare methods in the Information Systems and Software Engineering communities are the Technology Acceptance Model (TAM) and the Method Evaluation Model (MEM).

The Technology Acceptance Model (TAM) [6] is an information systems theory that models how users come to accept and use a technology and it suggests that when users are presented with a new technology, a number of factors influence their decision about how and when they will use it, notably:

- Perceived usefulness (PU) - This was defined by Fred Davis as "the degree to which a person believes that using a particular system would enhance his or her job performance".
- Perceived ease-of-use (PEOU) - Davis defined this as "the degree to which a person believes that using a particular system would be free from effort"

TAM is one of the most influential extensions of Ajzen and Fishbein's theory of reasoned action (TRA) in the literature. It was developed by Fred Davis and Richard Bagozzi [7]. TAM replaces many of TRA's attitude measures with the two technology acceptance measures— ease of use, and usefulness. TRA and TAM, both of which have strong behavioural elements, assume that when someone forms an intention to act, that they will be free to act without limitation. In the real world there will be many constraints, such as limited freedom to act.

The Method Evaluation Model, proposed by Moody in [13] is a theoretical model that is based on Technology Acceptance Model (TAM), and the Theory of Reasoned Action [14] and the Methodological Pragmatism from the philosophy of science [15].

The resulting theoretical model combines two different but related dimensions of method "success": actual effectiveness and adoption in practice. Actual efficacy is the pragmatic success of the method, i.e. the extent to which it improves the performance of the task in question. Adoption in practice is the extent to which the method is used in practice. These two dimensions are captured by the MEM as summarized in Figure 1. It consists of the following constructs.

- Actual efficiency: The effort required to apply a method;
- Actual effectiveness: The degree to which a method achieves its objectives;
- Perceived ease of use: The degree to which a person believes that using a particular method would be free of effort;
- Perceived usefulness: The degree to which a person believes that a particular method will be effective in achieving its intended objectives;
- Intention to use: The extent to which a person intends to use a particular method;
- Actual usage: The extent to which a method is used in practice.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

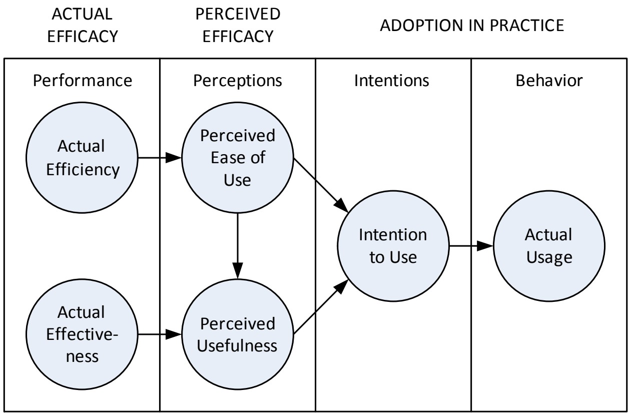

Figure 1: Method Evaluation Model

In MEM, Rescher's theory of Methodological Pragmatism predicts that methods that are more efficient and/or effective in achieving their objectives will be adopted in favour of other methods. This model proposes a slightly different view: those methods will be adopted based on perceptions of their ease of use and usefulness. Actual Efficiency and Effectiveness determine intentions to use a method only via perceptions of ease of use and usefulness. This is a subtle difference, but very important in human behaviour, subjective reality is more important than objective reality. While perceptions of ease of use and usefulness will be partly determined by actual efficacy, they will also be influenced by other factors (e.g. prior knowledge, experience with particular methods, normative influences).

The interesting MEM feature is that it provides hints and suggestion to support causal Explanations through "law's of interactions" for the observed phenomena. The idea of causality, or the relation between cause and event, is central to many conceptions of theory [15]. When theory is taken to involve explanation, it is intimately linked to ideas of causation. Often, to ask for an explanation of an event is to ask for its cause. Similarly, the ability to make predictions from theory can depend on knowledge of causal connections. The arrows between the constructs in Figure 1 depict the hypothesized causal relationships between the constructs. For example, perceived usefulness is determined by actual effectiveness and perceived ease of use. EMFASE investigates these constructs and causal relationships to understand which features or properties of SRA methods that may contribute to them.

Thus MEM is not just an evaluation model for methodologies, but it provides also causal and explanatory means to support analysis, to make prediction on efficacy and actual adoption of methods. MEM can also support a proper design of new methods or re-design/enhancement of old ones.

In fact, according to [5], five primary goals of theory discerned are: 1) Analysis, 2) Explanation, 3) Prediction, 4) Explanation and prediction, and 5) Design and action. These are reported and further explained in the following table.

And the MEM can be classified for sure as a theory for "Explanation and Prediction (see [5]), while its careful adoption can also inform design or re-design of methods under analysis. This is why we have selected MEM has candidate theory to explain the results of our empirical studies.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

THEORY TYPE	DISTINGUISHING ATTRIBUTES
Analysis	**Says what is:** The theory does not extend beyond analysis and description. No causal relationships among phenomena are specified and no predictions are made.
Explanation	**Says what is, how, why, when, and where:** The theory provides explanations but does not aim to predict with any precision. There are no testable propositions.
Prediction	**Says what is and what will be:** The theory provides predictions and has testable propositions but does not have well-developed justificatory causal explanations
Explanation and prediction	**Says what is, how, why, when, where, and what will be:** The theory provides predictions and has both testable propositions and causal explanations.
Design and action	**Says how to do something:** The theory gives explicit prescriptions (e.g., methods, techniques, principles of form and function) for constructing an artifact.

Table 1: A taxonomy of theory types in Information System Research (taken and adapted from [5])

2.2 Method Evaluation Theories for Security Methods

There are very few existing theories or frameworks that are specific for evaluating methods for security assessment or management similar to what we are developing in the EMFASE project. Such a theory should explain how various features or techniques of SRA methods contribute the success of the method, and which aspects of method success that are improved. The theory should be based on empirical findings, and it should allow the prediction of the comparative success of a given SRA method.

Although there is a lack of theory, there are nevertheless several works that has been done on comparing SRA and other methods for security assessment. Some of these include identified criteria for method evaluation and comparison, and some make attempts towards theory development.

One way of comparing methods for secure systems development in general (such as security requirements engineering, security design or security risk assessment) is to compare the features of the methods. For example, does the method target specific security properties such as confidentiality, does it come with support for threat identification and modelling, does it involve the identification of security mechanisms or controls, etc.? Such a comparison is done, for example, by Fabian et al. [16] on a number of security assessment and requirements engineering methods. The comparison is done with respect to a set of security requirements concepts, and therefore applies only to methods that cover such concepts. A more general evaluation framework for security engineering methods, tools and techniques is proposed by Busch et al. [17]. The framework shall aid researchers and engineers in identifying the engineering artefacts that are most suitable for solving a specific problem during the system development process. However, neither of these works propose any theories, and they do not evaluate or empirically investigate the extent to which given methods are successful.

Vorster and Labuschagne [18] present a framework for comparing different information security risk analysis methodologies using quantifiable criteria. There is no explanation of why the proposed criteria are used, which makes them seem somewhat arbitrary. The idea is nevertheless that analysts and other stakeholders shall weight the importance of each criterion and on that basis use the

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

framework to select the preferred method. The EMFASE framework rather uses criteria that we identified prior to the development of the framework, after which we empirically investigate how and the extent to which these criteria contribute to the success of SRA methods. A comparison of five methods is presented in [18], but only analytically and at a quite high level. No empirical studies are presented, and also no theories on what make an SRA method successful.

Hong et al [19] aim to develop such a theory for information security management, partly motivated by the observation that very few empirical studies have investigated the effectiveness of management tools and strategies. Five related theories are presented and combined covering information security policy, security risk management, security control and audit, security management and contingency management. The theory building is largely based on literature review and consists in decomposing each of the five parts of the theory into the elements that contribute to its fulfilment. The goal is that the theory shall help make predictions about what makes a method for security management effective, but no evidence of the validity of the theory is given. The authors rather state that empirical studies are required.

Fenz and Ekelhart similarly observe that most research on information security risk management (ISRM) aim to improve methods, but that there is a lack of thorough verification, validation and evaluation of the developed approaches. They moreover argue that "methodologically sound and comparable verification, validation, and evaluation results are crucial for measuring and understanding the implications of applied ISRM approaches" [20]. However, rather than presenting a framework or theory for evaluating or comparing SRA methods regarding their success, they discuss which strategies stakeholders should use for verification, validation and evaluation in each step of the standard risk assessment process.

Diallo et al. [21] make a comparative evaluation of thee specific approaches, namely the Common Criteria, misuse cases and attack trees. The reason for the selection of these three is that they are quite complementary and therefore cover different needs. The approaches were evaluated with respect to five evaluation criteria that were identified in a post hoc manner by the authors after applying the methods on an example case. The authors stress the need for established, standard criteria for comparison and evaluation of security, and the lack of such criteria is a weakness of the presented work. The authors themselves on a selected example do the evaluation, and it is purely qualitative and not very precise.

Other approaches to evaluating methods for security management are much more general. Siponen and Wilson [22], for example, make an analysis of established international standards. They argue that the specific needs of organizations must be taken into account, and that security standards need to be validated. The analysis is rather high level, and no empirical investigations are conducted.

Considering existing approaches to the evaluation of methods or techniques for security management, there is clearly a need for theory building and for empirical validation of the theory. Research and development within security management and security engineering commonly involves some kind of evaluation of the developed artefacts, but there is no established theory that researchers can use for predicting the success of the artefacts. In particular, there is not established theory similar to what we are developing in the EMFASE project with clearly identified theory concepts and causal relationships between them. Some of the attempts of theory building propose such concepts and relationships, but offer little or no empirical evidence or validation of the theory.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

3 Explaining the empirical ‘results based on Method Evaluation Model

In this section we provide a summary of the results obtained from the experiments we run in EMFASE and we discuss if the Method Evaluation Model can explain our results. We start from the experiments where we compared textual versus visual security risk assessment methods and we conclude with the experiments where we investigated the impact of using catalogs of threats and controls in conducting a security risk assessment.

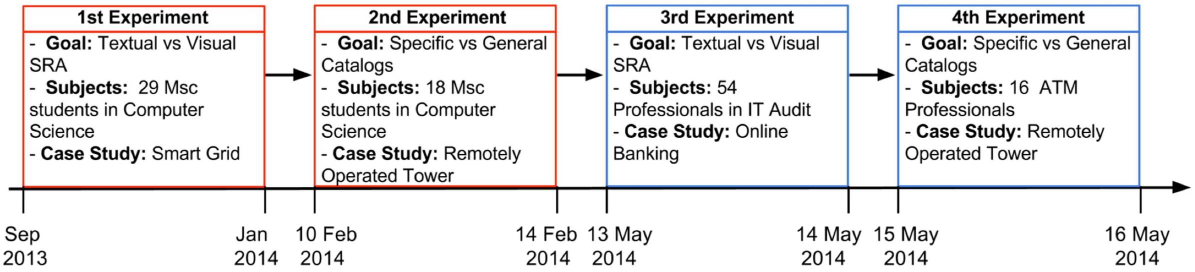
1st Experiment
- Goal: Textual vs Visual SRA
- Subjects: 29 Msc students in Computer Science
- Case Study: Smart Grid
2nd Experiment
- Goal: Specific vs General Catalogs
- Subjects: 18 Msc students in Computer Science
- Case Study: Remotely Operated Tower
3rd Experiment
- Goal: Textual vs Visual SRA
- Subjects: 54 Professionals in IT Audit
- Case Study: Online Banking
4th Experiment
- Goal: Specific vs General Catalogs
- Subjects: 16 ATM Professionals
- Case Study: Remotely Operated Tower
Sep 2013
Jan 2014
10 Feb 2014
14 Feb 2014
13 May 2014
14 May 2014
15 May 2014
16 May 2014

Figure 2: Experiments Timeline

3.1 Textual vs visual methods for security risk assessment

3.1.1 Students

The experiment involved 29 MSc students who applied both textual and visual methods to an application scenario from the Smart Grid domain. CORAS [23] was selected as instance of a visual method, and EUROCONTROL SecRAM [24] as instance of a textual method.

3.1.1.1 Results

The results show that there is no difference in the actual efficacy of the visual and textual methods because both methods generated a similar number of threats and controls (see Figure 3) and also of the same quality.

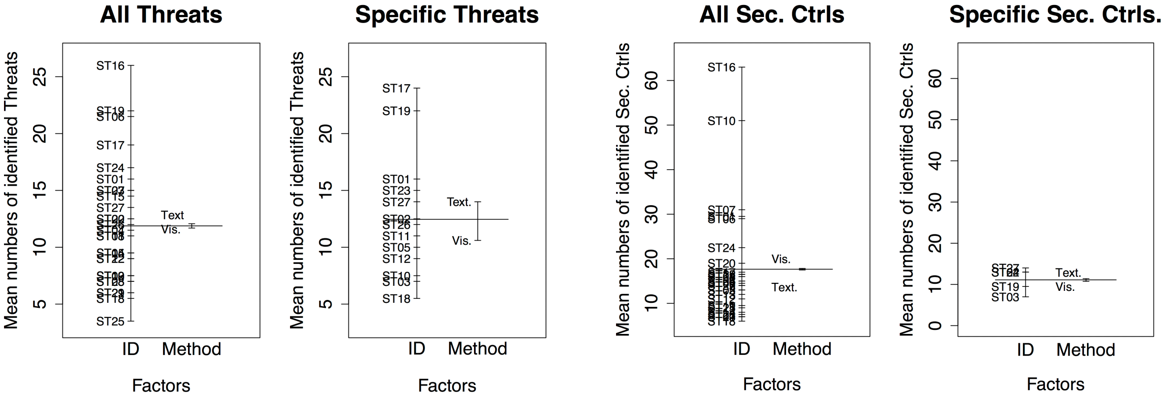

Figure 3: Actual Efficacy

With respect to perceived efficacy, participants prefer the visual method rather than the textual method. In fact, the visual method exhibits and higher perceived ease of use and perceived usefulness. The visual method also performed better with respect to participants’ intention to use the method.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

We also evaluated the causal relationship among methods' actual and perceived efficacy. We used the Kendall tau rank correlation coefficient. The test revealed that there is no causal relationship between the actual and perceived efficacy. The causal relationship among perception variables are instead supported as follows:

- Perceived Ease of Use → Perceived Usefulness (p=9,41*10^{-8}; tau=0,57)
- Perceived Ease of Use → Intention to Use (p=2,47*10^{-8}; tau=0,59)
- Perceived Usefulness → Intention to Use (p=1,43*10^{-12}; tau=0,76)

3.1.2 Professionals

The experiment involved 54 students with professional experience in IT Audit for Information Systems. The participants worked in groups to apply alternatively a visual method (CORAS) or a textual method (EUROCONTROL SESAR SecRAM) on a provided scenario targeting the Online Banking service.

3.1.2.1 Results

The analysis showed that the textual method has a slightly higher actual efficacy than the visual method because when the participants applied it they identified more threats and security controls, than with the visual method.

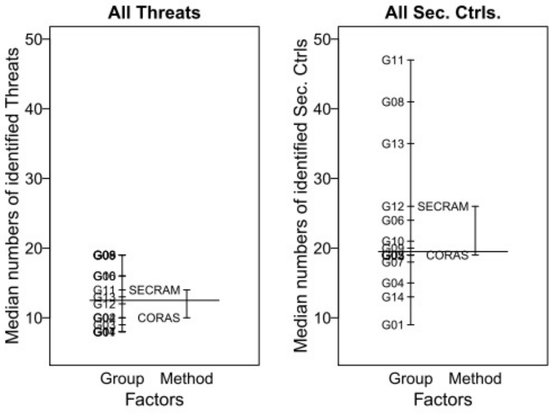

Figure 4: Actual Efficacy.

Although the textual method was more effective in the identification of threats and security controls, participants expressed a higher Preference towards the visual method over the textual method.

When we investigated the causal relationship among methods' actual and perceived efficacy with Kendall tau rank correlation coefficient, the test showed that there is no causal relationship between the actual and perceived efficacy. However, the causal relationship among perception variables are instead supported as follows:

- Perceived Ease of Use → Perceived Usefulness (p=0,02; tau=0,267),
- Perceived Ease of Use → Intention to Use (p=0,02; tau=0,25)
- Perceived Usefulness → Intention to Use (p=2,91*10^{-7}; tau=0,58)

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

3.2 Effect of using catalogues of threats and controls in the application of methods for security risk assessment

3.2.1 Students

The experiment involved 18 MSc students: half of them applied SESAR SecRAM with the domain-specific catalogues and the other half with the generic catalogues. Each group had to conduct a security risk assessment of the Remotely Operated Tower (ROT) operational concept.

3.2.1.1 Results

The actual efficacy of domain-specific and domain-generic catalogues was the same as they produced approximately a similar number of threats and controls and the quality of the threats and controls identified with the two type of catalogues did not significantly differ.

The overall perception of the catalogues for the domain-specific catalogue as reported by all participants. The same holds for the perceived usefulness of the method.

The analysis of the causal relationships in the MEM model showed that only the causal relationship between perceived usefulness and intention to use: the Kendall tau rank correlation coefficient has shown a strong positive correlation between methods' PU and ITU (p=0,017; tau=0,49).

3.2.2 Professionals

The experiment involved 15 professionals in the Air Traffic Management (ATM) who applied individually SESAR SecRAM on the Remotely Operated Tower (ROT) application scenario. The participants had all a good knowledge of the ATM domain but different level of knowledge in security: Only 5 out of 15 participants had an experience in security and thus were instructed to apply SESAR SecRAM without any catalogue. The other participants without security experience were divided in two groups: one group applied SESAR SecRAM with the support of the domain-specific catalogue while the second group worked with the support of the domain-generic catalogue.

3.2.2.1 Results

The analysis of the threats and controls identified by the three groups of participants shows that there is no difference in the actual efficacy of the two type catalogues because they produced a similar number of threats and controls and also of the same quality. However, an interesting result was that the participants who have not security knowledge performed the same as participants who had it but did not use the catalogues. In contrast, the perceived efficacy is higher for the domain-specific catalogue.

We also found that none of the causal relations in the MEM among actual efficacy and perceived efficacy variables and among perceived variables are supported by our data. This can be due to a small sample size (5 participant per type of catalogues).

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

4 Preliminary Theories

In this section we first discuss why the Method Evaluation Model cannot explain our experimental results and then we propose two theories that provide an explanation for our results.

The Method Evaluation Model asserts that actual efficacy of a method determines its perceived efficacy. This means that methods that are more efficient and effective in achieving their intended objects should also have a higher perceived ease of use and perceived usefulness respectively. However, our experiments results show that there is no causal relation between actual efficacy and perceived efficacy of security risk assessment methods and of catalogues. In fact our results have shown that even though there is no difference in the actual efficacy of textual and visual security risk assessment methods, the visual methods generally have a higher perceived efficacy. And the same holds for the difference between domain-general and domain-specific catalogues: there is no difference in the actual efficacy of the two types of catalogues but domain-specific catalogues have and higher perception. Thus, we came to the conclusion that there is no causal relation between actual efficacy and perceived efficacy of a method and that we need two different theories to explain which factors determine them.

4.1 A Theory for Actual Efficacy

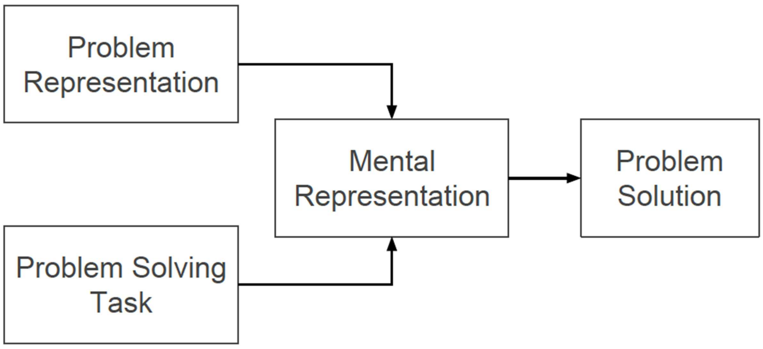

Figure 5: General problem-solving model

We resort on cognitive fit theory [25] to explain the results on actual efficacy. Cognitive fit theory links a problem representation with the efficiency and effectiveness of a problem-solving task. Figure 5 presents the general model of problem solving on which cognitive fit is based. The model views a problem solving as an outcome of the relationship between problem representation and problem-solving task. The mental representation is the way the problem solver represents the problem in human working memory. The mental representation is formulated using the characteristics of both the problem representation and the task. Specifically, it is derived from the interaction of the processes to act on the information in the problem representation and the problem-solving task. When the type of information emphasized in the problem-solving elements (problem representation and task) match, the problem solver can use processes (and formulate a mental representation) that also emphasize the same type of information. Consequently, the processes the problem solver users to both act on the problem representation and the task will match. The resulting consistent mental representation will facilitate the problem-solving process. Thus, cognitive fit theory suggests that performance of a task will be enhanced when there is a cognitive fit (match) between the information emphasized in the representation type and the information required by the task type.

In Section 5 we discuss how the cognitive fit theory can explain why our empirical studies have shown that visual and textual methods for security risk assessment methods have same actual efficacy and why the same results hold for domain-specific and domain-general catalogues.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

4.2 A Theory for Perceived Efficacy

The Method Evaluation model hypothesizes that Perceived Ease of Use of a method determines its Perceived Usefulness but does not provide an explanation of why there could be a difference among the two. To this end we propose our theory of causal explanations for perceived efficacy. The model is based on the qualitative results presented in [26] and [27]. Figure 6 sketches the theory with the main constructs and the causal relations among the constructs.

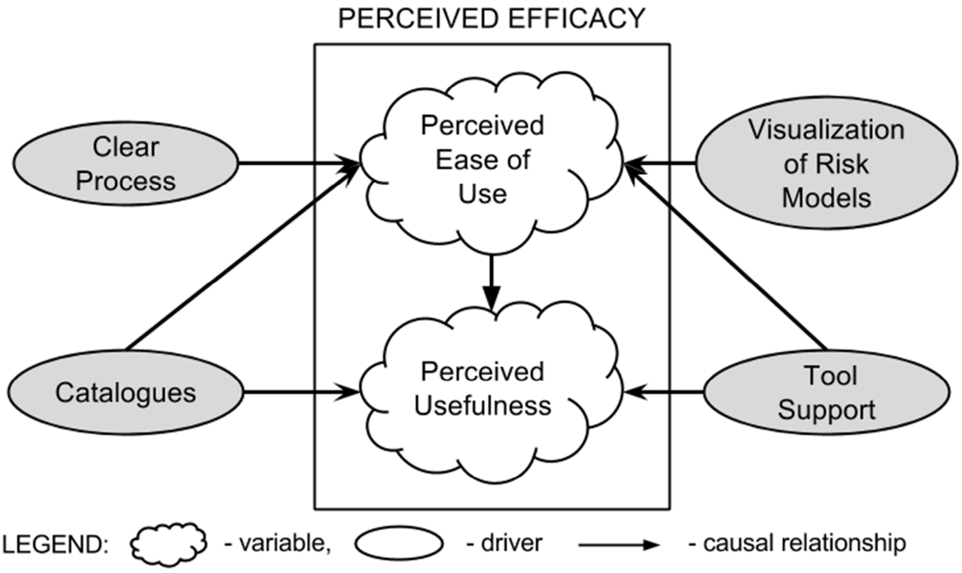

Figure 6: Preliminary Theory for Perceived Efficacy

The main constructs of our theory are:

- ***Perceived Ease of Use (PEOU):*** the degree to which a person believes that a particular method would be free of effort;
- ***Perceived Usefulness (PU):*** the degree to which a person believes that a particular method will be effective in achieving its intended objectives;
- ***Clear process***: a well-defined set of steps that guide analysts through security risk assessment process;
- ***Catalogues***: structured representation of state-of-the art security threats and controls;
- ***Visualization of Risk Models***: the representation of assets, threats and security controls and of the relations among them;
- ***Tool Support***: software solution that supports users in the execution of all the steps of the security risk assessment for example through the automation of some steps like risk level computation.

4.2.1 Causal Relationships

We assumed the following causal relationships between the constructs of the model:

- **Clear process may affect PEOU**. Clear process determines how easy is to understand and apply security risk assessment steps. Therefore, it affects users PEOU of method. If a method has clear process, this positively affects PEOU because user clearly understands how to follow the steps and conduct security risk assessment. On the contrary, if a user has doubts on how a step of the process should be executed, he will not perceive the method as ease to use.
- **Visualization of Risk Models may affect PEOU.** Risk model visualization may facilitate the overview of results of a security risk assessment, and thus my have a positive impact on

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

method's PEOU. However, if the visual notation does not scale, it may have a negative effect on method's PEOU.

- ***Catalogues may affect PEOU***. Catalogues can facilitate the identification of threats and controls especially for users who have no or limited security knowledge. Thus, the use of catalogues makes easier to conduct a security risk assessment and has positive effect on method's PEOU.
- ***Catalogues may affect PU***. Catalogues can improve the quality of threats and security controls and serve as checklist to control the completeness of security risk assessment results. Thus, the use of catalogues positively impacts a method's PU.
- ***Tool support may affect PEOU***. Tool can automatize the execution of a security risk assessment process (e.g. computation of risk level) or can facilitate the reporting of the results using an appropriate format (e.g. provide a set of tables that match method's steps). A well-designed tool can thus have a positive effect on method's PEOU. In contrast, a primitive tool can only have a negative impact on the user's PEOU of the method.
- ***Tool support may affect PU***. Tool can positively impact on a user's PU of the method because it automatizes the validation of the results of the security risk assessment process, i.e. using representation provided by tool it is easy to identified if something is missed or inconsistent. However, if the tool is poorly implemented, it decreases the user's perceived usefulness of the method.

5 Discussion

In this section we discuss how the results of our experiments can be explained by preliminary theories presented in previous section.

5.1 Visual vs. Textual security risk assessment methods

The fact that visual and textual methods showed similar actual efficacy can be explained by cognitive fit theory. Security methods that we studied in our empirical studies adopt similar security risk assessment process: first, you need to identify assets and possible threats that can harm them, then you need to evaluate impact and likelihood of each threat, and define a set of security controls to mitigate the most critical threats. Therefore, we can conclude that the task that users have to perform is the same for the studied methods, namely CORAS (visual), EUROCONTROL SecRAM (textual) and SESAR SecRAM (textual). The same holds for the information representation applied by the evaluated methods. Since for each method there is a cognitive fit between task and representation, the three methods exhibit similar actual efficacy.

Regarding perceived efficacy, in both experiments with students and professionals the visual method has higher PEOU than textual method. This is due to the fact that visual method supports visualization of risk models and has a clear process that have positive effect on method's PEOU. Statements made by participants during individual interview support this conclusion: "there are many summary diagrams which are useful to summarize what has been done", and "the advantage of CORAS is very clear structure".

Tool support plays an important role in method's PEOU. But it also explains why visual method has higher PU in the experiment with the students. In this experiment we compared CORAS (visual) and EUROCONTROL SecRAM (textual) methods. The visual method has some tool support with CORAS tool or Visio stencils, while the textual method does not have tool support at all. Participants had to create and maintain different tables manually: "It [tool support for textual method] is needed because it would save half of the time if the table were generated automatically".

5.2 Domain- specific vs Domain-General Catalogues

Both catalogue types have similar information representation, i.e. they provide a set of threats and security controls. The only difference is the scope of provided information. Thus, there is no difference in task type (threats and security controls identification) and representation type (two catalogues of

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

threats and security controls). This may explain why security risk assessment method applied with different types of catalogues showed similar actual efficacy.

The higher perceived efficacy reported for the domain-specific catalogues may be due to the fact that the catalogues structure facilitates the navigation of the catalogues and thus the identification of security threats and controls.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

6 Conclusions

MEM and TAM theoretical models are widely used in the Information System and Software Engineering community to evaluate and explain actual efficacy and perceived efficacy of methods. MEM hypothesizes that actual efficacy of a method determines its perceived efficacy. This means that methods that are more efficient and effective in achieving their intended objects should also have a higher perceived ease of use and perceived usefulness. However, our experimental results have shown there is no causal relation between actual efficacy and perceived efficacy of the evaluated methods and catalogues. Thus we proposed two different theories to explain our results on actual efficacy and perceived efficacy of security risk assessment methods and catalogues of threats and security controls. To explain the results on actual efficacy we rely upon the cognitive fit theory which suggests that performance of a task will be enhanced when there is a cognitive fit (match) between the information emphasized in the representation type and the information required by the task type. Since for both visual and textual methods there is a cognitive fit between task and representation, the methods exhibit similar actual efficacy. The same holds for the results on actual efficacy of domain-specific and domain-general catalogues.

To explain the results on perceived efficacy we designed our own theory of causal explanations. The theory claims that a clear process, a visual representation for risk models, a catalogue of threats and security controls, and tool support determine the perceived ease of use and perceived usefulness of a method. According to this theory, visual methods have a higher perceived efficacy because they have a clear process to identify security threats and controls and adopt a graphical representation for risk models.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

7 References

[1] Buenko, J.A. (1986): "Information Systems Methodologies - A Research View". Information Systems Design Methodologies: Improving The Practice, T.W. Olle, H.G. Sol and A.A. Verrijn-Stuart, (Eds.), North-Holland.

[2] Moody, D.L. and Shanks, G.G. (1998): "Evaluating and Improving the Quality of Entity Relationship Models: An Action Research Programme", Australian Computer Journal, November.

[3] Wynekoop, J.L. and Russo, N.L. (1997): "Studying Systems Development Methodologies: An Examination Of Research Methods", Information Systems Journal, 7, 1, January.

[4] M Jorgensen, K Molokken-Ostvold (2004), Reasons for software effort estimation error: impact of respondent role, information collection approach, and data analysis method, Software Engineering, IEEE Transactions on 30 (12), 993-1007.

[5] Gregor, S. (2006). The nature of theory in information systems. *Mis Quarterly*, 611-642.

[6] Davis, F. D. "Perceived Usefulness, Perceived Ease-of-Use, and User Acceptance of Information Technology," MIS Quarterly (13:3), 1989, pp. 319-340.

[7] Bagozzi, R. P.; Davis, F. D.; Warshaw, P. R. (1992), "Development and test of a theory of technological learning and usage.", Human Relations 45 (7): 660–686,

[8] Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. Information systems research, 11(4), 342-365.

[9] Dervin, Brenda. "The relationship of user-centered evaluation to design: addressing issues of productivity and power." *ACM SIGOIS Bulletin* 16.2 (1995): 42-46.

[10] http://www.eurocontrol.int/sites/default/files/publication/files/e-ocvm3-vol-1-022010.pdf

[11] TW Olle, HG Sol, AA Verrijn-Stuart (Eds.), Information systems design methodologies: a comparative review, North-Holland, Amsterdam (1982).

[12] Moody, Daniel L. "The method evaluation model: a theoretical model for validating information systems design methods." ECIS 2003 Proceedings (2003): 79.

[13] Madden, Thomas J., Pamela Scholder Ellen, and Icek Ajzen. "A comparison of the theory of planned behavior and the theory of reasoned action." *Personality and social psychology Bulletin* 18.1 (1992): 3-9.

[14] Rescher, Nicholas. "Methodological pragmatism: A systems-theoretic approach to the theory of knowledge." (1977).

[15] Kim, J. "Causation," in The Cambridge Dictionary of Philosophy (2nd ed.), R. Audi (ed.), Cambridge University Press, Cambridge, UK, 1999, pp. 125-127.

[16] Benjamin Fabian, Seda Gürses, Maritta Heisel, Thomas Santen and Holger Schmidt: A comparison of security requirements engineering methods. Requirements Engineering 15(1):7-40, 2010.

[17] Marianne Busch, Nora Koch and Martin Wirsing: SecEval: An Evaluation Framework for Engineering Secure Systems. In Modellierung 2014, Lecture Notes in Informatics (LNI) 225, pp.337-352, Gesellschaft für Informatik, 2014.

[18] Anita Vorster and Les Labuschagne: A Framework for Comparing Different Information Security Risk Analysis Methodologies. In Proceedings of the 2005 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries, pp. 95-103, SAICSIT, 2005.

[19] Kwo-Shing Hong, Yen-Ping Chi, Louis R. Chao and Jih-Hsing Tang: An integrated system theory of information security management. Information Management & Computer Security, 11(5):pp. 243-248, 2003.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

[20] Stefan Fenz and Andreas Ekelhart: Verification, Validation, and Evaluation in Information Security Risk Management. IEEE Security and Privacy, 9(2):58-65, 2011.

[21] Mamadou H. Diallo, Jose Romero-Mariona, Susan Elliott Sim, Thomas A. Alspaugh, and Debra J. Richardson: A Comparative Evaluation of Three Approaches to Specifying Security Requirements. In Proceedings of the 12th Working Conference on Requirements Engineering: Foundation for Software Quality (REFSQ'06), 2006.

[22] Mikko Siponen and Robert Willison: Information security management standards: Problems and solutions. Information & Management 46(5):267-270, 2009.

[23] Lund, M. S., Solhaug, B., & Stølen, K. (2010). Model-driven risk analysis: the CORAS approach. Springer Science & Business Media.

[24] EATM, ATM Security Risk Assessment Methodology, Edition 1.0, EUROCONTROL, May 2008.

[25] Vessey, I. "Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature*." Decision Sciences 22, no. 2 (1991): 219-240.

[26] Labunets, K., Massacci, F., and Paci, F. (2013, October). An experimental comparison of two risk-based security methods. In *Empirical Software Engineering and Measurement, 2013 ACM/IEEE International Symposium on* (pp. 163-172). IEEE.

[28] Labunets, K., Paci, F., Massacci, F., & Ruprai, R. (2014, August). An experiment on comparing textual vs. visual industrial methods for security risk assessment. In *Empirical Requirements Engineering (EmpiRE), 2014 IEEE Fourth International Workshop on* (pp. 28-35). IEEE.

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.

-END OF DOCUMENT-

©SESAR JOINT UNDERTAKING, 2011. Created by *[Member(s)]University of Trento, SINTEF, Deep Blue and University of Southampton* for the SESAR Joint Undertaking within the frame of the SESAR Programme co-financed by the EU and EUROCONTROL. Reprint with approval of publisher and the source properly acknowledged.