REPORT D1.3.1

Validation Strategy and Plan

PROJECT TITLE:	CONTRACT-BASED AIR TRANSPORTATION SYSTEM			
PROJECT ACRONYM:	CATS			
CONTRACT NUMBER:	TREN/07/FP6AE/S07.75348/036889			
PROJECT START DATE:	01.11.2007			
DURATION:	36 MONTHS			
PROJECT COORDINATOR:	FREQUENTIS AG	(1)	(FRQ)	AT
PRINCIPAL CONTRACTORS:	THE EUROPEAN ORGANISATION FOR THE SAFETY OF AIR NAVIGATION	(2)	(EEC)	BE
	AIR FRANCE CONSULTING	(3)	(AFC)	FR
	UNIQUE (FLUGHAFEN ZÜRICH AG)	(4)	(Unique)	CH
	UNIVERSITY OF LEIDEN, INTERNATIONAL INSTITUTE OF AIR AND SPACE LAW	(5)	(IIASL)	NL
	EIDGENOESSISCHE TECHNISCHE HOCHSCHULE ZÜRICH	(6)	(ETH)	CH
	LABORATORIO DI RICERCA OPERATIVA – DIPARTIMENTO DI ELETTROTECNICA ED INFORMATICA UNIVERSITÀ DEGLI STUDI DI TRIESTE	(7)	(ORTS)	IT
	ENAV SPA – SOCIETÀ NAZIONALE PER L'ASSISTENZA AL VOLO	(8)	(ENAV)	IT
	SKYSOFT-ATM S.A.	(9)	(SkySoft)	CH

DOCUMENT IDENTIFIER:	D1.3.1
ISSUE:	5.0
ISSUE DATE:	15/10/2010
PREPARED:	EUROCONTROL
APPROVED	EUROCONTROL
RELEASED	FREQUENTIS
DISSEMINATION STATUS:	PUBLIC
DOCUMENT REF:	CIFE01EN40001.21

Project co-funded by the European Community within the 6th Framework Programme.

History Chart

Issue	Date	Changed Page (s)	Cause of Change	Implemented by
0.1	10/12/07	All sections	New document	S. Guibert
1.0	02/03/2008	All sections	Review for final delivery	S. Guibert
2.0	02/06/2008	Sections 7 & 8	Update regarding HIL1	S. Guibert
2.0.2	11/08/2008	Section 9	HIL1 Experimental Plan	S. Guibert
2.0.3	18/08/2008	Section 12	Analysis Plan	S. Guibert
2.0.4	14/10/2008	Sections 7 & 8.2	IIASL inputs	S. Guibert
2.1	15/10/2008	All sections	Review for submission	S. Guibert
2.2	24/02/2009	§5.3, §7.2 & §10.1	ORTS inputs	S. Guibert
2.3	10/03/2009	§1.1; 3.1; 3.6 & 7.2.3	IIASL inputs	S. Guibert
2.4	19/08/2009	§9.2 & §13	Integration of HIL2 Validation Plan	S. Guibert
2.5	28/08/2009	§9.2	Integration of partners' inputs	S.Guibert
3.0	15/12/2009	§9.2.5	Explanation of SPT and SESAR follow-up	S. Guibert
3.1	19/02/2010	§9.3	HIL3 update	S.Guibert
3.2	24/02/2010	§11	Conclusion	S.Guibert
3.3	18/05/2010	All sections	Integration of partners' inputs	S. Guibert
4.0	04/06/2010	§9.3.3 & 9.3.4	Integration of HIL3 Validation Plan	S. Guibert
4.1	25/06/2010	All	Integration of partners comments and proof of reading	S. Guibert
4.2	29/06/2010	§5.3.2	Integration of ORTS comments	S. Guibert
5.0	15/09/2010	§1, 4, 7, 8, 9 & 10	Review after EC comments	S. Guibert L. Castelli

Authorisation

No.	Action	Name	Signature	Date
1	Prepared	S. Guibert /EEC		15/09/2010
2	Approved	L. Guichard/EEC		15/09/2010
3	Released	C. Rihacek / FRQ		

The information in this document is subject to change without notice.

Distribution List

This document is distributed as indicated below.

Additional copies held by unnamed recipients will not be updated.

Paper Copy Number	Name	Address
1	Katarzyna Gryc	EC, Brussels
2	Library	Frequentis, Vienna

Electronic Copy Number	Name	Address
1	Katarzyna Gryc	EC, Brussels
2-10	CATS consortium	

Contents

Illustrations

Tables

Executive Summary

This document defines the strategy and the plan followed to conduct investigations in order to validate the CATS hypotheses and to provide strong arguments and carry out a performance assessment, using a range of experimental techniques (modelling, simulation, etc.).

It is based on a rigorous process designed to establish a validation framework, and is based on the E-OCVM methodology which provides guidelines in order to structure the design and validation process in line with a precise lifecycle view.

As required by the SESAR Programme, performance assessment experiments will mainly address the stakeholders' needs and expectations, through key performance areas (KPAs) and the associated KPIs (key performance indicators). The aim is to assess whether the benefits are being delivered as proposed.

The second goal of this validation relates to how the concept performs regarding operators (e.g. working methods, activity, performance, etc.). This is also a key aspect of the validation; the aim is to ascertain whether the proposed concept is acceptable from an operator's perspective.

The CATS Project evaluates the Contract of Objectives concept through three human-centred experiments, ensuring that all the operational issues of the Contract of Objectives are addressed. At the same time, a safety and risk assessment and an analysis of the benefits and legal implications will be carried out.

1 Introduction

1.1 CATS Project

The CATS Project was selected during the 4th call of the European Commission's FP6. It was launched on 1 November 2007 for a period of three years. The CATS Project proposes an integrated decoupled ATM organisation, where all the actors negotiate and agree on their own objectives. This organisation is based on a breakdown of responsibilities between the various actors. The transfer of responsibilities between them is now explicit and formally contracted. This will reinforce arrival punctuality.

The new ATM paradigm proposed by CATS is based on the SESAR concept [6] and proposes one of the possible implementations of the Business Trajectory, namely the Contract of Objectives (CoO). This introduces an innovative way of managing ATM by mutually-agreed objectives leading to a market-driven ATS.

It addresses the entire air transport supply chain by reconciling operational links between air and ground services and thus enhances efficiency by increasing the predictability of ATS. Objective assignment and negotiation are carried out on the basis of the collaborative decision-making (CDM) process with a view to establishing operational agreement (the right balance between productivity and safety).

Guaranteed results are offered to the airline by the air traffic system to ensure punctuality at destination.

The trade-off process, driven by actors' interests and constraints, will reinforce autonomy and responsibility, while allowing a high level of integration between all stakeholders and the consistency of the ATM organisation, in order to optimise the use of scarce resources.

The main concerns are operational improvements and airspace users' benefits. The aim of the concept of operations is to improve system efficiency by means of enhanced collaboration focused on the same goal, namely meeting schedules.

The objectives presented below reflect the focus of the proposed project:

- Link ATM actors together through agreed objectives and interfaces: Reconciliation of air navigation service providers (ANSPs), airlines and airports by ensuring mutual awareness of the constraints (i.e. target window) and thus allowing the ultimate target (i.e. punctuality at destination) to be focused on. If these constraints are shared, everyone is aware of the possible scope for adaptation in relation to the current flight, thus allowing an efficient collaborative decision-making process.

- Integrate flexibility to cope with uncertainties: Target window (TW) modelling includes both technical-level constraints (e.g. aircraft performance, en-route control limitations) and scope for sufficient flexibility when disruptions occur. It gives air traffic controllers (ATCOs) and aircrews the managerial scope to deal with uncertainties arising while a flight is airborne, whilst still respecting the original schedule.

- Coordination of actors' resources in order to deliver the best service: The CoO is the fundamental unit of the collaborative planning process, and is drawn up, agreed and shared by all the actors. This contract represents a "guarantee of service results" from the actors involved in the flight. Each actor is thus able to provide the relevant resources and infrastructure to offer the appropriate service, taking into account the applicable and mandatory legal regime.

- Enhanced collaboration through the Single European Sky: The CoO and TWs represent a commitment on agreed interfaces between actors. The network is highly sensitive to amendments (the butterfly effect), so in order to be efficient it needs to be considered at continental level (i.e. on a European scale).

The main aim of the CATS Project is to assess the operational validity of the Contract of Objectives proposal through experiments.

This will be achieved by:

- defining the interfaces between the various actors and operators:
  - ATCO and ATCO in the airborne cycle (between two ANSPs)
  - ATCO and aircrew in the airborne cycle
  - Airport, airline and ANSP during the renegotiation process
- assessing, through human-in-the-loop experiments, operational acceptability and technical feasibility, particularly in the case of renegotiation
- evaluating the impact of this contract in terms of human factor issues, safety and operator workload
- evaluating the benefits (cost business analysis, safety & risk analysis) and the legal implications

1.2 Document Structure

This document represents the CATS Project delivery D1.3.1 forming part of WP1 as defined in the DoW [2]. This document is consistent with the E-OCVM methodology and presents the strategy for conducting investigations in order to validate the hypotheses. Starting with the State of the Art document [9] and the operational concept definition [8], the aim of this document is to clearly:

- define the high-level objectives addressed by the three experiments
- define the scope and content of each of them
- describe the expected outcome, in terms of the various themes to be assessed
- identify indicators and metrics and define the global validation strategy

This document is divided into 12 main chapters:

- Chapter 1 introduces the document
- Chapter 2 describes the problems relating to en-route traffic execution
- Chapter 3 provides an overview of the CATS concept and identifies its expected benefits
- Chapter 4 quantifies the key performance framework and SESAR targets
- Chapter 5 presents the Validation Methodology
- Chapter 6 identifies the CATS Validation Expectations
- Chapter 7 defines the CATS Validation Objectives
- Chapter 8 sets the CATS Validation Strategy

- Chapter 9 describes the Operational Validation Plan
- Chapter 10 defines the Systemic Validation Plan
- Chapter 11 concludes the document
- Chapter 12presents the references used in the document
- Chapters 13 & 14 introduce the operational Data Analysis Plan in Annex A & B.

1.3 Assumptions

The CATS Project is intended to validate the impact of implementing the CoO. This validation will demonstrate the operational acceptability of TW management and also identify the benefits and limitations of the CoO implementation. A pre-requirement for CATS operational assessments is that the CoO has already been created and validated. One of the assumptions is that all the stakeholders will play the game during the planning phases in order to build a stable RBT and CoO. Hence, assuming that the CoO is agreed, the three human-in-the-loop (HIL) experiments will be carried out to assess the impact of the proposed concept on operators, taking account of both system and human performance. The systemic assessment will approach the analysis from a more global viewpoint, including the planning and execution phase.

1.4 Document evolution

These activities will be carried out in the framework of an iterative approach focusing first on issues specific to controllers, then on issues related to the interactions between controllers and aircrews, and finally on broader issues involving airports, airlines and ANSPs in the renegotiation of the Contract. The Concept of Operations and then the Validation Strategy and Plan will be revised in the light of the results and conclusions of the previous iteration.

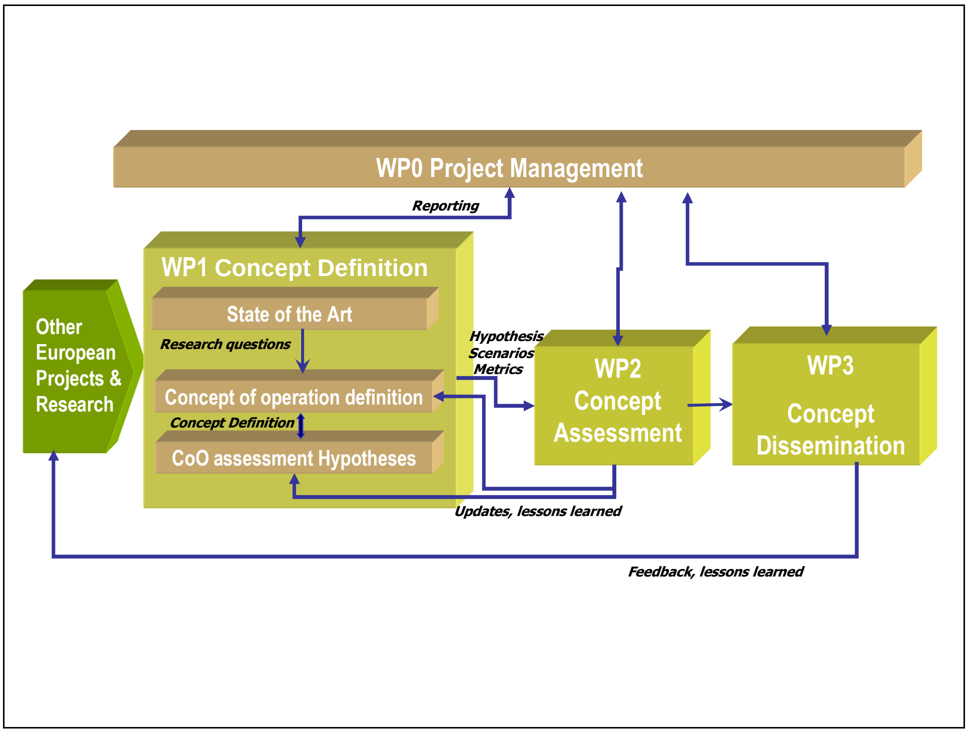

Figure 1: Concept of Operations iterative improvement process

1.5 Definition, abbreviations and acronyms

ACC	Area Control Centre
ANSP	Air Navigation Service Provider
AO	Airline Operation
AOs	Aircraft Operators
APT	Airport
ATC	Air Traffic Control
ATCO	Air Traffic Controller
ATM	Air Traffic Management
BEC	Behavioural and event checklist
CAP	Capacity
CLM	Concept Lifecycle Model
CONOPS	Concept of Operations
CoO	Contract of Objectives
CWP	Controller Working Position
DOW	Description of Work
ECAC	European Civil Aviation Conference

EFF	Efficiency
E-OCVM	European Operational Concept Validation Methodology
EP3	Episode 3 project
EXE	Executive Controller
FLEX	Flexibility
FMS	Flight Management System
FTS	Fast-Time Simulation
HF	Human Factors
HFE	Human Factors Expert
HIL	Human-In-the-Loop
HMI	Human-Machine Interface
ISA	Instantaneous Self-Assessment
KPA	Key Performance Area
KPI	Key Performance Indicator
LTF	Long Term Forecast
MTCD	Medium-Term Conflict Detection
NASA-TLX	NASA Task Load
NM	Network Manager
NOP	Network Operational Plan
OTS	Over the shoulder
PLN	Planning Controller (sometimes call also planner)
PI	Performance Indicator
PRED	Predictability
RBT	The reference business trajectory, which the airspace user agrees to fly and the ANSP agrees to facilitate. Most times indicated in the RBT are estimates, some may be target times of arrival (TTA) to facilitate planning and some may be constraints (Controlled Time of Arrival - CTA). The RBT consists of a 2D route, altitude and time constraints when required, altitude, time and speed estimates at way points and trajectory change points.
RTS	Real-Time Simulation
SAF	Safety
SASHA	Shape Questionnaire
SESAR	Single European Sky Applied Research
SME	Subject Matter Expert
STCA	Short-Term Conflict Detection
SWIM	System-Wide Information Management
SWOT	Strengths, Weaknesses, Opportunities and Threats

TC	Tactical Controller
TTA	Target Time of Arrival
TWs	Target Windows
SA	Situation Awareness
UDPP	User Driven Prioritisation Process
XFL	Exit Flight Level

2 Problem description

SESAR D1 identified the limitations of the current situation [4]. The major limitations can be summarised as follows:

- The future European ATM system must be able to cope with the expected growth in traffic.
- European airspace is, in the main, organised around the use of fixed volumes and rigid route structures which are managed in a fragmented manner. This results in aircraft being unable to fly their most efficient trajectory and creates unnecessary additional workload for air traffic control.
- Most aircraft operating today have the capability to fly with much greater precision in terms of position and time than is accommodated in the design, and supported by many of the systems in operational service for managing and controlling air traffic. This capability is currently not exploited.
- Throughout the processes, procedures and systems used by the stakeholders currently involved in planning, managing and executing flights, decisions are often taken in isolation by somebody on matters which have an impact on others. This leads to fragmentation and inefficient flight profiles. However, the stakeholders have a large amount of information which is currently not fully exploited.
- Pilots currently have a limited situational awareness of traffic, which can potentially affect them and restrict them from taking a more pro-active role in the ATM process.
- The design and management of the future European ATM system should be based on:
  - a common framework which links the economic, commercial and operational values of its stakeholders and the establishment of explicit service relationships between them
  - a comprehensive performance framework applied across the system as a whole
  - the use of the notion of a "network plan"

The European ATM network of the future will be structured around the use of a performance-based, service-oriented operational concept. In support of the network there will be a single European ATM system, which will provide a variety of ATM services to all types of airspace users to meet their respective needs.

On the basis of the SESAR analysis and the work performed in D1.2.1 [8], the CATS Project identifies and deals with three main problems:

- Punctuality is needed for passengers, airlines, and airports, as regards the turn-around of aircraft.
- The lack of predictability in traffic flows. ATCOs need to visualise the final flight objectives reported on their CWP, to consider downstream actions. This will be more efficient for the overall system [13].
- The improvement of predictability will allow optimisation of the scarce resources of all ATM actors. This will be more cost-effective, and will also have a good impact on safety, as a result of the decrease in traffic bunching.

Efficiency and punctuality are thus key elements on which the future ATM system should be built.

The proposed CATS concept could be seen as a means of increasing the predictability of aircraft arrival times while ensuring the organisation of actors' resources.

3 CATS concept description

3.1 Proposed solution

A new way of enhancing traffic efficiency is to improve functional and operational continuity in aircraft management, both on the ground and in the air, with a view to meeting safety and productivity objectives. Functional continuity has an airspace dimension (heterogeneity of European airspace) and a time dimension (from planning to execution). To this end, it is proposed to bring together all air navigation components by means of a Contract of Objectives.

The Contract of Objectives will provide well-defined objectives for each actor involved in a flight (air traffic controllers, aircrews, airports, airlines, air navigation service providers). These objectives represent the commitment from each actor to deliver a particular aircraft inside temporal and spatial intervals. Agreement on the Contract of Objectives is reached through a collaborative decision-making (CDM) process, allowing the right balance between productivity and safety.

Instead of all actors optimising their own processes locally according to their own constraints and business objectives, CATS promotes a comprehensive and holistic view based on highly collaborative and system-wide processes leading to a different and better distribution of constraints across the whole ATS. Air traffic controllers and all ATM operators should be informed about the impact of their actions on the overall performance of the flight. A clear transfer of responsibility between actors will ensure the full achievement of the planned business trajectories.

The Contract of Objectives, and associated Target Windows, could be seen as organisational elements constraining the business trajectory, so as to maintain efficiency and schedule integrity, while contributing to environmental standards. It will allow partners to better organise their own scarce resources, and then this optimisation of resources will positively impact the efficiency of the entire air-transport supply chain.

The Contract of Objectives is a way of improving ATM efficiency by providing a better organisation of the actors' resources and predictable traffic flows.

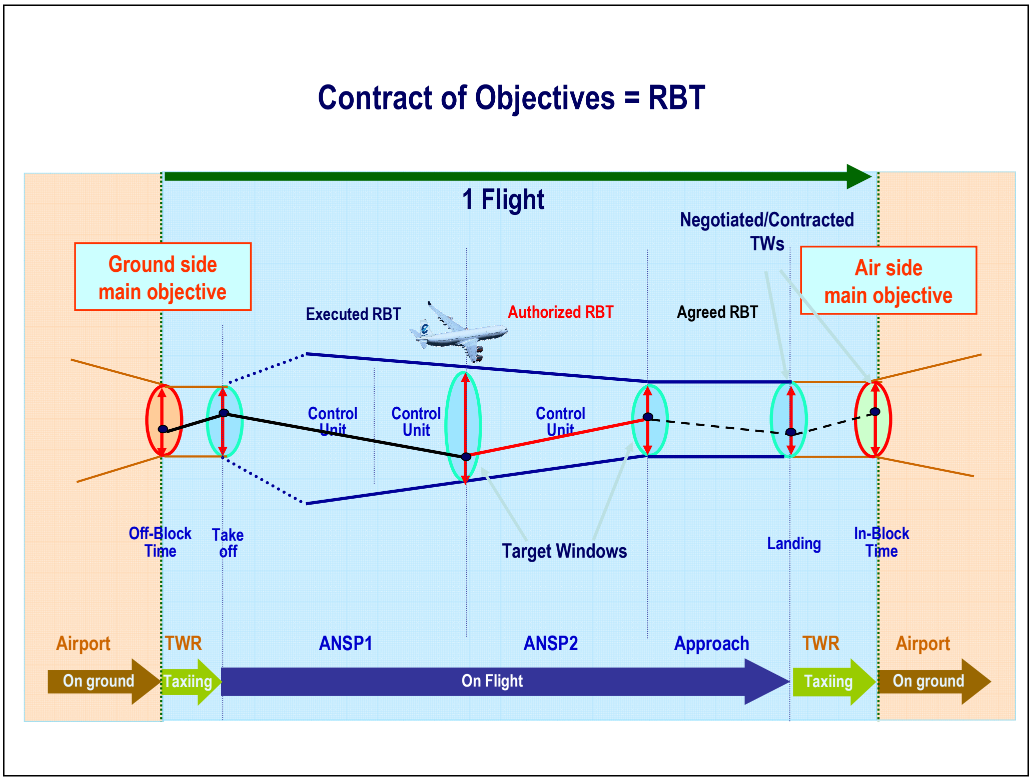

Figure 2: Contract of Objectives

3.2 Operational description

The objectives set by the CoO are negotiated through a CDM process supported by SWIM, in terms of punctuality at destination, taking into account all actors' constraints. This negotiation process could be described as follows:

- Long-term planning phase: development of an initial and not too detailed schedule, made up of TWs at departure and arrival airports, taking into account infrastructure and environmental constraints.
- Medium-term planning phase: development of business trajectories and negotiation of TWs through an iterative process. Integration of weather prediction.
- Short-term planning phase: Continuous refinement of TWs until signature of the CoO.

Then the Execution phase of the flight can start. The Contract of Objectives gives the controller and aircrew the means to manage the imprecision inherent in air traffic in accordance with their own objectives. The crews' objectives, therefore, are to adhere to an arrival schedule; controllers, on the other hand, must ensure aircraft safety while keeping aircraft within the envelope defined in the contract, which guarantees that the contract's objective will be respected.

If for any reason (weather, conflict, etc.) a TW cannot be fulfilled, a renegotiation process will start between the impacted actors, resulting in a new CoO.

3.2.1 Principles

The Contract of Objectives is defined by a collection of target windows, located at the areas of transfer between each actor. Thus, the TW represents the objectives assigned to the flight through responsibility transfer areas between control units (lateral or vertical boundary of responsibility airspace). These objectives can be expressed as a temporal window associated with a subset of the responsibility transfer area. In other words, the target window is the objective to be achieved at the end of the responsibility area, in order to meet the ultimate objective: punctuality at destination.

In this task, the ATCO has to manage the traffic in the best way for an aircraft to comply with the Target Windows. This task requires the ATCO to have the right information to accurately manage the aircraft in relation to the Target Windows, knowing where aircraft are planned to cross the Target Window, and the rooms for manoeuvre to manage achievement of the Target Window. The management of the Contract of Objectives at the Controller Working Position requires specific information to be provided to the ATCOs. The objective is the same for the aircrew.

In the CATS philosophy, target window management is completely integrated into the tasks of the pilots and ATCOs. Information relating to target window management has to be perceived and processed as soon as possible and directly it is available. The aircraft position in the target windows needs to be easy to perceive in order to identify what action needs to be taken or whether no action is required. So, this information needs to be displayed on the tools available (e.g. aircraft label on the radar screen) and has to be informative about the aircraft’s intended position. Three levels of aircraft state, regarding the TW, are relevant:

- Nominal state, where the aircraft is in the “core” of the target window
- Edge state, where the aircraft is inside the target window, but close to the target window boundaries, with a tendency to leave the window
- Outside state, where the aircraft is outside the target window, and so a renegotiation process has to be initiated

The status (In, Out or Edge) is shown by a colour change at the exit point in the label (or XFL to be checked) in order to keep the overall coherence of the platform, and the trend, extracted from the change in past positions, is shown by a + (diverging) or – (converging) on the right of the XFL. This information will always be displayed on the screen. More detailed information on the HMI proposed to display the 4D information for the various HIL experiments will be described in Deliverable D2.1.4.1. This additional information specifying exact deviations from TWs would be useful in allowing ATCOs and pilots to take the appropriate actions (e.g. conflict resolution, maintain the RBT).

3.3 Assumptions

The CATS concept of operations [8] covers the overall scope from the ATM planning phases to the post-flight analysis. Nevertheless, the CATS operational assessments will mainly focus on the impact of the CoO and TWs through the trajectory-based operations.

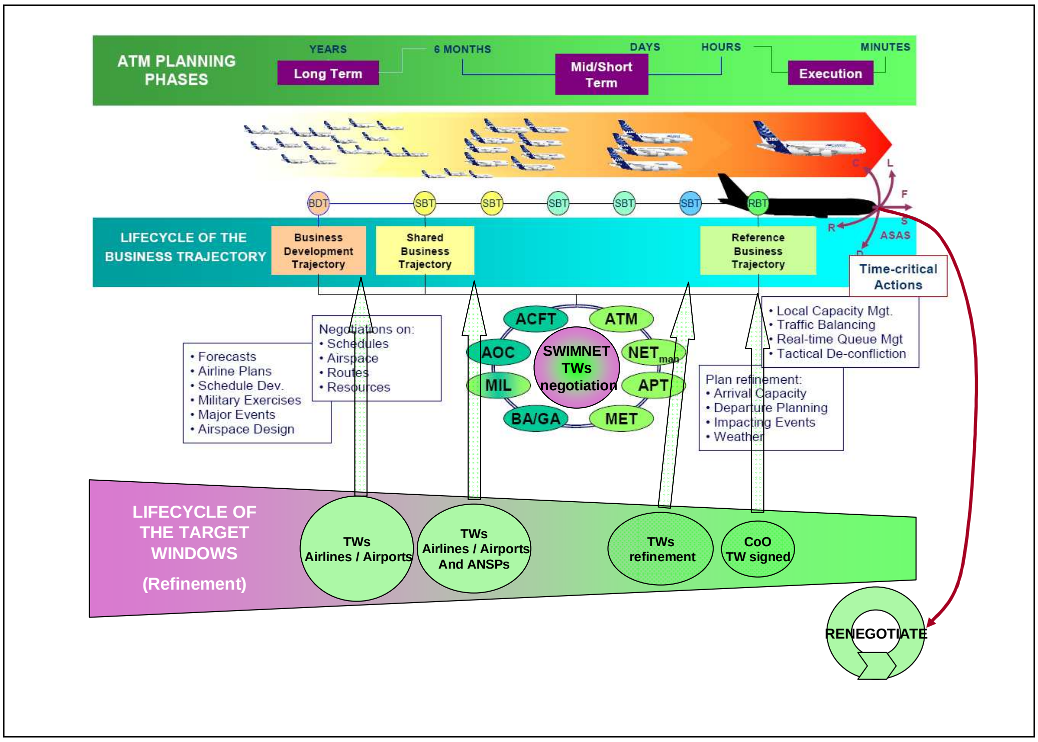

Figure 3: Target Windows Lifecycle

As described in the DOW [2], the CATS project goal is to assess the acceptability of the concept to operators, while focusing on the en-route execution phase, and also to demonstrate the benefits from a more global viewpoint (systemic assessment).

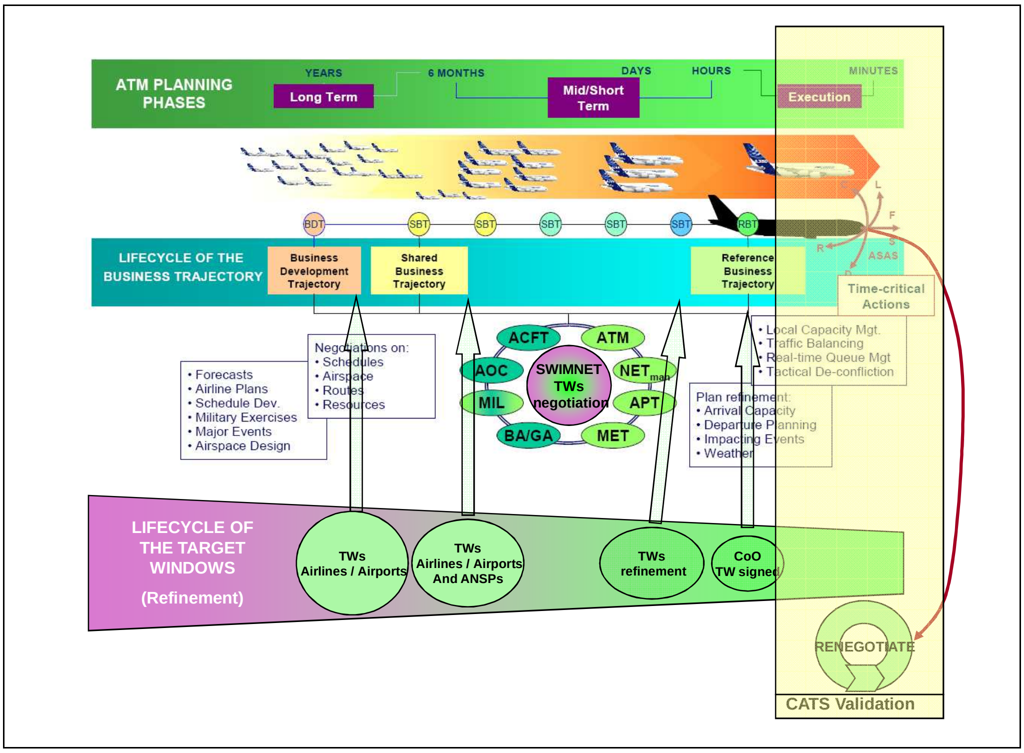

Figure 4: CATS validation scope

The main assumption of the operational assessments is therefore that the CoO and associated TWs have been negotiated and shared by all the actors involved. Another assumption made is that the air traffic controller is responsible for achieving the TW, even if the pilot can also monitor it and share with the controller the adherence of the flight to the TWs.

3.4 Technical requirement

As presented in the CATS Concept of Operations [8], the main enablers supporting the CoO in the SESAR environment are:

- Collaborative Decision-Making infrastructure supported by SWIM, allowing data sharing between airspace users/service providers (i.e. ATM systems, FMS, airport);
- Equipped aircraft with:
  - Data-link applications (i.e. ADS-C, CPDLC, ADS-B/out and ADS-B/in);
  - Target Window monitoring and renegotiation tools, including a "what-if" HMI.
- Ground systems providing:
  - Data-link applications supported by SWIM:
    - 4D trajectory uplink;
    - Clearance to proceed.
  - Enhanced monitoring and resolution tools to the controller:

- Clearance tools;
- Trajectory prediction;
- Medium-Term Conflict Detection (MTCD);
- Target Window monitoring and renegotiation tools.

3.5 Concept story board classification

Regarding the SESAR concept story board [24], the CATS concept will be placed in Step 2 (Trajectory-Based Operations) as explained in D1.2.2 Table 3[25]. It focuses on the business trajectory management.

3.6 Expected benefits

The successful implementation of the operational concept through the CoO and the TWs may produce positive effects for the various stakeholders, and for the European ATM system as a whole. From a theoretical point of view, when various actors operating on the same system adopt a cooperative attitude, this entire system is likely to achieve better performance than when the actors do not cooperate. In the latter case, advantages usually accrue to the actor which has more information or decides first and imposes its choices on the others. Instead, the CATS Project assumes that TWs have already been negotiated before the aircraft takes off, and could be renegotiated during the flight execution phase. Hence the evolution from the Initial Shared Business Trajectory to the Reference Business Trajectory and possible deviations from it during the flight are achieved through a collaborative decision-making process. The expected benefits for the ATM system and each individual stakeholder (i.e. ANSP, airline or airport) will be assessed mainly with regard to SESAR operational key performance areas, namely cost-effectiveness, capacity, (flight) efficiency, flexibility and predictability (see Section 4.2 for further details).

The achievement of a CDM process requires that all players are ready to share information and accept possible trade-offs in their performance. In this context, the CoO may play a major role in enforcing or encouraging the actors to participate in the negotiation and renegotiation phases by defining incentive schemes or performance-based rewards and/or penalties.

3.7 Legal implications

As the CoO is closely related to the responsibility of the actors for providing a service and adhering to the agreed objectives, given tangible form in the TWs, the legal framework for the various actors needs to be carefully evaluated. This will be the main goal of the specific study carried out in WP2.2.3.

One of the aims of the CATS Project, through WP2.2.3, is to clearly define the responsibilities and liabilities of each actor within the CoO.

The relationship between ANSPs and airlines is governed by and subject to a number of public laws, especially international public laws, instruments such as the Chicago Convention on Civil Aviation, standards developed by the International Civil Aviation Organization (ICAO), bilateral air services agreements, European Community law and international agreements concerning EUROCONTROL.

Bilateral air services agreements are concluded by States in order to regulate air services between, and sometimes beyond, the two States concerned. These agreements are especially relevant for non-European actors, in particular non-European airlines, as they fly into, from and via European airspace under mandatory rules laid down in these bilateral agreements.

The Warsaw and Montreal Conventions make air carriers liable for delays in the carriage of passengers. This falls under the framework of private air law issues. The Montreal Convention of 1999 is becoming the more interesting agreement as it is increasingly adhered to by European States and now also non-European States.

The relationship between ATM actors is governed by mandatory public law. Contractual agreements are rare. Airport-airline relations may have a contractual nature, but this relates to quality only and not generally to liability.

The only example of an airline-ATM contractual relationship can be found in New Zealand, where the local ANSP contracts with airlines by way of a standard form of contract. According to the terms of the contract, the ANSP retains the right to refuse to provide services to any carrier which refuses to abide by the terms or is in breach of them. The terms exclude the liability of the ANSP for delays. The major flaw of this system is that no law obliges the airlines to enter into a contract with the ANSP.

GALILEO (the European Satellite Positioning and Air Navigation System model) proposes that private-law contracts should be concluded between all actors involved and that public-law agreements should be concluded by all States concerned. These public-law contracts would ensure that the private-law contracts are harmonised and contain certain necessary provisions.

In any case, it is important to emphasise that the CoO concept cannot function properly without an enforceable public-law provision which allows airlines, airports and the ANSPs concerned to enter into a contractual relationship.

It can be concluded that, from a legal point of view, the CoO constitutes a real paradigm shift. The fields of application and related substantive provisions still have to be determined. The substantive provisions must be subject to the public and private laws in force, or certain provisions of these laws should be amended. This will be the goal of D2.2.3.

4 Stakeholder requirements & performance framework

4.1 Stakeholder objectives

As described in E-OCVM [10], following the description of the ATM problem and the proposed solution, the validation strategy can start to be defined. The strategy is based on the stakeholders' objectives and required performance.

The main stakeholders of the CATS Project can be split into two different groups:

- the first has a major impact on the CoO's implementation (ANSP, ATCO, aircraft and airport operator and aircrew)
- the second is involved in the decisions to adopt changes (regulatory authorities and passengers).

Stakeholder	Concerns
ANSP	Optimise their resources (manpower in line with level of traffic and consistent airspace design) to reduce costs
ATCO	Flexibility to manage traffic and solve potential conflicts, better predictability of traffic
Aircraft operator	Optimise aircraft turn-around, trust in the scheduling
Aircrew	Meet the schedule in an optimal way (fuel consumption, etc.)
Airport operator	Optimise ground operations and runway use
End-user customers (passengers, freight)	Need to arrive on time at destination, need to have the best choice of destinations and schedules, benefits from cost reductions
Regulatory authorities	Will consider the impact of CoO implementation on the rules and legal framework and compatibility with such rules

Table 1: Stakeholder concerns

The first group of stakeholders will be the direct user of the CoO's implementation. ANSPs, aircraft and airport operators will be directly involved in negotiation of the TWs and CoO, while ATCOs and aircrews will operate the aircraft safely in accordance with the CoO.

The regulatory authorities will provide the regulatory framework for the CoO's implementation, as a result of the recommendations made in WP2.2.3. And finally, the passengers will be the direct beneficiaries of the CoO's implementation.

4.2 Performance Framework

The performance framework is intended to structure the performance problems in terms of expectations, objectives, targets and indicators, the latter being expressed as quantifiable values. The CATS Project is fully aligned with the SESAR Performance Framework described in SESAR Deliverable 2 [5].

This Section describes the achievable SESAR expectations with regard to KPAs, as well as their definitions of applicability, and the CATS contribution.

KPA	Definition	Strategic Design Target	CATS Contribution
Capacity	Ability of the system to cope with air traffic demand (airspace, airport, network)	3 times more traffic in 2020, 50 000 flights / day in 2022	CoOs are finalised at a time close to the actual execution of flights. They take into account current traffic patterns and capacity distribution. Final delivery of CoOs close to the flight execution reduces flow restrictions in order to ensure better capacity. TWs are the result of the refinement process, which means that all known disruptions are taken into account. They thus promote maximisation of throughput in the ATS as a whole (including both airports and airspace). Depending on the regions concerned, TWs will ensure an optimal trade-off of constraints. Low-density traffic areas are included in the process and so will be able to give an immediate and appropriate answer linked to evolution in their situation.
Cost-Effectiveness	Cost of gate-to-gate air traffic management services in relation to the use that is made of those services	From 800/flight in 2004 to 450/flight in 2020 (in 2005 €)	CoOs are the result of a collaborative process involving all the actors. Every actor’s constraints and interests are thus part of the final trade-off. CoOs represent an improvement in airport operations (i.e. turn-around In Block Time and Off Block Time are truly predictable) leading to cost benefits for both airlines and airports (e.g. through reduced delays and optimised use of gates). The TWs are defined taking into account the various actors’ inherent “parameters” (e.g. preferred/economically desired routes, staff optimisation, avionics)
Efficiency	Measurement of the actual business trajectory against the shared business trajectory	Average departure/arrival delay of flight will not exceed 10mn Flight duration extension will not exceed 10 mn, less than	CoOs are the final output of the refinement process. This process takes into account every actor’s requirements in relation to their operational effort and leads to the “best” trade-off for everyone.

KPA	Definition	Strategic Design Target	CATS Contribution
		2.5% of flights suffering additional fuel consumption, average additional fuel consumption less than 5%	An initial constraint in the CoO definition is arrival punctuality, to allow delays to be minimised and to ensure on-time operations, which are especially critical for turn-around flights. The number of broken CoOs is an indicator of the quality and relevance of the initial mechanism. Corrective actions (e.g. parameter tuning, actor constraints, re-evaluations, etc.) will allow improvements to be made to overall efficiency. TWs represent potentially "acceptable" alternatives defined in terms of a "4D volume". Operations can thus be adapted in line with real-life events in order to obtain the agreed preferred trajectory.
Flexibility	Ability of the system to respond to sudden changes in demand or capacity	98% of non-scheduled flights accommodated with a delay penalty of less than 3mn Change request on scheduled flights: no more than 2% suffering from a delay, delay less than 5mn	CoOs represent agreed links between each of the actors involved. If a CoO is broken for whatever reason (e.g. airport closure, serious meteorological conditions, etc.), the renegotiation mechanism is triggered. This process operates at global level to ensure better solutions at network level. The network and schedule can then be quickly adapted to the new traffic pattern. As stated above, TWs are designed to provide scope for action and allow flights to be quickly adapted to new traffic patterns.
Predictability	Ensure a reliable and consistent level of 4D trajectory performance	Less than 2% of flights suffering departure delay of more than 3 mn, less than 5% of flights suffering arrival delay of more than 3 mn, less than 5% of flights suffering additional flight duration of more than 3mn, variability of flight duration around coefficient of 0.015	CoOs are major enablers for predictability, as all the involved actors jointly agree to comply with them and to "deliver" flights in accordance with them. TWs allow each actor to verify its own operations in relation to the agreed contract and thus ensure global predictability as soon as it is able to "deliver" the flight in accordance with them. Early detection of possible non-compliance with the agreement allows the overall schedule to be quickly

KPA	Definition	Strategic Design Target	CATS Contribution
			reorganised (the renegotiation phase) and thus guarantees the predictability level of the future adapted schedule.
Safety	Occurrence of accidents involving IFR traffic with an ATM contribution	1 per 100 million flight-hours	CoOs represent agreements between the actors which must be fulfilled. In the event of shortcomings during operations, a "report" is made and corrective action can be taken to ensure better and safer operations. The collaborative aspect of CoOs should permit the actors involved to share best practices in terms of safety. TWs have built-in disruption margins. It is thus possible to avoid dangerous conditions and thus minimise or prevent any impact on schedules.
Security	Attempts to interfere with aircraft trajectories, number of times unauthorized access to ATM information occurs	Zero attempts to interfere with aircraft trajectories, policy-driven access to ATM information in place in 2020	Not addressed.
Environment al Sustainability	Gaseous emissions, noise generated, operational processes to improve ATM environmental performance, degree to which traffic constraints imposed are respected	Noise emissions minimised, atmospheric emissions reduced, local traffic constraints fully respected.	Environmental constraints are introduced during the collaborative refinement process involving all the actors. CoOs are the final output of this process, and will comply with the associated environmental constraints (i.e. airport environmental constraints).
Access and Equity	Equitable treatment of airspace users	No quantified design targets	Not addressed
Participation by the ATM community	Level and methods of involvement of ATM community (informing, obtaining feedback, collaborative decision making, etc.)	No quantified design targets	Subject to provisions of mandatory law (international, European, national), CoOs are the final result in a highly transparent and collaborative refinement process. They require the participation of all the actors involved.

KPA	Definition	Strategic Design Target	CATS Contribution
			The final agreement which a CoO represents is also to the result of a participative trade-off between those actors.
Global Interoperabili ty	Extent to which global standards and uniform principles are applied	No quantified design targets	CoOs are the agreed and contractual interface between various actors. These elements are the outcome of negotiation. Even if the infrastructure and working methods are not the same for all the actors, the CoO guarantees interoperability between the actors involved, thus ensuring operational continuity. TWs represent negotiated and agreed transfers of responsibility between actors, irrespective of working methods.

Table 2: SESAR KPA

Central to the CATS assessment is the management of the RBT with the CoO. The validation of the concept does not attempt to answer the question whether or not all the targets set by SESAR have been met. The objective is to focus on the proposed improvement and to evaluate its effect regarding these targets. In other words, the objective of the CATS assessments is to assess the extent to which the introduction of the CoO and TWs can contribute to potential improvements.

If we refer to SESAR D4 [7], CATS assessments are mainly related to the Implementation Package 2 (IP2) - Accelerating ATM to implement the 2020 target concept. The objectives and targets are defined as follows:

- Cost-effectiveness: Reduce average European direct ATM gate-to-gate cost to 450€ per flight (up to 600€ in high-complexity airspace).
- Capacity: Ensure, by 2020, a European ATM network capacity increase of 73%, based on the LTF 2004 high scenario, while improving the QoS (efficiency, predictability and flexibility).
- Quality of Service
  - Efficiency
    - Fuel efficiency: eliminate the current inefficiencies as far as possible, taking into account the trade-off with capacity and other QoS indicators;
    - Time efficiency: ensure on-time operations, seeking to meet the D2 targets [5], taking into account the trade-off with capacity and other QoS indicators;
  - Predictability
    - Ensure on-time arrival
    - Reduce variability of operation times
    - Reduce the impact of disruption to the ATM network
- Safety level has to be kept commensurate with the improvement made in operational performance, to ensure that no degradation occurs.

5 Validation Methodology

This CATS Validation Plan defines a strategy and plan to conduct modelling and simulation activities in order to achieve assessment results regarding the performance of the SESAR concept and to develop arguments for the stakeholders and decision-makers.

The Validation Plan will be used to define and support the experiments to be conducted in WP2. The experiments and experimental conditions are determined by the KPA, presented in Chapter 4, and are in line with the E-OCVM methodology.

CATS uses the E-OCVM approach for validation [10]. The concept proposes several changes to the current ATM picture: various technological changes, changes in responsibilities and in working methods. The validation activity should cover all these aspects and relate them to the project goals.

As described in the DOW [2], the CoO and associated TW concepts will be assessed through two different approaches:

- Operational assessment, performed by way of three "case-focused" HIL experiments, devoted to the operational acceptability of the proposed concept. Particular attention will be paid to human factors aspects, since the role of ATCOs and pilots will be modified;
- Systemic assessment, based on a safety-case study and risk analysis, a legal study and also a cost-benefit assessment.

The CoO concept needs to be designed and assessed by means of a global approach involving all the ATM actors. We propose to evaluate this concept by means of a three-way approach, thus ensuring that all the CoO issues are addressed:

- evaluation of the impact of the CoO between ANSPs
- evaluation of the impact of the CoO between ATCOs and aircrews
- evaluation of renegotiation between airports, airlines and ANSPs

5.1 Validation definition

The validation strategy provides a framework for specifying and performing a set of validation activities, defining reference models, techniques, methods and rules to structure and standardise validation.

Operational Concept Validation could also be defined as the iterative process by which the usefulness of a newly developed system or operational concept is demonstrated.

5.2 Level of maturity

The identification of the level of maturity of the operational concept is a pre-requisite for developing the validation strategy/activities in accordance with the E-OCVM [10]. It structures the concept development and sets expectations concerning the evidence.

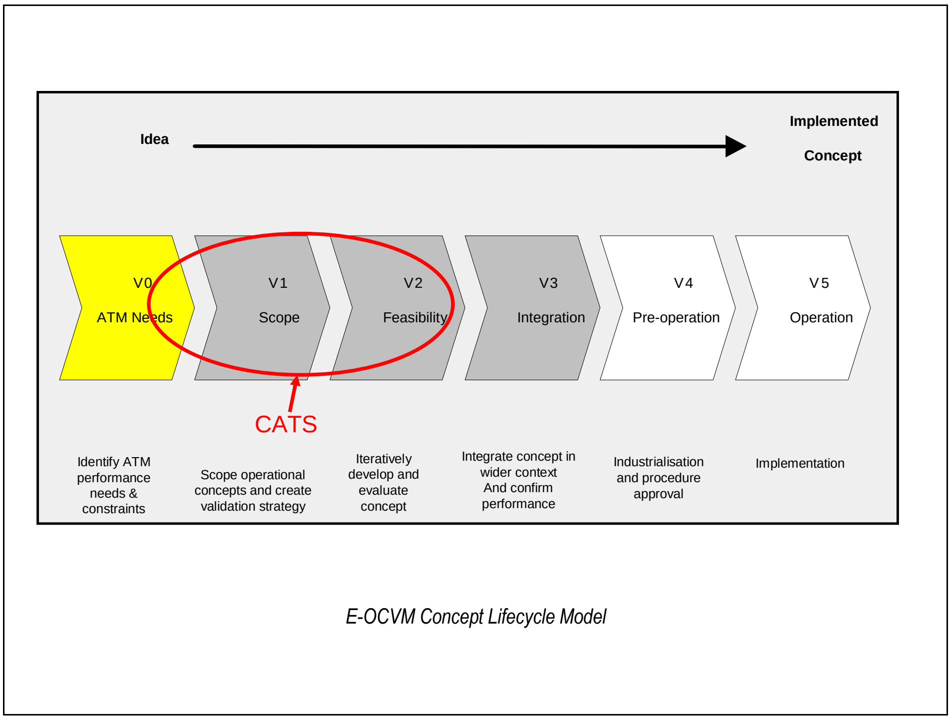

Figure 5: E-OCVM lifecycle model

As shown in Figure 5, the CATS Project will cover the following steps:

- V0: **ATM Needs**: The aim of this step is to identify the ATM performance needs, and current constraints/barriers. This task was mostly developed by SESAR D2 [5] and will be the basis of our work. CATS Deliverable D1.1 also added some interesting research questions.
- V1: **Scope**: The goal here is to describe the proposed solution and highlight anticipated benefits. This will be the role of WP1 through Deliverable D1.2.1 [2]. This step also covers the definition of the Validation Strategy. This document develops this issue.
- V2: **Feasibility:** The aim of this task is the evaluation of the proposed concept. It is an iterative process designed to show the operational feasibility of the proposed solution. Within the CATS Project, this task will be performed during the various assessments proposed in WP2 [2].

Regarding the SESAR Validation and Verification document [23], the level of maturity of HIL1& 2 will cover the V1 & V2 validation of the proposed concept using a prototype close to the operational platform; the last experiment, HIL3, as explained in 9.3, will only cover the V1 validation level.

5.3 Assessment techniques

Qualitative, quantitative and human factor techniques are required to perform a comprehensive assessment of the CoO. This assessment must also be underpinned by an in-depth legal analysis, identifying the feasibility of the objectives of a CoO.

5.3.1 Qualitative methods

When building the CATS Economic Feasibility Case, the analysis performed in Phases V0, V1 and V2 mostly relies on the judgment of independent experts. In this context, the following tools may be used [19]:

– *Analytical Hierarchy Process*: This technique looks at the decision, or decisions, to be made and builds a hierarchy of variables for comparison purposes based on the goals and objectives of the Economic Feasibility Case. Using mathematical formulas, the importance of each variable is calculated and used in the determination of the best outcome.

– *Influence Diagrams*: The influence diagram records the various factors and their relationship to each other which may influence the subject of the Economic Feasibility Case.

– *Decision Hierarchy:* This defines the scope of the decisions (i.e. resource allocation) both in terms of what is and what is not in the scope.

– *Causality Diagram*: This graphically documents the changes introduced in the ATM environment and delineates how those changes lead to benefits.

– *Cost Driver Representation*: This graphically represents the main cost drivers associated with the process by grouping them in different categories.

5.3.2 Quantitative methods

5.3.2.1 Fast-Time Simulations (FTS), and modelling exercise

These are typically performed to examine system performance, including the assessment of benefits (e.g. delay, fuel burn, time/distance flown) and the analysis of capacity, safety, risk and efficiency. These techniques are also useful for identifying potential problem areas where real-time simulations are necessary for further exploration. Fast-time simulation (FTS) is a type of virtual time simulation referring to the representation of the operation of a process or system (= target) over a finite time period by another process or system (= model) operating in a different but as short as possible timescale. FTSs are typically performed using computer or mathematical or statistical modelling.

- Computer models describe the target system through a computer program (state machine).
- Mathematical models describe the target system in terms of equations. Where these equations have closed-form solutions, these models are called analytical. In particular, mathematical programming or optimisation models identify the set of solutions used to minimise costs or maximise profits when constraints are imposed on the target system.
- Statistical models represent the target system through parameterised sets of probability distributions.

5.3.2.2 Human-in-the-Loop simulations (HIL)

HILs best enable the assessment of operational usability, acceptability and suitability. HIL simulations include rapid prototyping, small-scale research simulations and large-scale real-time simulations.

- Rapid prototyping studies can be used to develop HMIs for advanced concepts and conduct usability studies.
- Small-scale research simulations enable specific questions or issues to be examined in a relatively controlled, rigorous environment and are used mostly for experimenting.
- Large-scale real-time simulations (RTS). The boundary between small-scale and large-scale real-time simulations is not well defined. The main difference is obviously the number of controllers taking part in the exercise and hence the greater fidelity of the simulations.
- Gaming exercises can be used to evaluate strategic decisions of operational staff, e.g. traffic managers, airline fleet managers, in areas such as quality of services or capacity evaluation. Gaming exercises are essentially human-in-the-loop activities, suitable for concepts where the aim is to improve the strategic decision-making processes of the actors.

HIL experiments imply clear and direct feedback from the operational staff involved. They are supported by real-time simulations even if the real-time fidelity is likely to differ from a large-scale to a small-scale real-time simulation. Regardless of the real-time fidelity which contributes to the absolute fidelity, the main contributor component of HILs to operational assessment is functional fidelity.

Real-time simulations are used to predict system performance and to assess working methods, layouts, risk identification and training needs. The aim of real-time simulations is to produce an environment that represents the pertinent features of the realistic environment in which the system will operate, to be able to predict the system's operational performance.

The advantages of real-time simulation are that they:

- Provide an economical solution to life trials on operational systems, particularly if the studied system is in a validation phase.
- Assess operators' activity and behaviour which cannot be easily observed in real environments (accurate performance metrics, situational awareness, workload, error recovery, etc.).
- Provide a close link with the behavioural assessment methods and also the observational, interview and questionnaire methods which may be used to collect relevant data.
- Properly control the parameters and the variables of the assessment.

5.3.2.3 Optimisation or Mathematical programming models

Mathematical programming uses mathematical modelling and algorithms to arrive at optimal or good decisions in complex problems which are concerned with optimising the maxima (e.g. profit) or minima (e.g. cost loss, delay) of an objective function in order to improve or optimise the performance of the system under study.

Mathematical programming models have greatly contributed to the development in several important areas of the current global air transportation system [15]. The take-off of a specific aircraft at a given time in a given airport is the outcome of a long planning process that may have started several months before. The aim of such a process, generally referred to as schedule planning, is to maximise airline profitability [15]. Its main steps are the identification of which Origin-Destination pairs to connect, the frequency of departures and allocation to specific timings based on market preferences, the type and capacity of aircraft to be used, and the number of passengers served [17]. Further steps to be addressed by the airline are aircraft routing and crew scheduling. Because of the enormous complexity involved, even for airlines with only 15-20 aircraft, each stage of this process can only be efficiently solved by mathematical optimisation. Revenue management (i.e. the practice of determining the number of seats on each flight to be made available at each fare level, limiting low-fare seats and protecting seats for later-booking, higher-fare passengers) is another key area where airlines rely entirely on optimisation models and algorithms [15][18]. Finally, mathematical programming has also been successfully applied in optimising the design, development and operations of aviation infrastructure, i.e. airports and ATM [15]. An important application of such techniques in the ATM domain is the implementation of the ground-holding program (i.e., holding an aircraft on the ground when reduced capacity is anticipated) and its relationship with the collaborative decision-making initiative (see [16] and the literature therein cited).

5.3.3 Human Factors methods

5.3.3.1 Observational methods

Observational methods aim to obtain data by directly observing the activity or behaviour being investigated. They are used in research to assist in making accurate and relevant observations of events.

These methods are non-intrusive and can be achieved by:

- Recording visual and audible events using an observer in the field of the experimental setting. It is important for the validation process that the observer does not play an active role in the task along with the operator.
- Recording those events on appropriate media for later transcription into a format that is meaningful to the analysis required.

Observational methods require suitably-trained observers. Traditionally, observers are Subject-Matter Experts (SME) or Human Factors Experts (HFE).

Observations are not "free" but are structured by hypothesis and lists of indicators previously determined in the light of the experiment's purpose.

The advantages of observation are that it is possible to:

- Collect data which can only be recorded by observations (e.g. collective interactions, situational impacts, or operational strategies, deviations from rules, etc.).
- Collect objective data.
- Integrate the outcomes with others methods such as interviews or subjective methods.

The main challenge in observing operators is the way the data are collected in order to meet the experiment's purpose. Over-the-shoulder (OTS) observations must be conducted with specific observation forms to support the SMEs and HFEs. These forms are intended to structure the way the observation is performed, and to steer its focus toward clear and pre-defined objectives. The benefits of observation tools are that they:

- Make all members of the validation team adopt the same focus of observation.
- Ensure that relevant data are effectively collected.

- Ensure coherence and comparability of the data collected.
- Avoid bias caused by "free" observations.
- Make members of the validation team interchangeable during the observation.

Three techniques can be used to perform relevant observations:

- OTS rating scales.
- Behavioural and event checklists.
- Time-line scale.

5.3.3.1.1 OTS rating scales

The performance of operators is evaluated across broad dimensions. Potential performance dimensions for a set of rating scales are identified through reviews of previous literature involving air traffic control and aircrew activity, existing on-the-job-training forms, performance verification forms, and current project work on the development of behaviour summary scales.

For example, an OTS rating scale designed for ATCOs describes eight performance dimensions, and the behavioural descriptors developed for these dimensions help to provide a frame-of-reference for the evaluators. The eight dimensions are:

- Maintaining Separation.
- Maintaining Efficient Air Traffic Flow.
- Maintaining Attention and Situation Awareness.
- Communicating Clearly, Accurately, and Concisely.
- Facilitating Information Flow.
- Coordinating.
- Performing Multiple Tasks.
- Managing Sector Workload.
- An “overall” performance category is added at the end.

Each dimension is then graded on a 7-point effectiveness scale. The scale is clustered into three primary effectiveness levels; i.e. below average (1 or 2), fully adequate (3, 4, or 5), and exceptional (6 or 7).

5.3.3.1.2 Behavioural and Event Checklist

A second instrument requires the evaluators to focus on more detailed behaviour and activities, and record whether and how often each has occurred. The “Behavioural and Event Checklist” (BEC) requires evaluators to actively observe the subjects performing their tasks during each scenario and note behaviour such as:

- Failure to accept handover, to interpret weather information, to coordinate pilot requests, etc.
- Directive or rule violations.
- Read back / hear back errors.
- Unnecessary delays.

- Incorrect information input into the computer.
- Late frequency changes.
- Operational errors, deviations, and special use airspace violations.

5.3.3.1.3 Time-line scale

The way in which information is collected is driven by time. All events, behaviour or performance dimensions are collected in accordance with the time reference of the performed task.

Observation-specific tools have to be designed before each HIL in accordance with the HIL's purpose, the simulation platform features, and the available resources (human, techniques and materials) for performing the experiment and the data analysis.

The tools, as described above, have to be adapted to the nature of the operators' activity (ATCO, aircrew, airport, etc.) and can be combined.

5.3.3.2 Questionnaire

Questionnaires allow the gathering from a group of selected people of opinions about, or attitudes towards, specific things (including themselves) or events, and make it possible to collect specific items of data in an economic and structured way.

Questionnaires are designed to obtain a set of pre-determined, formatted questions administered to any number of people to elicit their views, for example as part of a survey.

The advantages of questionnaires are that they:

- Are economical to administer.
- Focus the survey on the issues in question rather than distracting respondents with secondary issues.
- Are robust, valid and reliable.
- Link with structured interview and observation outcomes methods to obtain structured data.

Questionnaire design must follow the following steps:

- Questions. The primary design principles are simplicity and clarity, avoiding technical terms, negatives and leading questions. This will ensure that all respondents interpret the items in the same way and avoid misinterpretation. Allow for a broad range of responses.
- Question types. These are dependent on the type of data required. The range of types includes multiple-choice questions, rating scales, paired comparisons, ranking according to a given criteria (all of which produce quantitative data), and open-ended questions (producing qualitative data).
- Administration. The appropriate population must be identified and an appropriate sampling strategy determined. Administration methods include self-completion questionnaires and structured interviews.
- Confidentiality. The respondents' data must be kept confidential.

The questionnaire, which will be used during the HIL, will be completed only by the operators performing the experiments. For this reason:

- It should be short to encourage sufficient thought to be given to each question.
- The questions should be accurate so that there is only one possible interpretation.
- The questions should avoid abstraction and concepts and focus on practical issues.

The questionnaires that will be submitted to the operators are:

- Post-run questionnaire. The questionnaire topics relate to feelings about the run, perception of the task, satisfaction, performance, risk management, safety, HMI features, the impact of tools, difficulties encountered, improvements required, etc.
- Post-experiment questionnaire. The questionnaire topics relate to general feelings about the concept, changes in the working methods, pros and cons, concept evolution, usability and acceptability.

5.3.3.3 Structured interviews

Structured interviews are meetings arranged for a particular purpose using a predetermined sequence of interview questions. The aim is to have a systematic collection of task-based data.

An interview generates general information on the issue in question whereas a survey interview has a more specific goal and is more systematic.

There is a need to design the interview format with the analysis in mind, to determine the type of data required (e.g. quantitative, qualitative, nominal, interval, etc.) and also to administer the interview in an environment that is likely to elicit response, usually an environment familiar to the interviewee or linked to the task under investigation.

As the interview data must be kept confidential, the interviewer must be informed to that effect. A rapport must be established to facilitate and elicit responses, with the interviewer coming across as confident and competent.

During the interview the relevance of the process must be maintained and the interviewer should guide that process.

The questions can be open, semi-open, or closed. The interview management can also be directive, semi-directive or open. In order to optimise the time allocated to interviews and meet the experiment objectives, a semi-rigid approach is recommended.

Semi-rigid interviews ensure that the relevant topics will be investigated for each operator, and that each operator will have the freedom to discuss topics which are not scheduled. Semi-rigid interviews are a way of being both systematic and open.

The advantages of interviews are that they:

- Provide a natural means of data gathering that is relatively economic.
- Can supplement other methods (e.g. critical incident and questionnaires).
- Are flexible in allowing interesting and unexpected information to be explored.
- Structure the interview, imposing consistency and standardisation on the process.

Interviews can also be structured around data collected during the HIL run (performance data, traces of activity, communications, verbal reports, etc.). This technique is called self-confrontation. The aim of self-confrontation is to rebuild the dynamics and the movie of the run in order to catch the situation awareness, the decision-making, the risk taking, the workload management, and the collective processes used during the run. The focus then is to describe and explain the performance achieved, and how the operator manages the constraints of the task (environment, HMI, tool, etc.) and his own cognitive limitations.

Interviews that will be submitted to the operators are:

- Post-run debriefing.
- Post-experiment debriefing.

During the post-run debriefing, the operators will be provided with different kinds of information and will be asked to:

- Discuss the performance of the concept.
- Reason their performance using the information provided by the CWP. Self-confrontation techniques will be used.
- Make the comparison between the activity performed with and without the CoO.

During the post-experiment debriefing, the operators will be provided with various kinds of information and will be asked to:

- Globally evaluate the pros and cons of the concept.
- Explain the limitations and benefits of the concept
- Envisage enhancements to the concept and working methods.
- Envisage the concept in the framework of SESAR development.

It is important to note that interviews and observational methods are closely interconnected techniques. This means that on one hand, data collected using the observation techniques are checked and discussed during the interviews, and on the other hand that insights emerging during the interviews are used to guide subsequent observations. The combination of methods has been proven to ensure the correctness, reliability and consistency of the outcome.

5.3.3.4 Psychometric scaling

Psychometric scaling is a subjective measurement technique using paired comparisons, ranking, sorting or rating to evaluate or measure things or events.

The aim of scaling is to quantify, both in degree and direction, subjects' views on the things or events in question.

A rating scale consists, usually, of a straight line with five, seven or nine equal divisions along with a series of adjectives (with or without a number) which describes the stimulus attribute. This Likert-type of rating scale, as it is known, is the one most often used.

Psychometric scaling is very attractive as it is fairly inexpensive and easy to administer. It can be applied to almost everything. The use of rating scales is by far the most frequently employed subjective measurement technique.

This method can be used by an observer and is then called observer rating (see observational methods above), or by the operator himself, and is then called self-rating.

Self-rating techniques are frequently used to assess situation awareness and workload. During HILs, situation awareness will be assessed by one technique, the SASHA-Q simplified version, and workload by two techniques, the ISA and NASA-TLX.

- SASHA-Q

The Situation Awareness for SHAPE Questionnaire (SASHA-Q) is a simplified version of the original SASHA questionnaire. The SASHA-Q is a self-rating scale technique in which the controller has to assess his own opinion regarding six items describing six dimensions of the ATCO's situation awareness. Each item is assessed with a seven-level Likert-scale.

- ISA

The Instantaneous Self-assessment of Workload (ISA) was developed as a tool that an operator could use to estimate his/her perceived workload during real-time simulations. The operator is prompted at regular intervals to give a rating of one to five of how busy he/she is (one means utilised, five means excessively busy). These data can be used to compare operators' perceived workload, for example, with and without a particular tool, or between different systems (e.g. in the CATS experiments, for with and without CoO scenarios).

The ISA is appropriate for use during completion of the run.

- NASA-TLX:

The NASA Task Load Index (NASA-TLX) is a multi-dimensional rating scale used by operators to report their mental workload. It uses six dimensions of workload to provide diagnostic information about the nature and relative contribution of each dimension in influencing overall operator workload. Operators rate the contribution made by each of the six dimensions of workload in order to identify the intensity of the perceived workload. The dimensions are:

  - Mental demands.
  - Physical demands.
  - Temporal demands.
  - Performance.
  - Effort.
  - Frustration.

A paired comparison procedure, in which the operator compares each of the six subscales with each of the others in turn, is used to identify the relative importance (weighting) of each dimension in contributing to the perceived workload.

Overall workload is then calculated by multiplying each rating by the weighting attributed to that dimension by the operator and the sum of the weighted ratings is then divided by fifteen (the number of paired comparisons).

The NASA-TLX can be used after the simulation trials as part of the post-simulation debriefing procedure. It is simple, quick and easy to apply.

5.3.3.5 Objective performance measurement

The aim of objective performance measurement is to obtain quantitative data on operator/team performance, but without the extreme subjectivity of observation and judgment in order to evaluate operator/team performance by means of objective measures such as timing and counting of events.

In a system of reasonable size the number of performance indicators that could be measured is enormous. However, when considered in terms of their similarity, they reduce to a few generic measures:

- Time (e.g. reaction time, duration).
- Frequency (e.g. number of responses, observations, communications, etc.).
- Accuracy (e.g. correctness of responses, observations, errors made, etc.).
- Quantity (e.g. response magnitude or quantity achieved and/or used).
- Traces of activity (e.g. gaze, actions, use of tools, etc.).
- Behavioural state (e.g. situation awareness).

Objective performance indicators can be recorded using observational techniques (e.g. data sheets, pen and paper) but mainly using automatic techniques linked to the running of the simulation platform (e.g. computer-based data-logging).

The list of collected data depends on the experiment hypothesis, simulation platform features and available technical resources. Where the list of performance indicators has been drawn up, the designer should be aware that a huge amount of data will generate a heavy load that cannot always be handled with the available resources. The choice of the relevant indicators is essential for achieving the experiment's objectives.

Among performance indicators, communication events are a useful addition for assessing collective work, as also are cognitive processes, situation awareness or workload. Communication events are individual transmissions that occur between pilots and controllers and between controllers and other controllers. Measures of communication events include number, timing, and content of transmissions, and identity of speaker.

Communication indicators are recorded routinely, are strongly correlated to other cognitive activities, and their content or other affective components may reflect the subjective effort experienced by controllers.

Performance measures need to be combined with other method outcomes to properly address topics such as activity, cognitive processes, human error, situation awareness, workload or collective work.

5.3.3.6 Integrated Task Analysis

Integrated Task Analysis is an analysis method covering various approaches addressing the behavioural and cognitive aspects of the operator's tasks in order to compare different work organisations or concepts. The aim of the Integrated Task Analysis is to investigate the cognitive aspects with an emphasis on the cognitive processes and working style of operators.

The task analysis is conducted by applying a set of methods, already described above, comprising of:

- Cognitive interviews.
- Behavioural observations.
- Post-run and post-experiment interviews.
- Self-confrontation.
- Additional questionnaires or self-rating scales.
- Performance data analysis.

The outcomes of the Integrated Task Analysis are to describe and explain the cognitive processes, cognitive profiles and working styles of the specific position/unit. Another outcome is identification of the users' needs in order to delineate the strengths and weaknesses of the concept, and finally propose recommendations for improvements.

6 Validation expectations

The planned assessments in the CATS Project are grouped into the following two approaches:

- operational acceptability, performed by three "case-focused" HIL experiments
- systemic assessments, based on a safety-case study and risk analysis, a legal study and also a cost-benefit assessment

The design and assessment of the CoO need to be commonly agreed by the actors involved through a global approach, consistent with the E-OCVM methodology [10]. The strategy is to involve all actors in participative design sessions and perform "restricted" HIL experiments (i.e. with a reduced number of actors) regarding the impact of the CoO between ANSP actors (i.e. ATCOs) (WP2.1.1), and between airlines and ANSP actors (i.e. aircrew and ATCOs) (WP2.1.2). Finally, an evaluation of the renegotiation process involving all actors (i.e. airlines, airports and ANSPs) is envisaged (WP2.1.3). This is the way proposed by the CATS Project to apply the iterative process common to the feasibility phase.

Looking in the DOW [2] the following high-level validation expected outcomes can be identified for the CoO's implementation:

- validate whether the CoO's implementation is acceptable for pilots, ATCOs, airport operators
- validate whether the CoO integrates flexibility to cope with uncertainty
- validate whether the concept improves system efficiency
- validate whether the concept will produce benefits
- validate whether the CoO's implementation will allow safe operations

The CATS concept of operations [8] covers the overall scope from the ATM Planning Phases to the post-flight analysis; nevertheless, the CATS operational assessments will mainly focus on the impact of the CoO and TWs through the trajectory-based operations. Consequently the CATS operational assessments will measure the contribution of the CoO and associated TWs in achieving some of the SESAR KPAs, even if not all the SESAR components are applicable.

7 Validation objectives

The assessment provided for in the CATS Project [2] focuses on two main aspects:

- Operational assessment: The aim of this assessment is to evaluate the operational acceptability of the CoO and associated TW concepts from the point of view of the various actors.
- Systemic assessment: This assessment will concern safety and risks, costs and benefits, and legal analyses. The safety and risks assessment will focus in particular on the TW concept oriented towards safety case studies. The benefits and legal assessments will focus on the CoO concept.

The high-level objectives of the operational issues are to assess the impact of the CoO on the various ATM actors. The first set of assessments proposes to evaluate this concept by means of a three-line approach, ensuring that all the issues of the CoO are addressed:

- impact of the CoO between ATCOs
- impact of the CoO between ATCOs and aircrew
- renegotiation process, involving airlines, airports and ANSPs

The main aim of the safety assessment is to determine the safety implications of the TW concept in terms of:

- the predictability it imposes on the traffic
- the flexibility it provides for conflict resolution manoeuvres
- the probability of renegotiation as a consequence of conflict resolution

The main aim of the benefit assessment is to identify benefits and drawbacks in monetary and non-monetary terms following the implementation of the Contract of Objectives for the various stakeholders (ANSPs, airports, and airspace users). The analysis starts from the Initial Shared Business Trajectory and considers both the ATM planning and flight execution phases.

The legal assessment is designed to identify the legal parameters regulating the provisions of the CoO. Those parameters can be found in public international law, private international law, European law, including the regulations of the European community and EUROCONTROL, and national law. The objective of this analysis is to determine to what extent the proposed provisions of the CoO are compatible with the parameters which can be found in the said legal framework.

7.1 Operational issues

On the basis of the stakeholders' concerns and validation expectations, two specific groups of validation objectives can be highlighted:

- objectives relating to system performance (ATM KPA)
- objectives relating to human performance

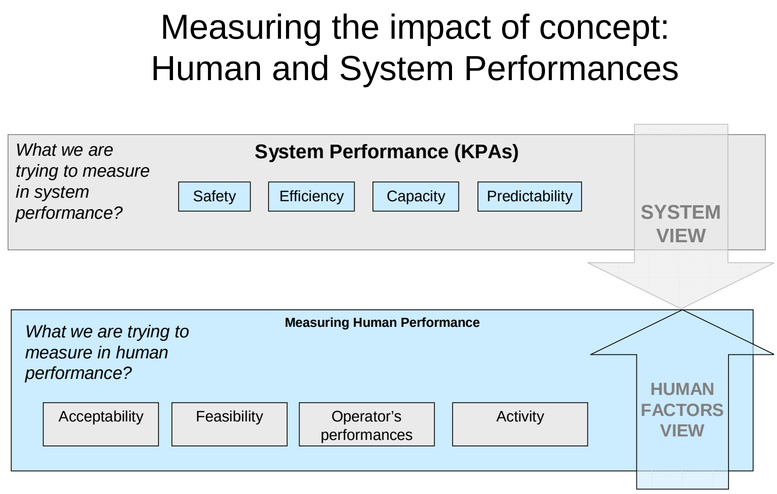

Figure 6: Measuring the impact of a concept

The first objective (system performance) is a key aspect of the validation. The aim is to assess whether the benefits have been delivered as proposed.

The aim of the second objective (human performance) is to assess whether the involvement of human beings in the proposed concept is consistent with their capabilities and characteristics (working methods, usability of the HMI proposed, task sharing, etc.). The goal is also to see whether the contribution of human beings to the overall system performance is within the expected limits (workload, stress, SA, etc.). Human performance could be seen as an enabler for achieving system performance.

7.1.1 Objectives relating to system performance (ATM KPA)

On the basis of stakeholder concerns and the performance framework [Chapter 4], five of the eleven SESAR KPAs were identified as being potentially improved by the CoO and introduction of the associated TWs, mainly:

- Safety (SAF)
- Capacity (CAP)
- Efficiency (EFF)
- Predictability (PRED)
- Flexibility (FLEX)

On the basis of the validation expectations [Chapter 6], the following high-level hypotheses were identified:

- Safety (SAF): the implementation of the CoO will not degrade safety.
- Capacity (CAP): the CoO is still manageable even with the increase in traffic expected in 2020.

- Efficiency (EFF): implementation of the CoO positively impacts aircraft output in the sector (flight duration, etc.).
- Predictability (PRED): implementation of the TWs will ensure compliance with schedules.
- Flexibility (FLEX): the renegotiation process is manageable and efficient and operators have the room for manoeuvre to deal with uncertainty.

7.1.2 Objectives relating to human performance

The objectives relating to human performance can also be seen as specific validation goals, related to the contribution of human beings to overall system performance. This is another way of achieving the validation expectations (ATM KPA).

Validation goals make it possible to carry out the high-level hypothesis assessments presented above (7.1.1). These are fundamental goals on the basis of which the HIL experiments are designed and performed, as shown in Figure 7.

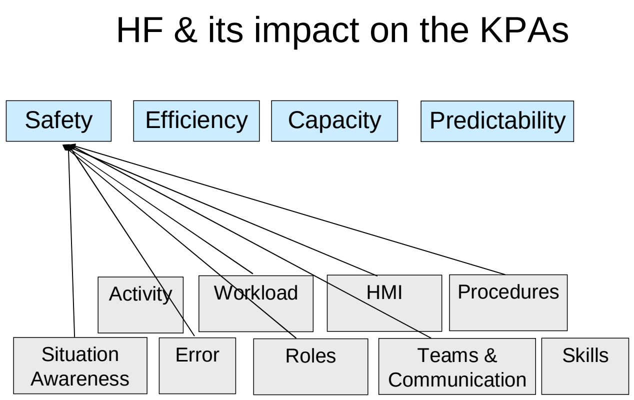

Figure 7: Example of HF impact on ATM KPAs

The impact of all potential individual HF elements on each KPA is considered:

The objectives of human performance validation are twofold:

- Describing and explaining factors affecting the high-level hypothesis assessment outcomes.
- Assessing actor-level hypotheses about working practices and activity.

Four validation goals have been identified:

- G1: Assess the feasibility and acceptability of the working methods in connection with execution of the CoO.
- G2: Assess the impact of CoO execution on ATCO and aircrew performance.
- G3: Assess the impact of CoO on ATCO and aircrew activity.
- G4: Assess strategic decisions in the event of renegotiation and their impact on traffic.

These validation goals can be broken down into several low-level objectives, which would then drive the various validation exercises.

7.1.2.1 Feasibility and Acceptability (G1) low-level objectives

The following low-level objectives concerning "Assess the feasibility and acceptability of the working methods associated with execution of the CoO" G1 are proposed:

- G1.1: Assess the feasibility and acceptability of the ATCOs' working methods associated with execution of the CoO, namely:
  - Task-sharing between executive and planner ATCOs.
  - Task-sharing between executive ATCOs and aircrews.
  - Task-sharing between planner ATCOs from different sectors.
  - Task-sharing between planner ATCOs from different ANSPs.
  - Task-sharing between ATCOs and ANSP (FMP, etc.).
  - ATCOs' working methods with CoO execution (with and without renegotiation of CoO).
  - Consistency and complementarities with classical ATCO tasks.
  - Controller working position HMI and tools to support the ATCOs' working methods.
- G1.2: Assess the feasibility and acceptability of the aircrew's working methods associated with execution of the CoO, namely:
  - Task-sharing between aircrews and executive ATCOs.
  - Task-sharing between aircrews and airline operators (AO).
  - Aircrews' working methods with CoO execution (with and without renegotiation of CoO).
  - Operational continuity along overflown sectors and ANSPs.
  - Consistency and complementarities with classical aircrew tasks.
  - Flight deck - cockpit HMI and tools to support the aircrews' working methods.
- G1.3: Assess the feasibility and acceptability of the airport, ANSP, and airline operator working methods associated with execution of the CoO, namely:
  - Task-sharing between airports, ANSPs and airlines.
  - Tools to support the airports, ANSPs and airline tasks.
  - Consistency and complementarities with classical ATM tools.
  - Airport, airline and ANSP working methods with CoO renegotiation process.
  - Task-sharing between ANSPs and ATCOs.
  - Task-sharing between airline operators and aircrews.

7.1.2.2 Performance (G2) low-level objectives

The following low-level objectives concerning "Assess the impact of CoO execution on ATCOs' and aircrews' performance" G2 are proposed:

- G2.1: Assess the impact of CoO execution on ATCOs' performance, namely:
  - Safety issues
  - Efficiency
  - Flexibility to cope with uncertainty
- G2.2: Assess the impact of CoO execution on aircrews' performance, namely:
  - Safety issues
  - Efficiency
  - Flexibility to cope with uncertainty

7.1.2.3 Activity (G3) low-level objectives

The following low-level objectives concerning "Assess the impact of the CoO on ATCOs' and aircrews' activity" G3 are proposed:

- G3.1: Assess the impact of the CoO on ATCOs' activity, namely:
  - Workload
  - Situation awareness
  - Human error and recovery
  - Individual impact
  - Collective impact (ATCO vs. ATCO; ATCO vs. aircrew)
  - Organisational impact (ATCO vs. ANSP; aircrew vs. airline)
- G3.2: Assess the impact of the CoO on aircrews' activity, namely:
  - Workload
  - Situation awareness
  - Human error and recovery
  - Individual impact
  - Collective impact (ATCO vs. ATCO; ATCO vs. aircrew)
  - Organisational impact (ATCO vs. ANSP; aircrew vs. airline)

7.1.2.4 Renegotiation (G4) low-level objectives

The following low-level objectives concerning "Assess the strategic decision in the event of renegotiation and the impact on traffic " G4 are proposed:

- G4.1: Assess the strategic decision process, namely:
  - Schedule
  - Data exchanges
  - Actors' interactions
  - Normal and stressful situations
- G4.2: Assess the impact of renegotiation on the tactical situation, namely:
  - Workload

- Situation awareness
- Collective impact (ATCO vs. ATCO; ATCO vs. aircrew)
- Organisational impact (ATCO vs. ANSP; aircrew vs. airline)

7.1.3 Performance indicators

This section provides a list of indicators and the associated metrics to be measured during the various HILs.

This section is in line with the indicators proposed by in the EP3 Framework Assessment Chapter 7.3.3[11].

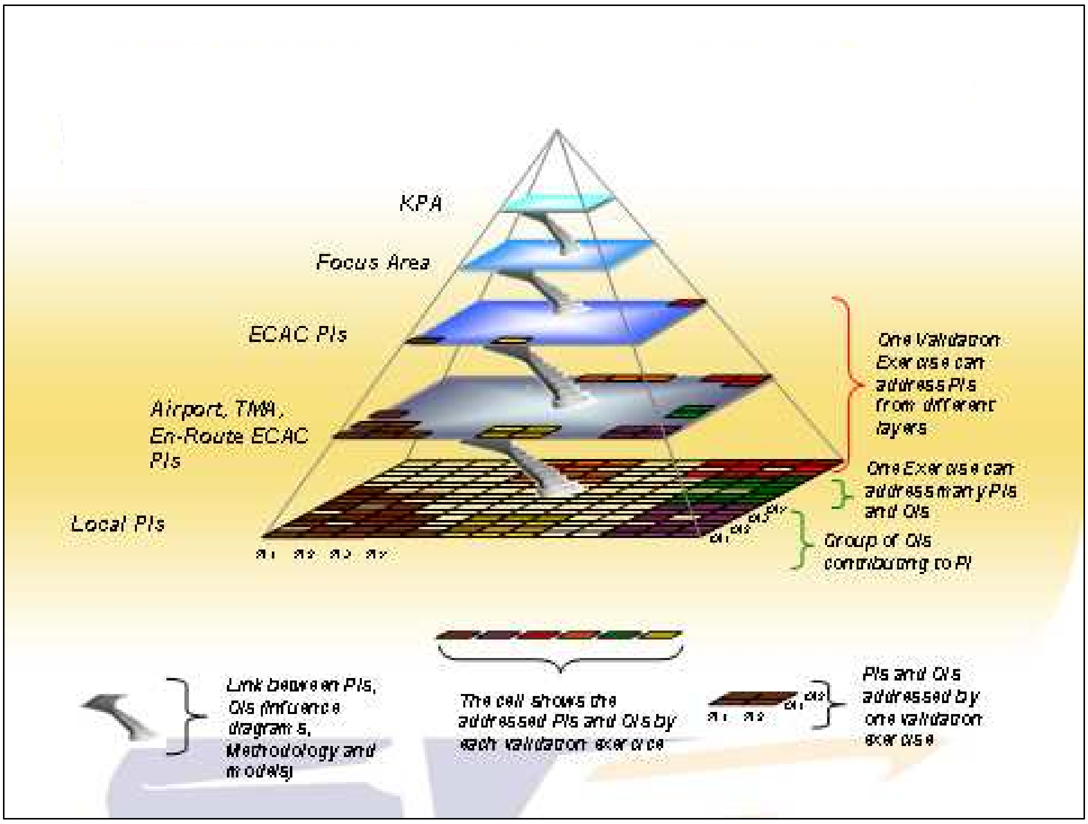

Figure 8: EP3 Performance Framework

As CATS operational assessments will evaluate the operational acceptability of the CoO and associated TW concepts from the point of view of the various actors between two ACCs and not the ECAC-wide performance of en-route, we will focus only on the Local Performance Indicators (PI) layer.

From the catalogue of performance indicators, CATS has extracted a number of local PIs relevant to its concerns and has also added indicators relating to the concept.

7.1.3.1 System performance - ATM KPA indicators

The various HIL experiments scheduled in WP2.1 will measure:

- SAFETY:
  - SAF.LOCAL.ER.PI 1: Number of hours with ATC workload overloads. Times the controller is saturated with various degrees of severity, therefore resulting in risky situations and then safety precursors.
  - SAF.LOCAL.ER.PI 2: Number of hours with ATC workload underloads. Not only overload but also underload situations can cause safety-critical situations, according to relevant studies.
  - SAF.LOCAL.ER.PI 3: Number of ATC tactical interventions. Number of times an ATC tactical intervention is needed in order to avoid a potential conflict in the air; Number of losses of minimum separation.
  - SAF.LOCAL.ER.PI 5: Conflict alerts. Number of STCA conflict alerts occurring during the simulation.
  - SAF.LOCAL.ER.PI 6: Separation between aircraft pairs in conflict:
    - Vertical
    - Horizontal
  - SAF.LOCAL.ER.PI 8: Number of resolutions.
- EFFICIENCY:
  - Number of TWs fulfilled.
  - Airspace crossing: time, distance, etc.:
    - EFF.LOCAL.ER.PI 1: Actual Flight Time. Actual aircraft flight time taking into account the restrictions in the system
    - EFF.LOCAL.ER.PI 7: Efficiency of routing service. Comparison of actual routing to initial shared business trajectory
    - EFF.LOCAL.ER.PI 8: Duration Increase. This is the ratio of actual trajectory duration divided by optimum trajectory duration
    - EFF.LOCAL.ER.PI 9: Lateral deviation. Mean lateral deviation between SBT route and actual flown route
    - EFF.LOCAL.ER.PI 10: Vertical deviation. Mean vertical deviation between SBT route and actual flown route
    - EFF.LOCAL.ER.PI 11: Number of flights able to fly the requested altitude. Number of flights whose max. altitude equalled the requested altitude in their SBT
  - Number of ATCO orders:
    - CAP.LOCAL.ER.PI 10: Speed clearances
    - CAP.LOCAL.ER.PI 11: Heading clearances
    - CAP.LOCAL.ER.PI 12: Altitude clearances
    - CAP.LOCAL.ER.PI 13: Number of directs
- CAPACITY:
  - CAP.LOCAL.ER.PI 8: Aircraft count per sector. A measure of the number of aircraft per sector
  - CAP.LOCAL.ER.PI 2: Maximum hourly throughput. Maximum number of controlled aircraft per hour in the airspace volume

- Load of control (number of A/C x airspace crossing time)

- PREDICTABILITY
  - PRED.LOCAL.ER.PI 1: Flight Time Deviation (min). For each flight, the difference between the actual flight duration and the optimum flight duration (no constraints)
  - PRED.LOCAL.ER.PI 2: Flight Time Standard Deviation (no unit)

7.1.3.2 Human performance indicators

For the G1, G2 and G3 validation goals, specific indicators and metrics are required to describe, understand and propose enhancements for the future CATS concept. These indicators are:

- WORKLOAD:

Workload is the strain placed on the actors (ATCOs or aircrew). In other words, the workload is the physical and mental cost required of the operator in order to perform the task. This cost depends on the task's features, but also on the knowledge, the experience and the level of skill of the operator. Finally, the level of fatigue and stress of the operator contributes greatly to the cost. Consequently, various metrics are required to assess the level of workload.

  - Self-assessment subjective techniques: ISA, NASA-TLX. These methods make it possible to rate the actors' opinions of their own workload.
  - Observations: OTS, BEC. Observational methods make it possible to assess the operator's performance, task management strategies, and traces of activity.
  - Post-run debriefings: questionnaire, interview, replays, self-confrontation. These methods make it possible to elicit perceived information on knowledge used, the decision-making process, risk management, operative modes, and mental conflicts.
  - Objective data from video, audio and platform (actions, traffic, etc.) recordings. These data make it possible to analyse the collected observational and subjective data by matching them.

- SITUATION AWARENESS ('PICTURE'):

Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future. Situation awareness encompasses three levels of awareness:

  - Level 1: perception of specific key elements in the situation
  - Level 2: comprehension and integration of that information in light of operational goals
  - Level 3: projection and anticipation of future states of the situation in the near future

Situation awareness depends on the environment, its complexity, its dynamics, its flow of information and its unexpected events. In addition, the way the operator manages his/her task contributes greatly to the way the situation is understood. The way the task is managed depends also on the knowledge, skills, and level of fatigue or stress of the operator. Consequently, the way in which situation awareness is assessed requires several metrics:

- Self-assessment subjective techniques: Simplified version of SASHA_Q. This method makes it possible to rate the actors' opinion of their situation awareness
- Observations: OTS, BEC. Observational methods make it possible to assess the operator's performance, task management strategies, and traces of activity
- Post-run debriefings: questionnaire, interview, replays, self-confrontation. These methods make it possible to elicit perceived information on knowledge used, the understanding of the situation, anticipations, hypotheses on future the situations, risk management, operative modes, and mental conflicts
- Objective data from video, audio and platform (actions, traffic, etc.) recordings. These data make it possible to analyse the observational and subjective data collected by matching them. They are very important in the situation awareness assessment in order to provide an objective view of the situation, which may be different from what is perceived by the operator

- ERROR PRODUCTION and RECOVERY:

Error is a generic term encompassing all those occasions when a planned sequence of mental or physical activities fails to achieve its intended outcome, and when these failures cannot be attributed to the intervention of some change agency. Error is an involuntary deviation from the preliminary intention to achieve. In other words, the operator had the intention to do what was necessary, but did not achieve it.

Errors are common in the activity of the operators, because they are the counterpart of the performance. However, errors are not synonymous with accidents because many of them are detected and recovered, or have no impact on safety and efficiency.

The analysis of error production and recovery is therefore a key element of safety management. It is also a way to better understand the cognitive constraints generated by the situation. Several indicators are useful for assessing error production and recovery:

- Observational methods: OTS and BEC. The purpose is to list the most frequent errors and their detection and recovery strategies
- Post-run debriefings: questionnaire, interview, replays, self-confrontation. These methods make it possible to elicit from the operator the errors made, their consequences for safety, and the detection and recovery strategies used. The post-run debriefings are useful for understanding the causes of errors in relation to the cognitive challenge posed by the situation
- Objective data from video, audio and platform (actions, traffic, etc.) recordings. These data make it possible to analyse the observational and subjective data collected by matching them. They are relevant in cases of error assessment for comparing what was perceived by the operator and what was the reality of the situation

- OPERATOR'S ACTIVITY:

The operator's activity is the way in which the operator performs the task. The activity results in mental and physical behaviour which is the outcome of the balance between the perceived information, the operator's goals and the available mental and physical operator's resources. The activity differs from the task because it is the result of what the operator is able to do under the circumstances of the task. The task is the formal way in which the goals are achieved, while the activity is the operator's adaptation to the reality.

The operator's activity knowledge is relevant for assessing the cognitive limitations arising from the concept and the HMI. It is a fruitful way of understanding the performance criteria and making relevant human-centred proposals for the next steps of the project.

The activity analysis requires several metrics:

- Observational methods for identifying the behavioural changes caused by the addition of the CoO to the operator's task. The expected changes may occur in perception, understanding, diagnosis, and decision-making.
- Post-run debriefings: questionnaire, interview, replays, self-confrontation. The metrics make it possible to investigate the knowledge used, the cognitive limitations caused by the HMI or the situational constraints, the dynamics for managing the situation, etc.
- Objective data from video, audio and platform (actions, traffic, etc.) recordings.

- COLLABORATIVE ACTIVITY (ATCO/ATCO - ATCO/aircrew):

Collaborative activity reflects the communicative and cooperative processes between operators to fulfil the task and meet the operational goals. Collaborative activity knowledge is relevant for assessing the needs of the collaborative work, the relevance of the available collaborative tools on the working position layout, the limitations caused by the CoO’s management for achievement of the task, and the initially proposed task-sharing between actors. Several metrics are required to assess collaborative activity:

- Communication assessment (flow, numbers, contents, media, etc.) in order to quantify communication performance and needs
- Responsibility field
- Observational methods for describing relationships
- Post-run debriefings: questionnaire, interview, replays, self-confrontation. These metrics are used to assess the cognitive mechanisms of collaboration, their limitations, and the way the constraints are managed
- Objective data from video, audio and platform (actions, traffic, etc.) recordings

- ORGANISATIONAL ASPECTS:

The organisational aspects of the task reflect the front-line operator's relationship with the upper operational management levels (e.g. the FMP for the ATCOs or the AOC for the aircrew). The organisational aspects of the work also affect the relationship between airport, airline and ANSP in the renegotiation process. The focus of HIL experiments is not an in-depth assessment of the organisational aspects. It is to enhance the recommendations on the organisational components derived from the outcome of the operational assessment. Several metrics are required:

- Observational methods to describe the necessary functional links with upper operational management
- Post-run debriefings: questionnaire and interview
- Objective data from video, audio and platform (actions) recordings

7.2 Systemic issues

The systemic assessments provided for in WP2.2 will bring evidence to the concept within a more global (macroscopic) approach.

The following high-level hypotheses have been identified:

- Risk and safety: the implementation of CoO will not increase risks and will not degrade safety
- Benefits: the implementation of CoO has positive benefits, including economic ones, for all partners, including airlines and airports
- Legal impact: CoO will make it possible to clearly define the responsibilities and liabilities of each partner

7.2.1 Safety

The current excellent safety levels need to be maintained and even improved to mitigate the safety risks posed by growing traffic and congestion. Safety risks quadruple when traffic doubles [14]. These risks fall into the scope of air traffic management (ATM) and air navigation services (ANS) as well as into the areas of aerodrome infrastructure and equipment and its operation. However, these subjects are not as yet included in the CATS Project.

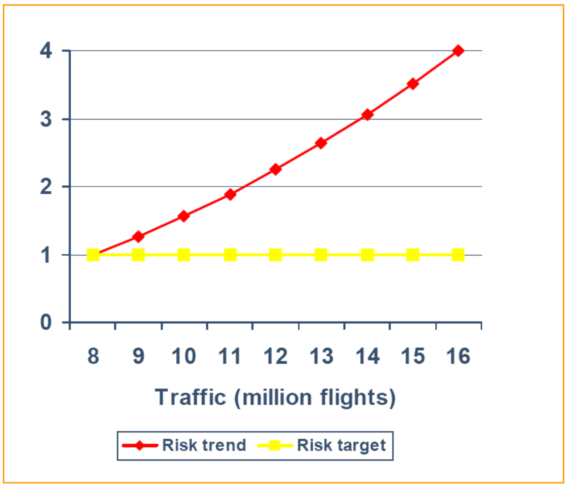

Figure 9: Risk quadruples when traffic doubles

The safety assessment provided for in WP2.2.1 involves fast-time simulations (FTS) of air traffic in a set-up comparable to that of the HIL experiments. The results of WP1.3 will be used to set up a simulation study on a fast-time simulator developed by ETH. The study will progress in a series of steps:

- TW and predictability. Monte-Carlo simulations will be used to determine the potential safety improvements as a result of the greater predictability of the air traffic afforded by TW. This study (together with the following one) is considered a precursor (base case) for central studies 3 and 4.

- Renegotiation probability. Monte-Carlo simulations will be used to determine the probability of aircraft missing their target windows, in the absence of intervention from air traffic controllers.
- TW constrained conflict resolution. Conflict resolution algorithms respecting the TW will be developed and introduced in the simulations. The aim will be to determine the effect of these algorithms on conflict probability (improvement expected) and renegotiation probability (deterioration expected).
- Conflict resolution with renegotiation. Conflict resolution algorithms will be developed for cases where resolution under the TW constraints is impossible. The aim will be to determine the safety implications of the concept, while minimising the knock-on effect of conflict avoidance on traffic.
- Human-in-the-loop effect. The algorithms of steps 3 and 4 will be modified to take into account the adaptation and preferences of air traffic controllers with respect to operations under the TW concept, identified in the HIL experiment. Studies 3 and 4 will be repeated to estimate the potential additional impact of controller actions.

7.2.2 Benefits assessment

CoO and the related TW concepts are expected to introduce some benefits previously introduced in Section 3.5, as well as some costs for each category of stakeholders. Benefits and costs are explicitly detailed in Section 0. An economic impact assessment both in quantitative and qualitative terms [19][20] will be conducted in order to support decision makers on the financial and economical viability of the implementation of the Contract of Objectives. To achieve this objective a performance assessment will be carried out, in line with the methodology adopted by SESAR and the performance architecture defined by E-OCVM [10], EP3[11] and CAATS II [19][20], in order to obtain the indicators for the various layers of the system.

After considering the extensive description of the concept of operations contained in this report and in the other available CATS documents, the practical implementation details involved in the new concept of operations are focused on. At each hierarchical level, the introduction of the CoO and of the TWs may imply that the organisation of each stakeholder needs to be modified. These adaptations are identified and analysed.

In particular, three stakeholders are considered, namely airport operators, aircraft operators and ANSPs. Moreover, three main hierarchical levels are considered, reflecting the various phases characterising the introduction of the CoO these are the organisational, flight planning and flight execution levels.

At organisational level, before introducing the CoO/TW concepts, some investment may be needed. Examples could be the purchase of new equipment, the reorganisation of airspace, or the cost of adapting the organisations to a new concept of operations. Moreover, training will be needed for the human resources.

Once the CoO/TW concepts are employed in the ATM system, it is possible to distinguish flight planning (long, medium and short-term) and execution phases in the process of preparation and execution of each flight. The long-term planning phase takes place several months before the day of operation. Flight request evolution and potential problems are analysed between airports and aircraft operators. During the medium term planning phase airlines, airports and ANSPs negotiate TWs, while the final requests of all airspace users and any constraints become delineated more and more precisely in the short-term planning phase. Finally, in the execution phase, TWs are contractualize through CoO signature and late flight requests are incorporated in the panorama of existing CoOs. The flights are actually executed, and CoO renegotiation is implemented when necessary (when TWs can not be reached).

A two-phase negotiation takes place before the CoO is signed.

The first step of the network planning is the elicitation of the airspace user request. In the light of the planned schedules and the experience of previous years regarding scheduled and unscheduled flights, a first draft of the airspace user request is made available to all airports. In a second step, airports assess their planned runway capacity, taking into account seasonal variability, expected events, and weather tendencies. This step can be described as the airport resource planning process which proposes an initial plan for the slot coordination process. The first set of TWs, related to departure and arrival airports, is then defined. If there are no mismatches between requests and resources, the initial plans are published, reviewed and accepted by the partners. If there are mismatches, a negotiation process is engaged between airports and users, until satisfactory solutions are found. If an agreement cannot be found between airspace users and airports, the ATM-NMF plays its role of moderator, facilitator or decision maker. In all cases, TWs drafting runs at regional or sub-regional levels under the supervision of the ATM-NMF. The solutions are, at this level of the Network Planning process, strategic, and this phase is referred to as First Round Negotiation.

Once the initial plan has been designed, it is made available to all ATM partners. Flight intentions are known in the form of Shared Business Trajectory and TWs relating to departure and arrival airports. The ANSPs now try to adapt the airspace structure and their working methods to the airspace user requests. The purpose is to avoid airspace bottlenecks, and to adjust the ANSP resources available to fit demand. The system will, however, be permanently optimised in order to satisfy demand as far as possible. This process takes place during the planning phase, and is referred to as Second Round Negotiation. It leads to updated TWs which is an evolution of the previous ones. This TWs set takes ANSPs’ constraints into consideration (airspace capacity, etc.). When ANSPs and airspace users have agreed on a trajectory, TWs are defined along the route, at each transfer of responsibility. These TWs also integrate ATC constraints. During the entire planning phase, until the CoO is signed, these TWs are refined to take account of known disruptions.

Finally, a Third Round Negotiation takes place. It is a further evolution of the current set of TWs and takes account of the constraints of the actual traffic. The outcome of this negotiation is the final CoO just before push-back from the gate.

In order to enter the negotiation phase with a better understanding of the situation and then obtain a more profitable result, it will be necessary to assign a value to the various TWs: stakeholders need to evaluate TWs according to their utility functions, or at least to rank them on a scale of importance. Using this assessment, any decision concerning possible compromises that will need to be agreed by the stakeholders will be correctly examined. This is particularly true for aircraft operators and ANSPs. The former, in fact, must know which TWs are of strategic importance, and which can be modified. The latter may be requested to increase capacity, which may either be impossible for safety reasons or may involve a certain cost. In the first case a priority criterion is required to balance demand and capacity, in the second case the cost must be compared with the value of the new TWs made available in order to perform an economic assessment. As far as airports are concerned, this evaluation of TWs will probably be substituted by that for airport slots. For coordinated and schedule-facilitated airports these can be considered as coincident. In all other cases, no economic assessments need to be made; otherwise airport slots would have been put in place.

The main implementation details for each stakeholder/hierarchical level are set out schematically in the following Table 3.

Implementation details

Implementation Details	**Stakeholder**		
Hierarchical level	Airport Operators	Aircraft Operators	ANSP
Organisational	Human Resources Training Equipment investment	Human Resources Training Equipment investment	Human Resources Training Equipment investment Organization of airspace
Long-Term Planning	First round negotiation Definition of the value of the TWs	First round negotiation Definition of the value of the TWs	
Medium-Term Planning	Second round negotiation	Second round negotiation	Second round negotiation Definition of the value of the TWs
Short-Term Planning	Third round negotiation	Third round negotiation	Third round negotiation
Execution	Fulfilment of TWs Renegotiation	Fulfilment of TWs Renegotiation	Fulfilment of TWs Renegotiation

Table 3: Implementation details- Main impacts of the CATS concept on stakeholders

If the TWs are considered as tradable goods, it will be necessary to build a suitable market mechanism, with a consequent regulation. Moreover, if some penalty or reward system is implemented in order to guarantee compliance, a procedure will need to be settled in order to strictly monitor the TW enforcement and to quantify the monetary return.

It should be noted that most of the investments needed by the implementation of the concept of operations proposed will be necessary regardless of which implementation of the SESAR guidance is chosen.

On the other hand, there are a number of constraints as regards the implementation of the proposed concept of operation.

At organisational level, various limitations concerning the available human resources and the rigidity of the stakeholders' organisation itself may arise.

At strategic and planning level, on the other hand, the constraints are more specific to each stakeholder. For example, airport operators need to deal with the limitations on the airport slots physically available, and on the current rules guiding their assignment process. On the other hand, aircraft operators need to optimise their schedule to comply with the airport slot allocation schemes in the light of demand forecasts and European rules (see EC Regulation 95/93, as variously amended). Moreover they must take into consideration their own resources, i.e. find the best aircraft available and define the best aircrew and station personnel scheduling, which is clearly constrained by actual availability and by the regulations. Finally, ANSPs have to respect the sectorisation of the air space and its consequent capacity.

The main constraints for each stakeholder/hierarchical level are set out schematically in the following Table 4.

Constraints

Constraints	**Stakeholder**		
Hierarchical level	Airport operators/slot coordinator/airport operator	Aircraft operators	ANSP
Organisational	Human resources Rigidity of the organisation	Human resources Rigidity of the organisation	Human resources Rigidity of the organisation
Long-Term Planning	Airport slots Flight demand forecast	Airport slots Airlines' own resources	Flight demand forecast
Medium- and Short-Term Planning	Airport slots	Airport slots Aircraft and aircrew availability and scheduling	Sectorisation Capacity
Execution	Airport slots Capacity	Airport slots Aircraft and aircrew availability and scheduling	Sectorisation Capacity

Table 4: Implementation details - Constraining factors in the TW assignment process

If the TWs are considered as tradable goods, or a penalty or reward system is implemented to guarantee compliance, a number of budget constraints need to be considered by all the actors.

As previously mentioned, the aim of the CoO/TW concept is the achievement of an agreement on the Contract of Objectives, reached through a Collaborative Decision-Making (CDM) process. This can be translated into a consensus on a set of TWs.

Moreover, it is possible to assign a value to the TWs, allowing the development of a market in which they are traded. On the other hand, a reward and penalty system may be implemented to ensure compliance with the TWs.

In order to portray the various factors likely to influence the benefits assessment and the relationship between them, an influence diagram is shown. It represents the main structural features of a situation and the key relationships existing between them. It presents an overview of the elements involved, and their main interrelationships.

The purposes of the diagram are to:

- Understand the various forces contributing to the outcome values from the variables in the decision.
- Help dissect a complex problem into its parts.
- Gain insights into the problem or opportunity.

- Identify constraints and dependencies.
- Help in developing a sound decision matrix.
- Aid in building an accurate cost model.

The symbols used for the various items in the Influence Diagram are:

- Oval or Circle – Uncertain Variable or Element
- Rectangle or Square – Decision
- Rounded Corner Rectangle – Terminal (Ultimate) Outcome Value
- Double or Shadowed Oval – Deterministic calculation
- Arrow – Arrowhead indicates the item acted upon, or influenced, by the source item

In the current analysis, the ultimate outcome value is represented by the Net Present Value (NPV) of the project. The NPV is the difference between the present value of cash inflows and the present value of cash outflows (see **Erreur ! Source du renvoi introuvable.**). It is influenced first of all by one-off and recurrent costs. The former can be subdivided into equipment investments and human resources training costs, and the latter into the effort devoted to evaluating the TWs, the negotiation and renegotiation costs, and the cost borne by the central authority for monitoring compliance with the TWs. Moreover, the Net Present Value of the system is impacted by several improvements allowed by the implementation of the concept of operations proposed, i.e. the aircraft scheduling buffer reduction, the improved use of airspace capacity, the increased predictability and flexibility, the interoperability of the ATC methods, and the increase in airport capacity.

If TWs are considered as tradable goods, stakeholders have the option of entering the market. If they do so, certain recurring costs are involved in the trading procedure, and some income or loss must be counted in the Net Present Value.

Finally, if a penalty or reward system is implemented to guarantee compliance with the TWs, stakeholders may choose whether or not to comply with them, and thus whether to incur the penalty or reward. If such a system is set up, certain expenditure on quantifying and managing these penalties and rewards will need to be taken into account.

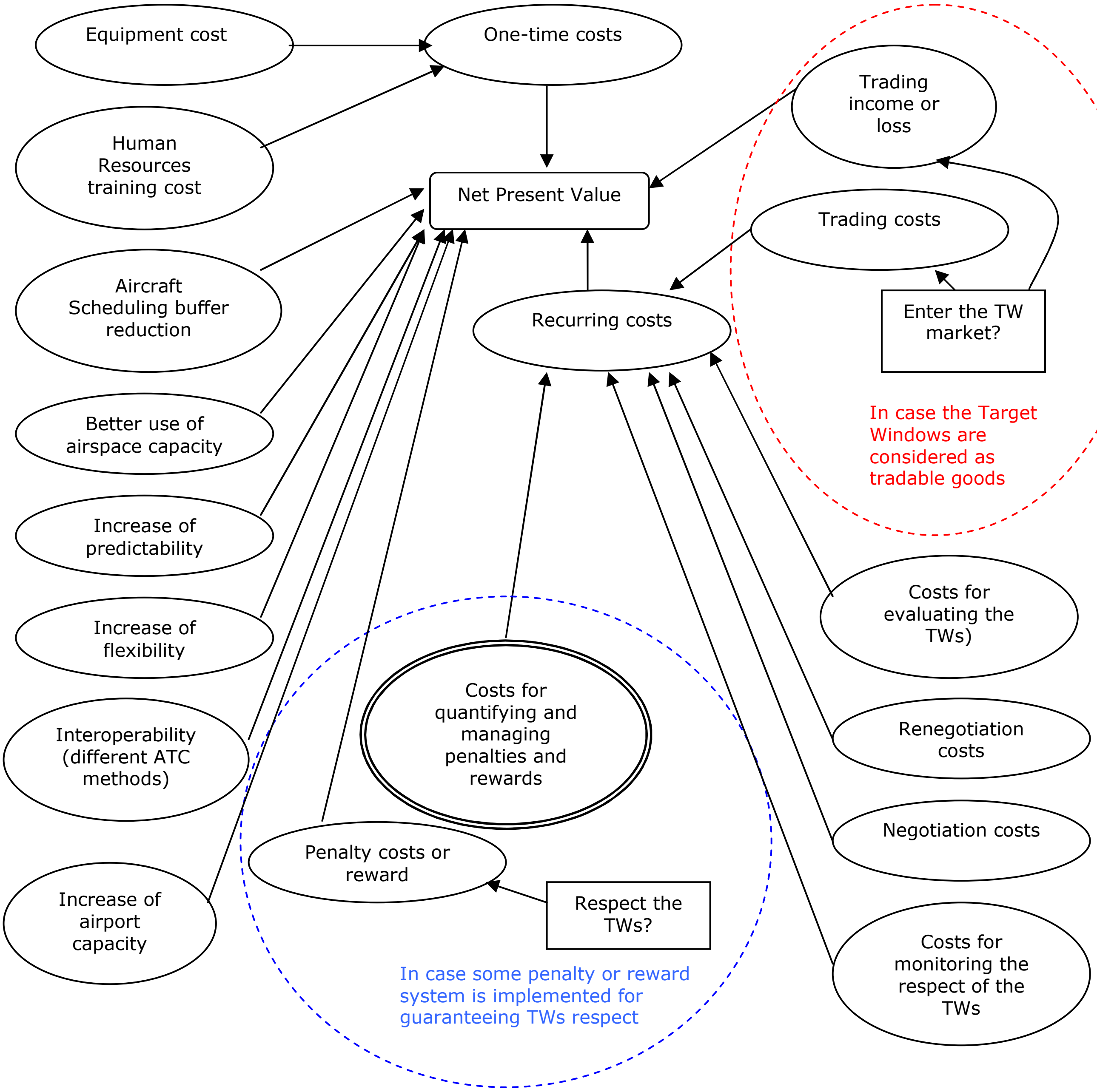

Figure 10: Influence Diagram

7.2.3 Legal implications

The CoO amounts to a real paradigm shift, as presented in §3.7. WP2.2.3 will determine the field of application and the related substantive provisions regarding the implementation of the CoO;

7.2.3.1 Objectives of the study

7.2.3.1.1 Identification of the stakeholders most involved

The first objective is to identify the stakeholders most directly involved in the process of drawing up a Contract of Objectives. They include, in the first place:

- operators of the air navigation services provider (ATCOs);
- operators of aircraft (airlines);
- aircraft crew, principally the aircraft commander;
- operators of airports.

7.2.3.1.2 Establishment of the role and functions of other parties

In the second instance, attention must also be paid to the role and functions of:

- Regulatory authorities, viz. governments and international organisations, principally the EC through its SES and SESAR programmes and agendas, whereas EASA may receive ATM-related tasks in the future; EUROCONTROL, because of its close involvement with air traffic management in Europe, and ICAO as the worldwide law-making body in various matters affecting the establishment of a CoO.
- The role and functions of non-public entities such as CANSO, IATA and ACI.
- Passengers, as their rights under a contract with the airline will be affected by the establishment of a CoO, for instance when it comes to defining liability for delay, which liability is governed by international law – see the aforementioned Montreal Agreement of 1999 on air carrier liability, as implemented in the EC through EC Regulation 889/2002, and EC Regulation 261/2004 on the compensation of passengers in cases of delay, denied boarding and cancellation of flights.

7.2.3.1.3 Establishment of responsibilities

The second objective would be to identify the current responsibilities and liabilities of each of the stakeholders (mentioned above) under the prevailing legal regimes. The regimes are established by the regulatory authorities referred to under 7.2.3.1.2 above.

As a corollary, the relationship between the responsibilities and liabilities of stakeholders and the public aviation law regime, set up by the governmental authorities identified under 7.2.3.1.2, will be assessed in order to avoid a conflict between the provisions laid down in the CoO and the current legal regime. As the CoO is not restricted to European actors, the implications of this exercise under international law will also be addressed.

7.2.3.2 Assumptions

It is assumed that contracts can be concluded between the various actors (see 7.2.3.1.1). As stated above, their current relations are basically governed by public rather than private law, including contractual arrangements. Hence, the legal framework must be organised or re-organised in such a fashion that public international law and public air law make room for the establishment of contractual arrangements, or at least do not forbid such arrangements.

It is also assumed that conflicts with current regulations, whether at a global, European or national level, must be avoided as it may appear difficult to change the law (*lex dura*), especially so when third-party – that is, non-European – interests are involved with the formulation of provisions leading to the establishment of a CoO. As stated above, non-European airlines fly from, to and via the EC and European airspace under international, that is, mostly bilateral, agreements concluded between States. Those agreements have the status of treaties under international law.

7.2.3.3 Methodology (approach/assessment)

The following steps will be taken in order to arrive at conclusions which will underpin the feasibility of and conditions for the establishment of a CoO:

- Identification of stakeholders, including their roles, tasks and responsibilities under public law;
- Analysis of existing contracts between stakeholders in so far as they are available – see, for instance, the "New Zealand" model contract between airlines and ATCOs, and service providers in Galileo;
- Carrying out of research on the current legal regime, with special reference to the questions referred to above; for instance, attention will be paid to the relationship between the aircraft commander and air traffic controller laid down in ICAO Annex 6 on the *Rules of the Air*;
- Case law regarding the above matters – see for instance, the Überlingen case ruled on in 2006 by the District Court of Konstanz, Germany (under appeal), and case law established in the US and other European countries;
- Attendance of meetings and communications with CATS partners and exchange of information on progress reached by those other partners;
- Examination of studies and publications touching on the above questions;
- Establishment of a number of conclusions.

The following table lists the targets of the CoO, and relates them to the legal provisions which may affect them. A translation is provided from 'CATS' language into legal language. The challenge will be to determine the compatibility between the two areas, that is, the objectives of the CoO under CATS with the provisions of the legal framework as identified below.

CONCEPT	WHERE IN CATS	WHAT IN CATS	WHAT IN LAW	WHERE IN LAW
CONTRACT OF OBJECTIVES COO	GENERAL (E.G. INTERNAL REPORT D1.2.1 – OPERATIONAL SOLUTIONS 4.1 – PAGE 4-1)	A 4D TRAJECTORY MANAGEMENT TOOL ASSOCIATED WITH A SINGLE FLIGHT (A FLIGHT ENVELOPE) THAT DEFINES A COMMON OPERATIONALLY REQUIRED MINIMUM SITUATION OF AWARENESS AMONG THE ACTORS, WHILE CREATING "BONDS OF RECIPROCAL RESPONSIBILITY" WITH THE AIM OF	A SERVICE AGREEMENT BETWEEN SEVERAL PARTIES, WHERE ALL THE PARTIES ARE BOUND TO PERFORM THEIR OBLIGATIONS IN ORDER TO ACHIEVE A COMMON DESIRED RESULT AS ESTABLISHED THEREIN.	NAT. LAW

		GUARANTEEING A CERTAIN RESULT: PUNCTUALITY AT DESTINATION.		
AGREEMENT OF THE PARTIES	GENERAL	A STEP IN THE NOP	AN ACT WITH LEGAL RELEVANCE IN THE PROCESS OF NEGOTIATING A CONTRACT (E.G. *CULPA IN CONTRAHENDO*); IT MAY EVEN CONSTITUTE THE CONTRACT ITSELF DEPENDING ON THE FORM THAT SAID CONTRACT SHOULD TAKE.	NAT. LAW
PLANNING PHASES	GENERAL (E.G. INTERNAL REPORT D1.2.1 – FACILITATING THE BUSINESS TRAJECTORY 4.2.4 – PAGE 4-8 *ET SEQ.*)	A NETWORK PLANNING PROCESS: LONG-TERM: FROM YEARS TO 6 MONTHS BEFORE EXECUTION; MEDIUM-TERM: FROM 6 MONTHS TO DAYS BEFORE EXECUTION); SHORT-TERM: FROM DAYS TO HOURS/MINUTES BEFORE EXECUTION	NEGOTIATION PHASE OF THE CONTRACT.	NAT. LAW (AS COMPLEMENTED OR HARMONISED BY COMMUNITY LAW)
EXECUTION	GENERAL (E.G. INTERNAL REPORT D1.2.1 – EXECUTION OF A FLIGHT THROUGH A COO 4.2.4.3 – PAGE 4-13 AND REPORT D1.3.1 CATS VALIDATION STRATEGY – OPERATIONAL DESCRIPTION 3.2 – PAGE 3-2)	EXECUTION OF THE FLIGHT IN ACCORDANCE WITH THE RBT: EXECUTED: FLOWN PORTION OF THE RBT; AUTHORISED: PORTION OF THE RBT GRANTED WITH A CLEARANCE TO FLY OVER; AGREED: REMAINING PORTION OF THE RBT NOT YET CLEARED. AND TWS AGREED: NOMINAL STATE (A/C	PERFORMANCE OF THE CONTRACT.	NAT. LAW (AS COMPLEMENTED OR HARMONISED BY COMMUNITY LAW)

		AT THE CORE OF TW); EDGE STATE: A/C INSIDE TW BUT CLOSE TO BOUNDARIES; OUTSIDE STATE: A/C OUTSIDE TW – NEED FOR RENEGOTIATION.		
RENEGOTIATION	GENERAL (E.G. INTERNAL REPORT D1.2.1 – 4.2.4.5 – PAGE 4-15 AND REPORT D1.3.1 CATS VALIDATION STRATEGY – OPERATIONAL DESCRIPTION 3.2 – PAGE 3-2)	A SPECIFIC DECISION PROCESS AT A STRATEGIC LEVEL THAT CAN BE TRIGGERED BY ANY OF THE ACTORS AND AS A RESULT OF SEVERAL EVENTS, LEADING TO THE PRODUCTION OF A NEW RBT AND NEW TWS.	MODIFICATION OF THE CONTRACT; IT MAY ALSO LEAD TO THE TERMINATION OF THE CONTRACT IN FORCE AND THE COMING INTO FORCE OF A NEW ONE.	NAT. LAW (AS COMPLEMENTED OR HARMONISED BY COMMUNITY LAW)
DEPARTURE	GENERAL	AN OPERATIONAL EVENT RELATING TO THE FLIGHT.	APART FROM CONSTITUTING AN OPERATIONAL EVENT RELATING TO THE FLIGHT, THEY NEED TO BE LEGALLY DEFINED AND FACTUALLY FRAMED, FOR THEY MAY CONSTITUTE MILESTONES IN THE PERFORMANCE AND FULFILMENT OF THE CONTRACT (E.G. *EN-ROUTE* CONCEPT). MOREOVER, THEY ARE ALSO EVENTS WITH LEGAL RELEVANCE REGARDING THE RESPONSIBILITY AND ENTITLEMENTS OF THE ACTORS AND EXPOSURE TO LIABILITY UNDER NATIONAL AND INTERNATIONAL LAW (E.G. GROUND ACTORS).	N/A
ARRIVAL				
PUNCTUALITY *AT DESTINATION*	GENERAL (E.G. INTERNAL REPORT D1.2.1 –	ARRIVING ON TIME	LEGALLY, IT DOES NOT NECESSARILY MEAN NON-EXISTENT	SEE NEXT LINE

	Problem identification 3.1.5 –page 3-3)		delay, for delay may occur only at origin (departure delay) and still have legal relevance and give rise to liability.	
Delay	General	Lack of punctuality upon arrival.	Several possibilities depending on the source of law, but generally means a time differential between the time agreed and the actual time performed. (*Please note* that in the contract of carriage, delay leads only to the non-timely execution of the contract but never to its non-performance). Conversely, regarding the CoO, delay may lead to the non-fulfilment of the contract, depending on its *time extension* and on where it occurs, e.g. a TW or upon arrival).	Please note that *delay* here is a concept relating to the contract of carriage and therefore, the primary potential liable entity will be the carrier, irrespective of possible recourse action by the latter against other actors (see e.g. recital 8 and art. 13 of Reg. 261/2004). Art. 16 of Reg. 261/2004 refers to delay as a time "beyond [a flight] scheduled time of departure", using time units (in combination with distances) in order to apportion due compensation. In this case, regardless of the time of arrival, delay is taken into account with reference to the departure time. Neither the WC (original and as amended) – art. 19 - nor the MC – art. 19 - define delay. This legal concept is left for the nat. courts to determine and to be subsumed in accordance with the applicable int. or EC law.
Constraints	General	Elements to take	Facts with legal	N/A

	(E.G. INTERNAL REPORT D1.2.1 – EXECUTIVE SUMMARY –PAGE VII)	INTO ACCOUNT DURING THE NEGOTIATION, EXECUTION AND RENEGOTIATION PHASE THAT MAY LEAD TO A CHANGE IN THE INITIAL REQUEST (BDT).	RELEVANCE FOR THE ESTABLISHMENT OF THE PARTIES' INTERESTS AND THE SCOPE OF THE CONTRACT.	
DISRUPTIVE FACTORS	GENERAL (E.G. INTERNAL REPORT D1.2.1 – PROBLEM IDENTIFICATION 3.1.2 –PAGE 3-2)	UNEXPECTED EVENTS THAT CAUSE UNCERTAINTY FOR THE MANAGEMENT TOOL (COO): AD-HOC: E.G. WEATHER (EXOGENOUS), SUDDEN LIMITATION OF RW CAPACITY; AIRCRAFT FAILURE (ENDOGENOUS); CONSTANT IMPRECISION: INACCURACY OF TECHNOLOGY; SYSTEM-WIDE: GENERATED OVER THE INTERFACES OF THE VARIOUS ACTORS.	CAUSES (FACTS) THAT MAY LEAD TO INEXACT FULFILMENT OR IMPOSSIBILITY OF PERFORMANCE OF THE CONTRACT. THESE CAUSES PLAY A PARAMOUNT ROLE REGARDING THE SUBSEQUENT ANALYSIS FOR THE APPORTIONMENT OF LIABILITY IN THE CASES MENTIONED ABOVE. THEY MAY MAKE THE DIFFERENCE BETWEEN EXONERATION AND EFFECTIVE LIABILITY.	REGARDING THE COO: THESE CAUSES/FACTS NEED TO BE EVALUATED BY THE COURTS IN THE LIGHT OF THE CONTRACTUAL OBLIGATIONS OF EACH OF THE ACTORS AS ESTABLISHED IN THE COO, IN ACCORDANCE WITH THE APPLICABLE LAW, IN ORDER TO ASSESS WHETHER THEY WILL RESULT IN CONTRACTUAL LIABILITY FOR ANY OF THE ACTORS. REGARDING THE CONTRACT OF CARRIAGE, THESE CAUSES/FACTS WILL NEED TO BE SUBSUMED UNDER ARTS. 20, 21 AND 25 OF THE WC OR ART. 20 AND 21 OF THE MC. THESE PROVISIONS PERTAIN ONLY TO THE CARRIER'S LIABILITY. IF ANY OTHER ACTOR IS THE ORIGIN, CAUSE OF OR IN ANY WAY LIABLE (BY LAW OR CONTRACT) FOR THE FACTS THAT LEAD TO DELAY, THE CARRIER WILL HAVE TO LEGALLY ACT AGAINST HIM/IT IN A RECOURSE ACTION.
RESPONSIBILITY *BETWEEN THE ACTORS*	GENERAL (E.G. INTERNAL REPORT D1.2.1 - ROLES AND RESPONSIBILITIES 4.2.5 – PAGE 4-17)	TASKS AND DUTIES OF EACH OF THE ACTORS OF THE COO; *EXPLICIT SERVICE RELATIONSHIPS*.	CONTRACTUAL OBLIGATIONS OF EACH ACTOR IN ORDER TO ENSURE THE PROPER AND PERFECT PERFORMANCE OF	INTERNATIONAL LAW STANDARDS OF EUROCONTROL (ESARRS) EU LAW

			the contract. Legal duties of the actors as stated in international and national law.	Nat. law Contracts in force (e.g. CoO or SLA or LoA)
liability	General (e.g. internal report D1.2.1 – Liability and Accountability 4.2. 7 – page 4-25)	has a legal, economic and operational aspect.	A legal concept relating to the apportionment of pecuniary damages to an entity as a result of its conduct (leading to the obligation to compensate, if proved to exist).	International Law EU Law Nat. law (tort and/or contract law: CoO; SLA; LoA)
accountability	General (e.g. internal report D1.2.1 - Liability and Accountability 4.2. 7 – page 4-25)			

Table 5 : Concept table

8 Setting the strategy

The E-OCVM provides a life-cycle view of the validation process, where the system engineering process runs in parallel with this validation process. The validation conclusions must drive system-engineering decisions and the validation must evaluate the system-engineering results.

To assess a concept, the E-OCVM methodology [10] proposes three main sub-processes:

- Definition of validation expectations,

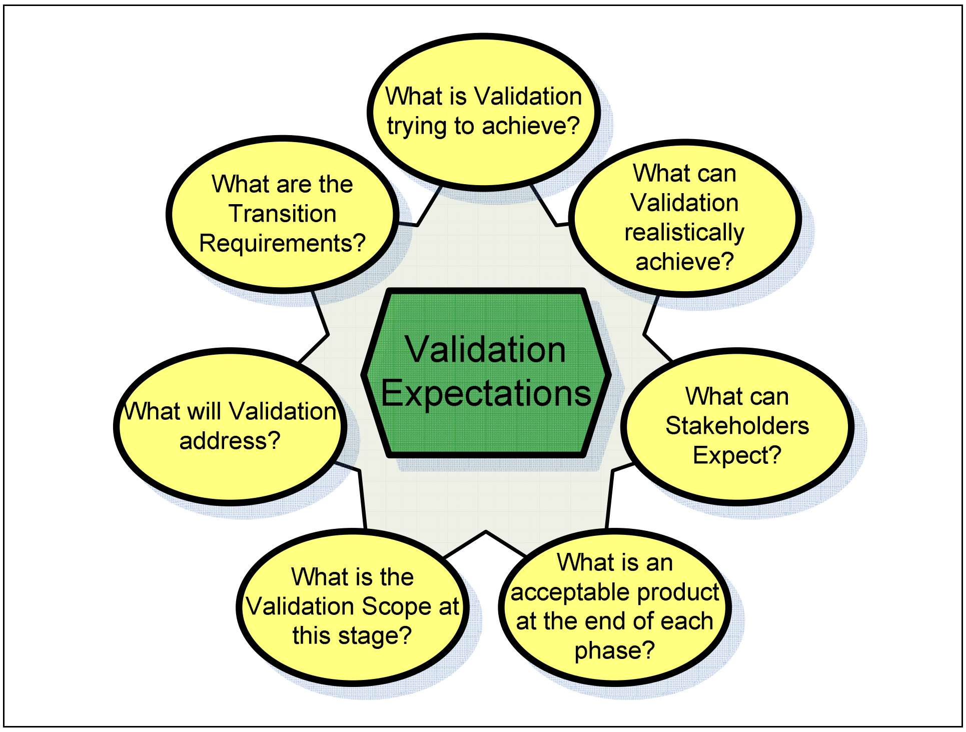

Figure 11: E-OCVM validation expectations

The objectives of the CATS assessments are defined Chapter 6

- Definition of validation requirements,

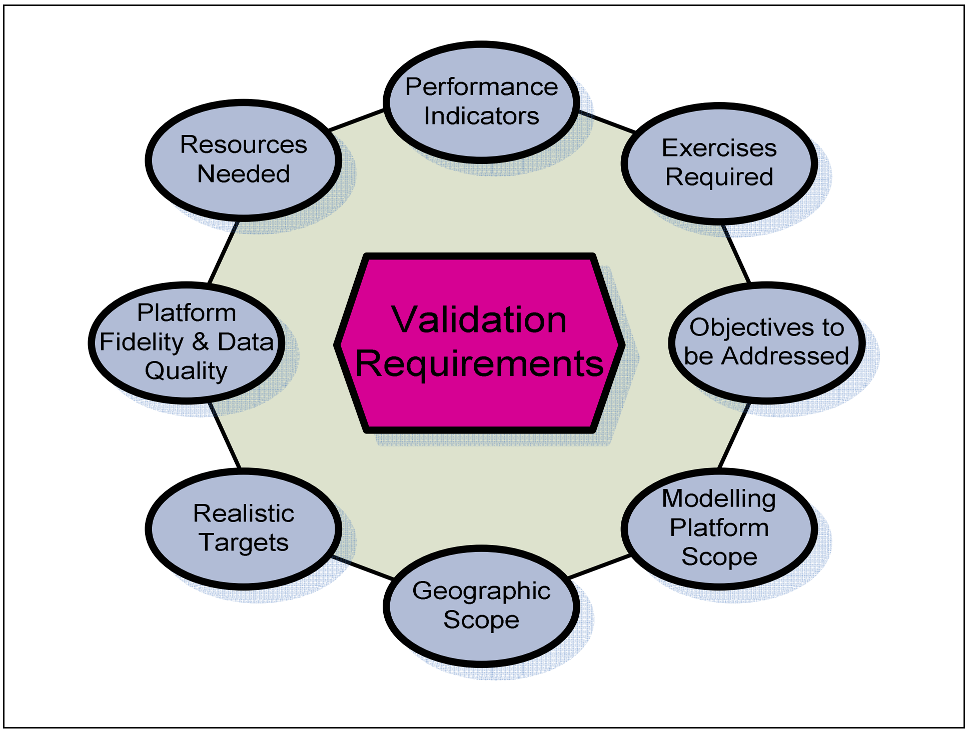

Figure 12: E-OCVM validation requirements

- and then performance and analysis.

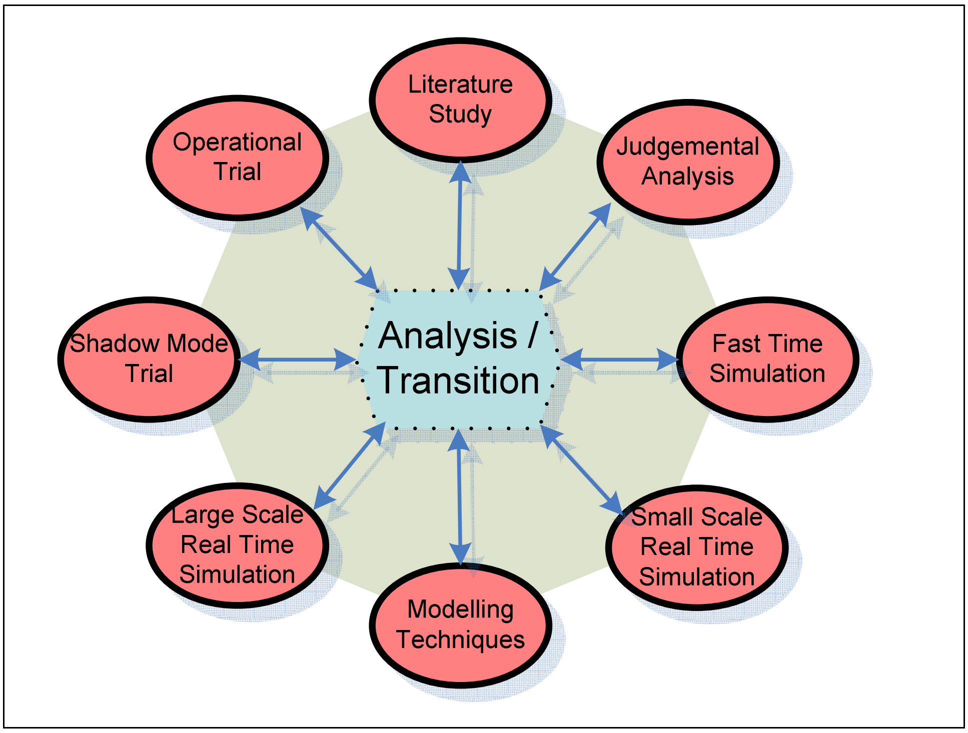

Figure 13: E-OCVM analysis

The validation plan presented in Chapter 9 covers the last two sub-processes. It describes in detail not only the objectives addressed, but also how to perform the validation and the corresponding schedule and associated resources.

Building on the idea of value, the validation and verification should increase confidence that the CoO is fit for purpose (Figure 11). A performance-based approach will bring this confidence, as the assessment will be against performance expectations (as described in 7.1.3).

Furthermore the validation and verification expected by SESAR [23] should be iterative, small-scale and close to the operational environment. This will be the case throughout the operational assessment (WP2.1), performed by way of three iterative small HILs, and developed in a real operational environment used in SkyGuide. The systemic assessment (WP2.2.2) will ensure that the project will seek benefits..

8.1 Experimental operational settings

The assessment provided for in WP2.1 [2] will focus on the impact of the CoO regarding the various actors involved in the flight execution.

It is proposed to evaluate the operational acceptability of the CoO and associated TW concepts from the point of view of the various actors and to obtain feedback from the stakeholders through HIL experiments. Three human-in-the-loop experiments will be carried out to assess the impact of the proposed concept on operators, considering both system and human performance. For the third experiment, it will be assumed that for some reason, the CoO cannot be fulfilled and needs to be renegotiated. The third experiment will then involve all actors, i.e. airports, airlines and ANSPs in a gaming exercise to evaluate strategic decisions. The output of this collaborative process will be re-introduced into the HIL experiment to confront operators with this decision.

The assessment activities will be enabled by using a rapid prototyping platform. The current Skyguide simulator will be the basis for the three HIL experiments. This simulator includes all the current Skyguide ATC centres functionalities (including MTCD) and a number of tools, e.g. to manage and display the TWs.

The simulated airspace will involve two ACCs, as described in the concept [8], mainly cross-border operations between Swiss and Italian airspace, as this simulation will look specifically at the en-route environment. The simulation airspace will be based on the border area between Geneva and Rome, with appropriate feed sectors, as presented below.

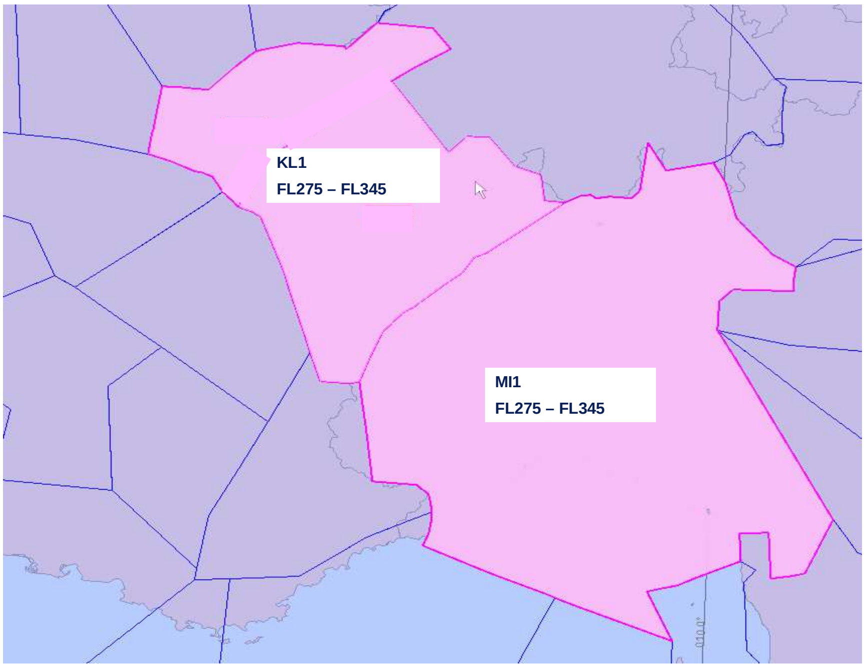

Figure 14: Airspace

ACC	Sector name	Min FL	Max FL
LSAG	LSAGKL1	FL275	FL345
LIRR	LIRRMI1	FL275	FL345

Table 6: Sectors

The Geneva and Milan sectors have been regrouped into a larger sector, in order to be compliant with SESAR expectations and also to facilitate TW adherence.

The traffic samples used in the human-in-the-loop simulations are derived from the traffic samples used in the FASTI simulations. Actual traffic in July 2006 was recorded and modified to meet the SESAR 2020 objectives (real traffic samples, increased by STATFOR). Training samples will also be created. Some non-nominal scenarios will also be taken into account. Traffic samples and scenarios will be further described in Deliverable D1.3.3 [12] and D1.3.4 [21].

8.2 Systemic validation settings

The assessment provided for in WP2.2.1 will focus on the safety implications of the TW in terms of conflict probability and renegotiation probability. The systemic study will be based on a fast-time simulation of an air traffic scenario of a scale comparable to the HIL experiments. The model used will be probabilistic and Monte-Carlo simulations will be employed to estimate the conflict and renegotiation probabilities. Conflict resolution algorithms will be developed to:

- exploit the flexibility provided by the TW for carrying out conflict resolution manoeuvres
- provide renegotiation suggestions for the TWs in the event of conflict resolution, while respecting their constraints, if possible

The impact of these algorithms on conflict probability and renegotiation probability will be assessed. Finally, based on feedback from the first HIL experiment, the preferences of controllers will be introduced as a bias in the conflict resolution algorithms and an assessment including this additional HIL consideration will be carried out.

The assessment provided for in WP2.2.2 follows the guidelines introduced in the E-OCVM [10] and CAATS II for building a Business Case [19][20] and aims to evaluate the positive and negative impacts of the implementation of the Contract of Objective on the various ATM stakeholders. The focus is on the ATM planning and flight execution phases.. The weightings of the various factors are determined through techniques based on expert judgments [19][20]. WP2.2.2 also quantitatively assesses the economic feasibility of a negotiation mechanism to define the CoO using mathematical programming techniques (see 5.3.2.3). The time frame is from the Initial Shared Business Trajectory to the last Reference Business Trajectory when the CoO is agreed (see 10.1.3).
The assessment provided for in WP2.2.3 relates to a legal analysis of the Contract of Objectives. As the CoO relates closely to the responsibility of the actors to provide a service and adhere to the agreed objectives, set out in the TWs, the legal framework for the various actors needed to be carefully evaluated. This is the main goal of the specific study performed in WP2.2.3.

Currently, relations between the various actors are governed by public law, that is, global, multilateral conventions and annexes thereto (Chicago; ICAO); regional arrangements (whether made by EUROCONTROL or in the context of the Single European Sky) and national law.

There are also private-law conventions and arrangements, made at a global or regional, i.e. European level, impacting on for instance questions pertaining to liability for damage to airlines in cases of delay and cancellations of flights, taking into account the "extraordinary circumstances" defence which airlines may invoke.

8.3 Validation schedule and resources

As described in the DOW [2] three HIL experiments are planned for the operational assessment, and three different studies will focus on benefits, safety and risk, and legal assessments.

The three HILs will be performed at the Skyguide premises, on a rapid prototyping and off-the-shelf platform (Skyguide simulation platform). This platform was also used for a FASTI experiment involving cross-border operations between Swiss and Italian airspace.

EEC, ENAV and SkySoft will be in charge of the preparation of the experiments. HIL1 will be led by ENAV, HIL2 by AFC and HIL3 by EEC.

ENAV will deploy the ATCOs for the experiments, while AFC will provide the aircrew.

SkySoft will implement the simulation environment.

EEC will be in charge of the analysis of the three HIL results.

FRQ and SkySoft will implement a workflow engine to evaluate the strategic decision-making process during renegotiation of the CoO. This will be done first with gaming exercises and the information obtained will be re-introduced into the HIL3 platform to evaluate the impact of this renegotiation.

For the systemic assessment, AFC will lead WP2.2 and mainly the three independent studies:

- The safety study of WP2.2.1 will be carried out by ETH on a fast-time simulator. Input is anticipated from WP1.3 (in terms of the scenario developed for the HIL experiment), WP2.4 (in terms of the TW modelling) and, for the final study, from WP2.1 (in terms of the preferences and adaptations of controllers with respect to the TW concept).
- The cost benefits analysis will be performed by ORTS, at three hierarchical levels (strategic, organisational and operational).
- The legal assessment will be performed by IIASL.

The link between the systemic assessment, the operational assessments and the concept description was continuously monitored by the Validation Manager to ensure coherence and consistency and a common understanding of the concept elements.

For the safety and risk analysis, the scenario and traffic used to perform the various modelling were extracted from the HIL1 and HIL2 scenarios. This was not the case for the benefits assessment, as the approach was more global (from the planning to the execution phases) than for the HIL experiments, restricted to the en-route area during flight execution. Some of the outputs of the safety and risk analysis were also taken as a basis for the Business Assessment, i.e. WP2.2.1 found that TWs are easy to hit in space and not always in time; this information was the basis of the last study of WP2.2.2 working on the temporal dimension of the target windows.

Even at this stage of the CLM, the CATS Consortium attached importance to the identification and review of the rules and regulations impacted by the proposed concept. Some of the legal analysis results feed back into the refinement of the concept and scenarios (e.g. responsibility for the fulfilment of TWs was given to the crew in order to respect the current regulations).

Name of the assessment	Deliverable Number	WP	Delivery date
Execution of the HIL1 experiment		WP2.1.1	20/31-Oct-2008
HIL 1 primary results analysis	D2.1.1	WP2.1.1	09-Jan-2009
Report on the selection of safety-critical and risk-sensitive case studies	D2.2.1.1	WP2.2.1	28-Feb-2009
Report on Legal Assessment	D2.2.3	WP2.2.3	04-May-2009
Execution of the HIL2 experiment		WP2.1.2	- Oct-2009
HIL 2 primary results analysis	D2.1.2	WP2.1.2	15-Jan-2010
Report on modelling and quantitative risk analysis	D2.2.1.2	WP2.2.1	15-Mar-2010
Execution of HIL3 experiment		WP2.1.3	-June-2010
Report on the HIL effects	D2.2.1.3	WP2.2.1	02-Aug-2010
Report on cost-benefit analysis	D2.2.2	WP2.2.2	02-Aug-2010

Name of the assessment	Deliverable Number	WP	Delivery date
HIL 3 preliminary results analysis	D2.1.3	WP2.1.3	02-Sept-2010

9 Operational validation plan

To measure the impact of a concept, showing the ATM community and the main stakeholders that benefits are being delivered as planned, focus was given on the impact of this concept on system performance, expressed by KPAs. Then the issue will be to show how various enablers such as humans will interact with the concept.

9.1 First validation experiment (HIL1)

The first validation activity provided for in the DOW [2] is HIL1. This HIL experiment will be prepared in WP1.3 and performed in WP2.1.1.

The objective is to assess the impact of the CoO between ATCOs.

As presented in 7.1, the aim of the validation is twofold (Figure 15):

- Assess whether the benefits for system performance are being delivered as proposed.
- Assess whether the contribution of human beings to overall system performance is within the expected capabilities (workload, stress, SA, etc.).

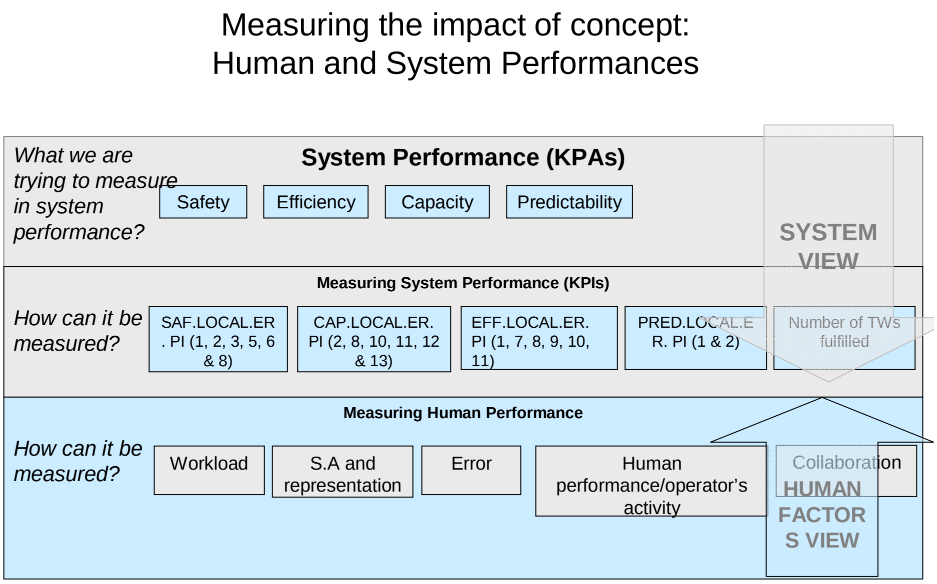

Figure 15: Human and System Performance

Table 7 below represents the various KPAs and KPIs which will be used to assess system performance, and how human performance will interact with this.

<table>
<tr><th>HIL 1</th><th>Assess the impact of the CoO between ATCOs</th></tr>
<tr><th>Objectives relating to system performance</th><th>Objectives relating to human performance</th></tr>
<tr><td>SAF</td><td rowspan="4">G1.1 : Feasibility and acceptability of the ATCOs' working methods as a result of the CoO execution
G2.1 : Impact of CoO execution on ATCOs' performance
G3.1 : Impact of the CoO on ATCOs' activity</td></tr>
<tr><td>CAP</td></tr>
<tr><td>EFF</td></tr>
<tr><td>PRED</td></tr>
<tr><th>Indicators</th><th>Indicators</th></tr>
<tr><td>SAF.LOCAL.ER. PI (1, 2, 3, 5, 6 & 8)</td><td rowspan="5">Workload: ISA, NASA-TLX, Interviews, Observations, Performance outcomes, Questionnaire
Situation Awareness: SASHA_Q, Interviews, Observations, Performance outcomes, Questionnaire
Error production and management: Observations, Questionnaire, Interviews, Performance outcomes
Operator's activity: Cognitive processes, Decision-making, Risk management, Constraints, etc.
Collaborative activity: Communications (number, time, content, speaker and receiver, etc.)</td></tr>
<tr><td>CAP.LOCAL.ER. PI (2, 8, 10, 11, 12 & 13)</td></tr>
<tr><td>EFF.LOCAL.ER. PI (1, 7, 8, 9, 10, 11)</td></tr>
<tr><td>PRED.LOCAL.ER. PI (1 & 2)</td></tr>
<tr><td>Number of TWs fulfilled</td></tr>
</table>

Table 7: HIL1 Validation Plan

9.1.1 HIL1 experimental plan

The experimental plan has been designed with the relevant variables to comply with the HIL1 validation plan presented in Table 7.

Variables are properties or characteristics of some event, object, or person that can take on different values or amounts.

An experimental plan uses two types of variables:

- Dependent variables
- Independent variables

Dependent variables are those which are seen to change in response to independent variables. Independent variables are those used deliberately to cause a change in dependent variables: "if *x* is given, then *y* occurs", where *x* represents the independent variables and *y* represents the dependent variables.

Variables may have different properties. The most common properties encountered in experimental variables are:

- Discrete (possible scores of discrete points on a scale) or continuous (the scale is continuous and not made up of discrete steps).
- Qualitative (Variables that express a qualitative attribute) or quantitative (variables that are measured in terms of numbers).

9.1.1.1 Independent variables

The independent variables determine the experimental plan (i.e. the number of runs and the experimental duration)

Three independent variables will be controlled and used during the HIL1 experiment:

- Use of the Contract of Objectives and Target Windows.
- The traffic load.
- The occurrence of operational events during the scenario.

9.1.1.1.1 The Contract of Objectives

The use of the Contract of Objectives at the ATCO working positions is the first goal of HIL1. The assessment of the Target Windows impact on ATCO and traffic performance requires scenarios to be tested with and without Contract of Objectives implementation.

In the HIL1 experiment, each scenario will be assessed with and without Target Windows.

9.1.1.1.2 The traffic load

All projects involved in the SESAR setting seek to propose new ways to cope with the planned traffic growth in the future. The purpose of HIL1 is to assess the impact of the Contract of Objectives in an operational environment meeting the 2020 requirements. Consequently, several traffic loads must be used in the simulation scenarios.

Two traffic load levels will be used:

- Current traffic load (in 2008).
- Expected traffic load (in 2020).

According to the EUROCONTROL forecasts for the sectors measured for the HIL1 experiment, traffic is expected to increase by 40% in 2020. Traffic growth is expressed in terms of the hourly throughput for each of the two measured sectors.

9.1.1.1.3 Conflicts

The difficulty in managing traffic depends on several features such as traffic load, but also on the number of conflicts, the nature of conflicts, the way the traffic comes into the sector, the constraints on meeting the performance objective, etc.

Several traffic scenarios will be designed and the inherent conflict situations need to be resolved. This means that the number and the nature of conflicts will be calibrated for each scenario.

The way the ATCOs manage different types of conflict with and without the Target Windows will be very informative and will show how the Contract of Objectives impacts on traffic management.

Two types of conflict will be used:

- Merging conflicts
- Catch-up conflicts.

Also, two natures of conflict will be used to assess the level of ATCO anticipation in traffic management:

- Existing conflicts when the aircraft enters the sector
- Conflicts triggered by orders given by ATCOs so as to fulfil the flight plan. This kind of conflict is more difficult to anticipate because it needs the ATCOs to assimilate the flight plan data to a considerable extent.

9.1.1.1.4 Occurrence of events

The occurrence of unexpected events in the scenario is a disruptive factor in the ATCOs' activity. Such events occur daily in the operational setting and require fast adaptation from the ATCOs to keep the traffic safe and allow for efficient air traffic management.

The purpose of having a disruptive factor in the scenarios is to assess the impact of having a Contract of Objectives on the management of this factor.

Several disruptive events may occur. They may occur for technical (ATC or aircraft failure, etc.) or environmental (weather disruptions, etc.) reasons.

For the experiment, only one event will be used and that will be a storm. The storm will occur during the heavy traffic load phase of the scenario. The purpose is to delineate the ATCOs' capacity for managing simultaneously high constraint levels (i.e. traffic and events).

9.1.1.2 Dependent variables

The dependent variables determine the range of the collected data and the analysis plan.

Two types of dependent variables will be collected:

- System performance data.
- Human performance data.

The dependent variables used for the HIL1 experiment are described in detail in 7.1.

9.1.2 HIL1 Conduct of the experiment

The experiment will take place over ten days at the SkySoft premises in Geneva, Switzerland:

- Five days from 20 to 24 October 2008.
- Five days from 27 to 31 October 2008.

These ten days will include:

- ATCO training relating to the CATS scenario (target windows management), the simulation environment and interfaces, and traffic management in the measured sectors.
- Conducting of the measured experimental simulations.

The experiment will be performed with operational air traffic controllers in active service at Rome and Brindisi-Italy En-Route ACC.

9.1.2.1 Scenario design

Scenarios must meet the experimental hypothesis, as defined in 7.1, and allow the study and analysis of the independent variables.

To cope with the fixed experiment duration and the statistical requirements, the following choices were made by the CATS team:

- A minimum of 8 measurements for each experimental condition is needed to satisfy statistical analysis.
- The length of the measured scenarios should be one hour at most.
- The "TW" independent variable will be assessed in specific scenarios.
- The "traffic load" independent variable will be assessed in a one-hour scenario, the scenario being split into two phases (the first for the current traffic load in 2008, and the second for the expected traffic load in 2020).
- The "event" independent variable will be assessed in all scenarios, and one incident will occur per scenario.

These choices result in a scenario script which encompasses the following components for each scenario:

- First phase: current traffic load in 2008.
- Second phase: expected traffic load in 2020
- Third phase: event occurring during the 2020 traffic load

An overall description of the HIL1 scenarios is given in Deliverable D1.3.3 CATS HIL1 Operational Scenario description [12].

9.1.2.2 Experiment design

Sixteen scenarios or runs will be necessary to perform the experiment:

- Eight in a "TW" condition.
- Eight in a "non-TW" condition.

Four ENAV ATCOs are available for the experiment and 16 runs will be performed. In each run, 2 sectors are measured (MI1 and KL1). For each sector, a team of ATCOs, planner (PLN) and executive (EXE), is required

The comparison between "with" and "without CoO" scenarios requires a comparison of the same pairs of ATCOs in the two sectors. Consequently eight combinations of pairs of ATCOs are necessary for the two sectors, each combination performing the "with" and the "without CoO" condition.

In addition, the traffic learning effect must be avoided by proposing various traffic patterns to the ATCOs. Because each ATCO will have different positions following the measured sector and the function on the controller working position (EXE or PLN), four different traffic samples are needed.

A synthesis of the experimental design constraints is presented in Table 8 in which the combination of ATCOs, measured sectors, traffic, and "TW" conditions is expressed.

Without CoO			With CoO		
	MI1 Sector	KL1 Sector	MI1 Sector	KL1 Sector	
Traffic 1	AB	CD	AB	CD	Traffic 4
Traffic 2	BA	DC	BA	DC	Traffic 3
Traffic 3	CD	AB	CD	AB	Traffic 2
Traffic 4	DC	BA	DC	BA	Traffic 1
Traffic 2	AC	BD	AC	BD	Traffic 3
Traffic 1	CA	DB	CA	DB	Traffic 4
Traffic 4	BD	AC	BD	AC	Traffic 1
Traffic 3	DB	CA	DB	CA	Traffic 2

- A,B,C,D: ATCOs

- AC, DB: pairs of ATCOs where the first is the EXE and the second the PLN

- 8 traffic samples to cover the 2 "TW" experimental conditions (with and without CoO)

Table 8: Combination of ATCO positions.

9.1.2.3 Experiment schedule

The experiment timetable encompasses three main periods:

- Pre-experiment period.
- Experiment period.
- Analysis period.

9.1.2.3.1 Pre-experiment period

During the pre-experiment period an overview of the HIL1 experiment will be presented to the ENAV ATCOs involved. Two main tasks will be performed during the pre-experiment period:

- A controller handbook will be produced to present the CATS concept, experiment's purposes, simulation devices, HMI, working methods, and measurements. This document will be delivered before the HIL1 preparation meeting.
- A preparation meeting will be held in Rome three weeks before the experiment. During that meeting, an oral presentation about the concept, objectives of the experiment, experiment environment and CWP device will be given. The purpose is to ensure that ATCOs have fully understood the expectations of the experiments. This meeting is also the occasion to dynamically familiarise the ATCOs with the HMI.

9.1.2.3.2 Experiment period

The experiment period will last ten days. The experiment period timetable is shown in Table 9:

- Half a day for presentation of the simulation devices: functions and limits.
- One day for familiarisation with the simulation devices, HMI, airspace, and the CoO. The aim of this training is pedagogical. It will be interspersed with explanations, comments and even on-line adjustments of the working methods. It must therefore be possible to stop the simulation. The scenarios proposed there are not operational but represent the different situations that could occur in the experiment, to show to the operators the HMI features.
- One and half days for training on operational scenarios and the experimental environment. The training will be run with the same conditions as the measured simulation (training scenarios with traffic loads and occurrences equivalent to those in the measured simulation). These training sessions will allow the experiment to be frozen, in order to give explanations and make final changes and adjustments.
- Six days for performing the 16 experimental runs with a maximum of 3 runs per day. Each run encompasses:
  - Short run presentation and briefing.
  - Run performance (one hour).
  - Completing questionnaires and performing self-evaluation scales (15 minutes)
  - Debriefing (20 minutes)
- One spare day to cope with unexpected circumstances relating to the simulation devices (i.e. breakdown, etc.) or to test different TW modelling hypotheses.
- Final debriefing with all attendees will close the HIL1 experiment period.

Day	Morning	Afternoon
20th October	Simulation devices presentation	Familiarisation with tuition scenarios (session 1)
21st October	Familiarisation with tuition scenarios (session 2)	Operational training (session 1)
22nd October	Operational training (session 2)	Operational training (session 3)
23rd October	Experimental runs #1, 2 & 3	
24th October	Experimental runs #4, 5 & 6	
27th October	Familiarisation and Training	Experimental runs #7 & 8
28th October	Experimental runs #9, 10 & 11	
29th October	Experimental runs #12, 13 & 14	
30th October	Experimental runs #15 & 16 – Final Debriefing	
31st October	Spare day –	

Table 9: Run plan for HIL1 experiment period

The scenarios are twofold:

- Training scenarios.
- Experiment/measured scenarios.

The training scenarios are built like the experiment/measured scenarios. Their aim is to develop the controller habits and skills needed when faced with the measured scenarios. The situations simulated will need to be equivalent in terms of traffic load and complexity to those in the measured scenarios. The controllers should not discover, during the measured sessions, working constraints unknown to them and for which they need to learn *in situ*. In such cases, measurements will assess learning more than usage. Controllers will clearly not be able to acquire complete and perfect mastery of new traffic organisation and management concepts with only half a day's training, but an attempt must be made to control these parameters as tightly as possible in the light of organisational, planning and staff availability constraints. Four training scenarios will be required for the half-day of training. The content of these scenarios will be determined in such a way as to be representative of the experiment/measured scenarios in terms of traffic density and occurrences.

The experiment/measured scenarios will be the situational and operational support for the controllers' comments and the measurements of their activity in order to achieve the experiment objectives. Each scenario will last approximately 60 minutes and will start with a change-over situation. An overall description of the scenario can be found in D1.3.3.[12]

9.1.3 HIL1 data analysis plan

Automatically recorded, system and human performance data will be processed by statistical methods.

Questionnaires and self-evaluation scale data will be processed by statistical methods.

Subjective data gathered via observations and verbal comments will be processed by analysis of their content. The debriefings will be recorded, but a thorough review of them is not envisaged. The recordings will be used only in the event of doubts relating to certain situations or explanations.

The qualitative and quantitative data will be collated on the basis of the phases of the scenario and the constraints encountered in the course of the scenarios.

Data will be analysed on the following basis:

- By controller, in order to find behavioural invariants between controllers, their analyses and their understanding of the scenario's various phases.
- By scenario phase, in order to search for usage invariants associated with situational constraints.

The human performance results will be analysed in relation to the system performance results in order to analyse the CoO's acceptability by the ATCOs and regarding future traffic evolution in 2020. This level of analysis is also very relevant in order to understand the advantages and deficiencies of the CoO concept from an operational point of view. The results provide promising ways for adapting the concept, assessing its compliance with the SESAR objectives, and refining the ATCOs' working methods and the HMI proposed for the following experiments. The complete analysis plan can be found in Annex A (13).

9.2 Second validation experiment (HIL2)

The second validation activity provided for in the DOW [2] is HIL2, following HIL1 (October 2008). This HIL experiment will be prepared in WP1.3 and performed in WP2.1.2.

The HIL1 experiment was the first opportunity to apply the CATS Validation Strategy. The results were satisfactory and the Validation Strategy will be quite similar for HIL2. Nevertheless, a number of limitations were identified and improvements will be applied in HIL2, as described below:

- Safety: the calculation of STCA occurrences is not satisfactory for measuring the impact of the CoO on safety, as most of them were under control. The recent progress made at the EEC on designing a new tool (EUROCONTROL SPT tool) measuring separation performance will be taken into account in HIL2.
- Predictability: It was difficult to catch up this indicator, as the CoO was designed with the main goal of improving predictability. Fulfilment of the CoO guarantees a high level of predictability. Measurement of predictability may be summed up with the measurement of fulfilment of the CoO.
- ATCO collaboration: To manage a high level of traffic, new collaborative tools for the EXE and PLN will be implemented in HIL2.
- Aircrew – ATCO communication: communication between ATCOs and aircrew was not evaluated during HIL1 due to use of the automated pseudo-pilot. This was a limiting factor in the evaluation of ATCO workload. This will be improved in HIL2 by

adding RT between ATCOs and manned cockpits and use of data-link for other flights.

Despite the remark made during the mid-term review, the validation team would reaffirm the need to use a small number of variables, to maintain confidence in the results. This is the main reason why the airspace and route structure will remain the current one. Nevertheless it should be emphasised here that the traffic level in this area has been increased to the forecast 2020 level, as given by STATFOR data (40% in the measured area).

The objective of HIL2 is to assess the impact of the CoO between ATCOs and aircrews.

As explained in 7.1, the aim of the validation is twofold (Figure 16):

- Assess whether the benefits for system performance are delivered as proposed.
- Assess whether the contribution of human beings to overall system performance is within their capabilities (workload, stress, SA, etc.).

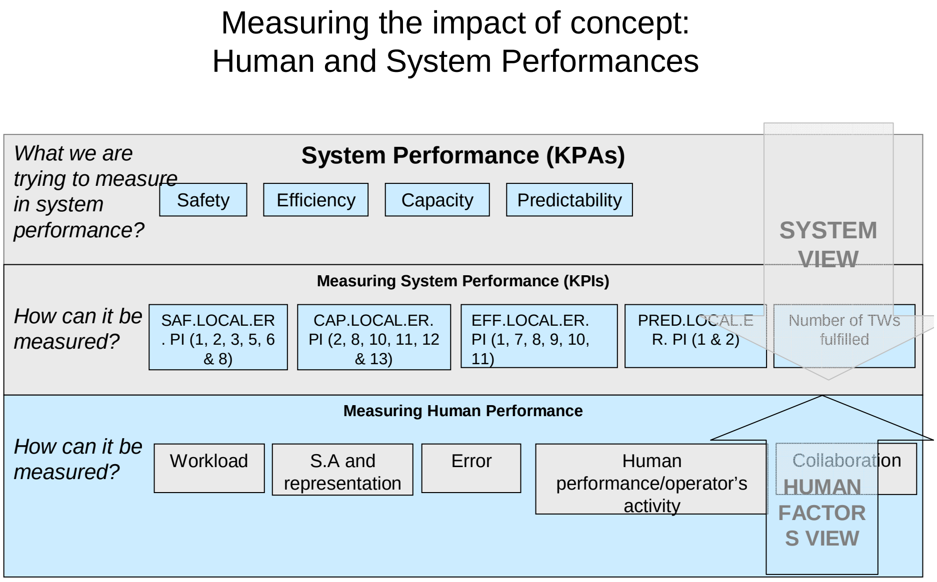

Figure 16: Human and System Performance

Table 10 below shows the various KPAs and KPIs which will be used to assess system performance, and how human performance will interact with this.

<table>
<tr><th>HIL 2</th><th>Assess the impact of the CoO between ATCOs and aircrew</th></tr>
<tr><td>Objectives relating to system performance</td><td>Objectives relating to human performance</td></tr>
<tr><td>SAF
CAP
EFF
PRED</td><td>▪ G1.1 & G1.2 : Feasibility and acceptability of the aircrew and ATCOs' working methods as a result of the CoO's execution
▪ G2.1 &G2.2 : Impact of CoO execution on aircrew and ATCO performance
▪ G3.1 & G3.2 : Impact of CoO on aircrew and ATCO activity</td></tr>
<tr><td>Indicators</td><td>Indicators</td></tr>
<tr><td>SAF.LOCAL.ER. PI (1, 2, 3, 5, 6 & 8)
CAP.LOCAL.ER. PI (2, 8, 10, 11, 12 & 13)
EFF.LOCAL.ER. PI (1, 7, 8, 9, 10, 11)
PRED.LOCAL.ER. PI (1 & 2)
Number of TWs fulfilled</td><td>▪ Workload: ISA, NASA-TLX, Interviews, Observations, Performance outcomes, Questionnaire
▪ Situation Awareness: SASHA_Q, Interviews, Observations, Performance outcomes, Questionnaire
▪ Error production and management: Observations, Questionnaire, Interviews, Performance outcomes
▪ Operator's activity: Cognitive processes, Decision-making, Risk management, Constraints, etc.
▪ Collaborative and R/T activity: Communications (number, time, content, speaker and receiver, etc.)</td></tr>
</table>

Table 10: HIL2 Validation Plan

9.2.1 HIL2 experimental plan

The experimental plan has been designed with the relevant variables to comply with the HIL2 validation plan, presented in Table 10.

Variables are properties or characteristics of some event, object, or person that can take on different values or amounts.

An experimental plan uses two types of variables:

- Dependent variables
- Independent variables

Dependent variables are those which are seen to change in response to independent variables. Independent variables are those used deliberately to cause a change in dependent variables: "if *x* is given, then *y* occurs", where *x* represents the independent variables and *y* represents the dependent variables.

Variables may have different properties. The most common properties encountered in experimental variables are:

- Discrete (possible scores of discrete points on a scale) or continuous (the scale is continuous and not made up of discrete steps).
- Qualitative (Variables that express a qualitative attribute) or quantitative (variables that are measured in terms of numbers).

9.2.1.1 Independent variables

The independent variables determine the experimental plan (i.e. the number of runs and the experimental duration)

Four independent variables will be controlled and used during the HIL2 experiment:

- Use of the Contract of Objectives and Target Windows.
- The traffic load.
- The occurrence of operational events during the scenario.
- Aircraft piloted by aircrew or by the simulation system.

9.2.1.1.1 The Contract of Objectives

The goals of HIL2 are mainly to assess the use of the Contract of Objectives by the aircrews, and the relationship between ATCOs and aircrews, to manage the target windows. Assessing the target windows' impact on ATCOs, aircrews and traffic performance requires scenarios to be tested with and without Contract of Objectives implementation. Then, in the HIL2 experiment, each scenario will be assessed with and without Target Windows.

9.2.1.1.2 The traffic load

All projects involved in the SESAR setting seek to propose new ways to cope with the planned traffic growth in the future. The purpose of HIL2 is to assess the impact of the Contract of Objectives in an operational environment meeting the 2020 requirements. Consequently, several traffic loads must be used in the simulation scenarios.

Two traffic load levels will be used:

- Current traffic load (in 2008).
- Expected traffic load (in 2020).

According to the EUROCONTROL forecasts for the sectors measured for the HIL2 experiment, traffic is expected to increase by 40% in 2020. Traffic growth is expressed in terms of the hourly flow for each of the two measured sectors.

9.2.1.1.3 Aircraft piloted by aircrew or by the simulation system

The objective of the HIL2 experiment is twofold:

- Analyse the collaboration process between controllers and aircrew regarding TW management.
- Analyse the impact of TW management on the aircrew's activity in the cockpit.

The collaboration between aircrews and ATCOs will be assessed through communications. In order to facilitate the communication process, aircrews and ATCOs do not have to cope with any language or communication device limitations. The use of professional vocal language is the best way to encourage comprehensive and fruitful exchanges. In this way, voice (RT) communication will be used between aircrews and ATCOs. Particular attention will be paid to these communications, in particular as regards TW management, which should help provide recommendations for future data-link message preparation.

However, as regards simulation limitations, not all aircraft will be piloted by aircrews. The automatic pseudo pilot, implemented in the platform, will execute the control orders for some flights. In order to avoid an excessive decrease in ATCO workload, a data link device will be implemented at the ATCOs' working position. The data link will integrate latency delays between the orders and their execution by aircraft. The latency delays will randomly vary between 10 seconds and 30 seconds.

9.2.1.1.4 Conflicts

The difficulty in managing the traffic depends on several features such as traffic load but also on the number of conflicts, the nature of conflicts, the way the traffic comes into the sector, the constraints on meeting the performance objective, etc.

Several traffic scenarios will be designed and the inherent conflict situations need to be resolved. This means that the number and the nature of conflicts will be calibrated for each scenario.

The way the ATCOs manage different types of conflict with and without the Target Windows will be very informative and will show how the Contract of Objectives impacts on traffic management.

Two types of conflict will be used:

- Merging conflicts
- Catch-up conflicts.

Also, two natures of conflict will be used to assess the level of ATCO anticipation in traffic management:

- Existing conflicts when the aircraft enters the sector
- Conflicts triggered by orders given by ATCOs so as to fulfil the flight plan. This kind of conflict is more difficult to anticipate because it needs the ATCOs to assimilate flight plan data to a considerable extent.

9.2.1.1.5 Occurrence of events

The occurrence of unexpected events in the scenario is a disruptive factor in the ATCOs' activity. Such events occur daily in the operational setting and require fast adaptation from the ATCOs to keep the traffic safe and allow for efficient air traffic management.

The purpose of having a disruptive factor in the scenarios is to assess the impact of having a Contract of Objectives on the management of this factor.

Several disruptive events may occur. They may occur for technical (ATC or aircraft failure, etc.) or environmental (weather disruptions, etc.) reasons.

As HIL2 is an exploratory experiment, only one event will occur by run. The results of the HIL1 experiment demonstrated that weather disruptions (CB and wind variation) are good disruptive factors. Disruptive factors will occur in the medium traffic load phase of the scenario. The purpose is not to saturate the ATCOs with high constraint levels (i.e. traffic and events).

9.2.1.2 Dependent variables

The dependent variables determine the range of the collected data and the analysis plan.

Two types of dependent variables will be collected:

- System performance data.
- Human performance data.

The dependent variables used for the HIL2 experiment are described in detail in 7.1.

9.2.2 Subjects

9.2.2.1 Air traffic controllers

The experiment will be performed with operational air traffic controllers in active service at Italy ENAV En-Route ACC.

At least four controllers are required. To satisfy the required profile, they should:

- Be in active operational service at an en-route control centre.
- Have experience with the measured sectors.
- Be familiar with experimental air traffic control simulation environments.
- Be aware of the problems of innovative research and the upstream study of concepts.
- Have knowledge and an understanding of the CATS study.

Knowledge of the CATS Project will be brought by the consortium in a one-day meeting which will be held before the simulation.

9.2.2.2 Aircraft pilots

The experiment will be performed with Air France professional pilots.

At least two pilots are required. To satisfy the required profile, they should:

- Be in active operational service at Air France. Pilots who are active or who retired less than one year ago are required.

- Have experience with glass cockpit aircraft in order to be familiar with and favourable to the use of modern aircraft systems and interfaces. In addition, the future of ATM, and the SESAR Project, relies on aircraft navigation aids such as the Flight Management System (FMS). The experiment has to take full account of these aspects. The aircraft simulation platform is a glass cockpit aircraft.
- Be aware of the problems of innovative research and the upstream study of concepts.
- Have knowledge and an understanding of the CATS study.

Knowledge of the CATS Project (concept and simulation environment) will be brought by the consortium in a half-day meeting which will be held before the simulation.

9.2.3 HIL2 Conduct of the experiment

The experiment will take place over ten days at the SkySoft premises in Geneva, Switzerland:

- Five days from 19 to 23 October 2009.
- Five days from 26 to 30 October 2009.

These ten days will include:

- Pilot and ATCO training relating to the CATS scenario (target windows management), the simulation environment and interfaces, and traffic management in the measured sectors.
- Conduct of the measured experimental simulations.

9.2.3.1 Scenario design

Scenarios must meet the experimental hypothesis, as defined in 7.1, and allow the study and analysis of the independent variables.

To cope with the fixed experiment duration and the statistical requirements, the following choices were made by the CATS team:

- A minimum of 8 measurements for each experimental condition is needed to satisfy statistical analysis.
- The length of the measured scenarios should be 70 minutes.
- The "TW" independent variable will be assessed in specific scenarios.
- The "traffic load" independent variable will be assessed in the 70-minute scenario, the scenario being split into two phases (the first for the current traffic load in 2008, and the second for the expected traffic load in 2020).
- The “piloted aircraft” independent variable will be assessed in each scenario.
- The "event" independent variable will be assessed in all scenarios, and one incident will occur per scenario.

These choices result in a scenario script which encompasses the following components for each scenario:

- First phase: current traffic load in 2008.

- Second phase: expected traffic load in 2020.
- Third phase: medium traffic load with one weather event (CB or wind variation).
- Piloted aircraft in each scenario. Regarding the duration of a run, the experimental team has two possibilities:
  - To have, per run and per cockpit position, 3 "short-haul" aircraft, flying for less than 20 minutes through the measured sector.
  - or to have, per run and per cockpit position, 2 "long-haul" aircraft, flying for between 20 and 35 minutes through the measured sector. These "long-haul" aircraft will make possible an in-depth assessment of the "piloted aircraft" transfer process between the two ANSPs.

An overall description of the HIL2 scenarios is given in Deliverable D1.3.4 CATS HIL2 Operational Scenario description [21].

9.2.3.2 Experiment design

Eight scenarios are required for each independent variable which is "TW" (this variable requires a specific scenario), and "Traffic load" (variable integrated in the "TW" scenarios).

The "Event" and "Conflicts" variables are integrated in the same scenarios.

The "Piloted aircraft" variable is also integrated in the same scenarios. In each scenario, the two values of the "piloted aircraft" condition ("short-haul" and "long-haul") will be assessed. Therefore, there will be 3 "short-haul" aircraft and 2 "long-haul" aircraft per scenario.

Consequently, sixteen scenarios or runs will be necessary to perform the experiment:

- Eight in a "TW" condition.
- Eight in a "non-TW" condition..

9.2.3.2.1 Controllers:

Four ENAV ATCOs are available for the experiment and 16 runs will be performed. In each run, 2 sectors are measured (MI1 and KL1). For each sector, a team of ATCOs, planner (PLN) and executive (EXE), is required

The comparison between "with" and "without CoO" scenarios requires a comparison of the same pairs of ATCOs in the two sectors. Consequently eight combinations of pairs of ATCOs are necessary for the two sectors, each combination performing the "with" and the "without CoO" condition.

In addition, the traffic learning effect must be avoided by proposing various traffic patterns to the ATCOs. Because each ATCO will have different positions following the measured sector and the function on the controller working position (EXE or PLN), four different traffic samples are needed.

A synthesis of the experimental design constraints is presented in Table 11 in which the combination of ATCOs, measured sectors, traffic, and "TW" conditions is expressed.

	Without CoO		With CoO		
	MI1 Sector	KL1 Sector	MI1 Sector	KL1 Sector	
Traffic 1	AB	CD	AB	CD	Traffic 4
Traffic 2	BA	DC	BA	DC	Traffic 3
Traffic 3	CD	AB	CD	AB	Traffic 2
Traffic 4	DC	BA	DC	BA	Traffic 1
Traffic 2	AC	BD	AC	BD	Traffic 3
Traffic 1	CA	DB	CA	DB	Traffic 4
Traffic 4	BD	AC	BD	AC	Traffic 1
Traffic 3	DB	CA	DB	CA	Traffic 2

- A,B,C,D: ATCOs

- AC, DB: pairs of ATCOs where the first is the EXE and the second the PLN

- 8 traffics samples to cover the 2 "TW" experimental conditions (with and without CoO)

Table 11: Combination of ATCO positions.

9.2.3.2.2 Pilots:

Two pilots from Air France Consulting are available for the experiment and they will fly "piloted aircraft" during the 16 runs.

TW information will be displayed on the navigational display during the "TW" runs. During the "non-TW" runs, the pilot will fly the aircraft as usual.

In each run, 2 "medium-haul" aircraft will be consecutively flown by one pilot, and 2 other aircraft will be consecutively flown by the other pilot. Each pilot will fly only one aircraft at a given time. During each run, 4 aircraft will be flown and measured.

Table 12 shows how the aircraft will be split into the various traffic samples according to whether they are "short / long-haul" aircraft, the order in which the flight will pass through the measured sectors (FL type), the traffic start time and the flight simulator position.

Day			22 Oct	22 Oct	22 Oct	23 Oct	23 Oct	23 Oct	26 Oct	26 Oct		27 Oct	27 Oct	27 Oct	28 Oct	28 Oct	28 Oct	29 Oct	29 Oct
Traffic			1	5	9	6	1	5	9	6		5	1	6	9	5	1	6	9
TW			No	Yes	No	Yes	No	Yes	No	Yes		No	Yes	No	Yes	No	Yes	No	Yes
Callsign	**FL type**	**Start time**	01	02	03	04	05	06	07	08		09	10	11	12	13	14	15	16
AZA201Y	MI->KL	XX:04:00	P1	P2	P2	P1						P2	P1	P1	P2				
AFR211Y	MI->KL	XX:04:00					P2	P1	P1	P2						P1	P2	P2	P1
SWR123X	MI->KL	XX:36:00	P1		P2		P2		P1				P1		P2		P2		P1
AFR235X	MI->KL	XX:36:00		P2		P1		P1		P2		P2		P1		P1		P2	
AFR551X	MI->KL	XX:04:00	P2	P1			P1	P2				P1	P2			P2	P1		
AFR553X	KL->MI	XX:04:00			P1	P2			P2	P1				P2	P1			P1	P2
SWR432Y	KL->MI	XX:46:00	P2		P1		P1		P2				P2		P1		P1		P2
AZA205Y	KL->MI	XX:46:00		P1		P2		P2		P1		P1		P2		P2		P1	

Table 12. Combination of aircraft and flight simulator positions in the HIL2 experiment for the pilots.

9.2.3.3 Experiment schedule

The experiment timetable encompasses three main periods:

- Pre-experiment period.
- Experiment period.
- Analysis period.

9.2.3.3.1 Pre-experiment period

The pre-experiment period encompasses the HIL2 experiment overview presented to the ATCOs and pilots involved. Two main tasks will be performed during the pre-experiment period:

- A controller handbook will be produced to present the CATS concept, experiment's purposes, simulation devices, HMI, working methods, and measurements. This document will be delivered before the HIL2 preparation meeting.
- Presentation meeting for ATCOs and pilots.
  - An ATCO presentation meeting will be held in Rome two weeks before the experiment period. This meeting will last one day, and will consist of oral presentations regarding topics, such as a concept description, the objectives of the experiment, the experiment environment and CWP devices. The purpose is to ensure that ATCOs have fully understood the expectations of the

experimenters. This meeting is also the occasion to dynamically familiarise ATCOs with the HMI.

- The pilot presentation meeting will be held in Paris three weeks before the experiment period. This meeting will last half a day, and will consist of oral presentations regarding topics, such as a concept description and the objectives of the experiment. The purpose is to ensure that pilots have fully understood the expectations of the experimenters.

9.2.3.3.2 Experiment period

The experiment period will last ten days. The experiment period timetable is shown in Table 13:

- Half a day for presentation of the simulation devices: functions and limits.
- One day for familiarisation with the simulation devices, HMI, airspace, and the CoO. The aim of this training is pedagogical. It will be interspersed with explanations, comments and even on-line adjustments of the working methods. It must therefore be possible to stop the simulation. Tuition scenarios will be used (small scenarios to present only the HMI features).
- One and half days for training on operational scenarios and the experimental environment. The training will be run with the same conditions as the measured simulation (training scenarios with traffic loads and occurrences equivalent to those in the measured simulation). These training sessions will allow the experiment to be frozen, in order to give explanations and make final changes and adjustments.
- Six days for performing the 16 experimental runs with a maximum of 3 runs per day. Each run encompasses:
    - Short run presentation and briefing.
    - Run performance (70 minutes).
    - Completing questionnaires and performing self-evaluation scales (15 minutes)
    - Debriefing (20 minutes)
- One spare day to cope with unexpected circumstances relating to the simulation devices (i.e. breakdown, etc.) or to test different TW modelling hypotheses.
- Final debriefing with all attendees will close the HIL2 experiment period.

Day	Morning	Afternoon
19th October	Simulation devices presentation	Familiarisation with tuition scenarios (session 1)
20th October	Familiarisation with tuition scenarios (session 2)	Operational training (session 1)
21st October	Operational training (session 2)	Operational training (session 3)
22nd October	Experimental runs #1, 2 & 3	
23rd October	Experimental runs #4, 5 & 6	
26th October	Familiarisation and Training	Experimental runs #7 & 8
27th October	Experimental runs #9, 10 & 11	
28th October	Experimental runs #12, 13 & 14	
29th October	Experimental runs #15 & 16 – Final Debriefing	
30th October	Spare day –	

Table 13: Run plan for HIL2 experiment period

The scenarios are twofold:

- Training scenarios.
- Experiment/measured scenarios.

The training scenarios are built like the experiment/measured scenarios. Their aim is to develop the pilot and controller habits and skills needed when faced with the measured scenarios and the aircraft cockpit. The situations simulated will need to be equivalent in terms of traffic load and complexity to those in the measured scenarios. The controllers and pilots should not discover, during the measured sessions, working constraints unknown to them and for which they need to learn *in situ*. In such cases, measurements will assess learning more than usage. Controllers and pilots will clearly not be able to acquire complete and perfect mastery of new traffic organisation and management concepts with only half a day's training, but an attempt must be made to control these parameters as tightly as possible in the light of organisational, planning and staff availability constraints. Four training scenarios will be required for the half-day of training. The content of these scenarios will be determined such a way as to be representative of the experiment/measured scenarios in terms of traffic density and occurrences.

The experiment/measured scenarios will be the situational and operational support for the controllers' and pilots' comments and the measurements of their activity in order to achieve the experiment objectives. Each scenario will last approximately 70 minutes and will start with a change-over situation for the controllers. For the aircrews, the flight will be initiated 3 minutes before the first measured sector, and will be stopped 3 minutes after overflying the last measured sector. An overall description of the scenario can be found in D1.3.4 [21]

9.2.4 HIL2 data analysis plan

Automatically recorded, system and human performance data will be processed by statistical methods.

Questionnaires and self-evaluation scale data will be processed by statistical methods.

Subjective data gathered via observations and verbal comments will be processed by analysis of their content. The debriefings will be recorded, but a thorough review of them is not envisaged. The recordings will be used only in the event of doubts relating to certain situations or explanations.

The qualitative and quantitative data will be collated on the basis of the phases of the scenario and the constraints encountered in the course of the scenarios.

Data will be analysed on the following basis:

- By controller and pilot, in order to find behavioural invariants between controllers and pilots, their analyses and their understanding of the scenario's various phases.
- By scenario phase, in order to search for usage invariants associated with situational constraints.

The human performance results will be analysed in relation to the system performance results in order to analyse the CoO’s acceptability by the ATCOs and aircrews, and regarding future traffic evolution in 2020. This level of analysis is also very relevant in order to understand the advantages and deficiencies of the CoO concept from an operational point of view. The results provide promising ways for adapting the concept, assessing its compliance with the SESAR objectives, and refining the ATCOs' and aircrews' working methods and the HMI proposed for the following experiments at the CWP and in the cockpit. The complete analysis plan can be found in Annex B (14).

9.2.5 Safety analysis

As explained in 9.2, HIL1 identified a number of limitations regarding measurement of the impact of the CoO on safety through STCA occurrences. Measurement of separation performance derived from failure to maintain separation, such as STCA alarms, would not seem to be relevant for small-scale experiments.

A new approach was proposed within the EUROCONTROL Experimental Centre [22], based on successful separations and failures, providing a more balanced view of controller performance. The idea is to measure how separation is achieved during successful operations (i.e. the majority of operations).

The proposed Separation Performance Tool (SPT) [22] predicts the aircraft trajectory and dynamically defines potential losses of separation. When a potential loss is detected, the tactical controller’s actions to resolve that conflict will be monitored.

The indicators of separation assurance are:

- Intervention type for resolving predicted conflict (heading instruction, speed change, altitude change)
- Time of intervention before potential loss of separation
- Distance before potential loss of separation when the tactical controller intervenes
- Localisation of controllers’ interventions

- Localisation of predicted losses of separation (LoS)[1]

The advantage of using the SPT tool is mainly that the data are captured automatically and are objective data, not dependent on controller opinions or occurrences of exceptions.

The SPT tool gives an indication of how the traffic was managed during the experiment. The tool gives an indication of the safety buffers applied by the controllers around the aircraft, and how these buffers can be affected by the implementation of the proposed concept. However, this tool does not allow a better understanding of traffic management strategies used by ATCOs (e.g. aircraft close to separation limits could all be under control and fully managed by the ATCOs).

The SPT tool was used for the CATS HIL2 run in Geneva in October 2009. The results of the investigations will be available in the D2.1.2 report.

9.3 Third validation experiment (HIL3)

The third and last validation activity provided for in the DOW [2] is HIL3, following HIL2 (October 2009). This HIL experiment will be prepared in WP1.3 and performed in WP2.1.3.

The objectives of HIL3 are to evaluate the CoO concept in the event that the TWs cannot be fulfilled. This will be a first attempt at a renegotiation process.

The goals of this experiment are twofold:

- To investigate the impact of the renegotiation on the operators, in terms of feasibility, usability, acceptability for pilots and controllers
- To assess the renegotiation process from the actors' point of view:
    - Gain a better understanding of the actors' role and responsibilities, in particular how they interact
    - Define system functions needed to support renegotiation and provide a demonstrator

Erreur ! Des objets ne peuvent pas être créés à partir des codes de champs de mise en forme.

Figure 17: HIL3 features

The validation will focus on the initial principles of the renegotiation process, addressing the V1 maturity level. The outputs of this assessment will highlight the elements of the renegotiation mechanism and identify specifications for the future system and tools needed to continue in V2. The validation team has decided to use gaming techniques to assess the renegotiation mechanism. These gaming techniques, coupled with the HIL platform, will help to clarify the collaborative decision-making process and its impact on the operators.

[1] Before the separation minima, so before 5NM

For the first goal, the impact of the renegotiation on the operators, a similar approach to the validation applied in HIL1 and HIL2 will be used, assessing the benefits for system performance (KPA) and the human contribution. It should be noted that due to the duration constraints (5 days of experiments), it will not be possible to apply a quantitative comparative methodology based on statistical analysis, as developed in HIL1/2.

Table 14 below shows the various KPAs and KPIs which will be used to assess the system performance, and how human performance will interact with this.

<table>
<tr><th>HIL 3</th><th>Assess the impact of the renegotiation of a CoO from the actors' point of view</th></tr>
<tr><td>Objectives relating to system performance</td><td>Objectives relating to human performance</td></tr>
<tr><td>SAF
CAP
EFF
PRED</td><td>■ G4.1: Assess the strategic decision process
■ G4.2: Assess the impact of renegotiation on the tactical situation</td></tr>
<tr><td>Indicators</td><td>Indicators</td></tr>
<tr><td>SAF.LOCAL.ER. PI (1, 2, 3, 5, 6 & 8)
CAP.LOCAL.ER. PI (2, 8, 10, 11, 12 & 13)
EFF.LOCAL.ER. PI (1, 7, 8, 9, 10, 11)
PRED.LOCAL.ER. PI (1 & 2)
Number of TWs renegotiated</td><td>■ Workload: ISA, NASA-TLX, Interviews, Observations, Performance outcomes, Questionnaire
■ Situation Awareness: SASHA_Q, Interviews, Observations, Performance outcomes, Questionnaire
■ Error production and management: Observations, Questionnaire, Interviews, Performance outcomes
■ Operator and actor activity: Cognitive processes, Decision-making, Risk management, Constraints, etc.
■ Collaborative and R/T activity: Communications content</td></tr>
</table>

Table 14: HIL3 Validation Plan

For the second goal of the experiment, the validation of the renegotiation process, one of the main activities will be to delineate the decision-making process, by understanding how each actor selects his/her priorities, how each actor expresses his/her constraints and how they evaluate the proposed new TWs, or to summarise, how the actors identify collectively solutions for a global objective.

To do this, particular attention will be paid to the system functions needed to support this renegotiation, in terms of tools, visualisation, etc.

9.3.1 HIL3 experimental plan

The experimental plan has been designed with the relevant variables to comply with the HIL3 validation plan, presented in Table 14.

Variables are properties or characteristics of some event, object, or person that can take on different values or amounts.

An experimental plan uses two types of variables:

- Dependent variables
- Independent variables

Dependent variables are those which are seen to change in response to independent variables. Independent variables are those used deliberately to cause a change in dependent variables: "if *x* is given, then *y* occurs", where *x* represents the independent variables and *y* represents the dependent variables.

Variables may have different properties. The most common properties encountered in experimental variables are:

- Discrete (possible scores of discrete points on a scale) or continuous (the scale is continuous and not made up of discrete steps).
- Qualitative (variables that express a qualitative attribute) or quantitative (variables that are measured in terms of numbers).

The independent variables are linked with the experimental hypothesis, and determine the experimental plan (i.e. the number of runs and experimental duration).

The dependent variables determine the range of the collected data and the analysis plan.

9.3.1.1 Renegotiation process (G4.1)

9.3.1.1.1 Independent variables

Three independent variables will be controlled and used during the HIL3 experiment:

- Collaborative or cooperative process.
- Renegotiation trigger
- Traffic load.

Renegotiation process

Renegotiation is a collaborative process between actors which may be achieved in different ways. In one approach, each actor issues its own requests, and one of the actors (most probably the airline actor) decides on the best solution by taking into account all the expressed constraints. In another approach, all the actors share with other actors their own constraints, and the final decision is taken cooperatively by all actors. These two approaches to the renegotiation process outline the features of various processes which could have an impact on the decisions and support for the actors.

Consequently, one of the HIL3 experiment objectives will be to assess whether one of these two renegotiation processes is more efficient than the other.

Two renegotiation process methods will be used:

- Collaborative process, in which each actor makes its own requests to the airline actor.
- Cooperative process, in which a decision is taken by agreement between all actors.

Renegotiation trigger

The reason for the renegotiation may impact on the extent of the renegotiation and also how difficult it is. Renegotiation will also be impacted by traffic and operational conditions. However, the trigger for the renegotiation is a key element in how the renegotiation will be led. In addition, SESAR proposes various operational scenarios with which each new ATM concept will have to cope.

Taking account of these elements, various renegotiation triggers will be assessed during the HIL3 experiment. Five triggers will be used and controlled:

- Trigger 1: airport closure.
- Trigger 2: airport closure and operational management of two aircraft from the same airline (SWAP).
- Trigger 3: forbidden airspace.
- Trigger 4: aircraft windshield crack.
- Trigger 5: entry in the ECAC area modified.

More information on these triggers can be found in the D1.3.5 HIL3 Experimental Scenarios Description [26].

Traffic load

All projects involved in the SESAR setting seek to propose new ways of coping with the planned traffic growth in the future. The purpose of HIL3 is to assess the renegotiation process in an operational environment, meeting the 2020 requirements. Consequently, several traffic loads must be used in the simulation scenarios.

Two traffic load levels will be used:

- Current traffic load (in 2008).
- Expected traffic load (in 2020).

According to the EUROCONTROL forecasts for the sectors measured for the HIL3 experiment, traffic is expected to increase by 40% in 2020. Traffic growth is expressed in terms of the hourly flow for each of the two measured sectors.

9.3.1.1.2 Dependent variables

The dependent variables are twofold: quantitative and qualitative

Quantitative variables

Quantitative variables measure the renegotiation process duration. The temporal aspects of the renegotiation process are crucial for assessing the feasibility of such renegotiation. Three durations will be collected:

- Duration from the trigger to the TW generator launch.

- Duration of new target windows validation, from the TW generator delivery to the validation of this new set of TWs by all the actors involved.
- Full duration of the renegotiation process, from the trigger to the target windows validation, including the TW generation.

Qualitative variables

Qualitative variables assess two elements:

- The content of the renegotiation
- The actors' feelings.

The content of the renegotiation will be assessed using an observation list on which the interactions between actors and between actors and operators will be noted, as will problems understanding one another, conflicts, deconflicting strategies, and deviations from the renegotiation process. Particular attention will be paid to the formulation of actors' constraints and how they decide on their own priorities in the decision-making process. This will then help to identify potential functionalities and tools for supporting the actors' decisions.

The actors' feelings will be assessed using two techniques:

- Post-run debriefing and post-experiment debriefing
- Post-experiment questionnaires

9.3.1.2 Impact of renegotiation on ATCOs and pilots (G4.2)

9.3.1.2.1 Independent variables

Aircrew

For the aircrews, no independent variables will be used. The purpose of the experiment will be to assess the impact of the renegotiation process, when it occurs, on the pilot's activity. Renegotiation induces a specific task for the pilots which cannot be compared with another operational situation.

Controllers

For the controllers, two independent variables will be used: traffic load, and renegotiation reasons / triggers.

Two traffic load levels will be used:

- Current traffic load (in 2008).
- Expected traffic load (in 2020).

These two conditions are the same as those described above for the renegotiation process by actors.

The reasons for/ triggers of the renegotiation may have different effects on the number of aircraft impacted and the problems generated by the new operational situation. In order to better understand how the controllers will cope in these situations, the nature of the trigger has to be considered.

As for the renegotiation process (§9.3.1.1.1), five triggers will be used and controlled:

- Trigger 1: airport closure.

- Trigger 2: airport closure and operational management of two aircraft from the same airline (SWAP).
- Trigger 3: forbidden airspace.
- Trigger 4: aircraft windshield crack.
- Trigger 5: entry in the ECAC area modified.

9.3.1.2.2 Dependent variables

Aircrew

The aircrews' dependent variables will be quantitative and qualitative.

Quantitative data will comprise measurement of the workload and situation awareness of the pilots involved.

Qualitative data will comprise the assessment of the impact of this renegotiation from the pilots' point of view by means of post-run and post-experiment debriefings as well as post-experiment questionnaires.

Controllers

The ATCOs' dependent variables will comprise system performance and human factors indicators, as described for HIL1 and HIL2.

System performance indicators encompass the measurement of:

- Safety.
- Efficiency.
- Predictability.

The human factors indicators will be:

- Workload and situation awareness.
- Observation data from controller activity.
- Collaborative work between executive and planning controllers.
- Collaborative work between controllers and pilots.

9.3.2 Subjects

9.3.2.1 Operators

9.3.2.1.1 Air traffic controllers

The experiment will be performed with operational air traffic controllers in active service at Italy ENAV En-Route ACC.

Four controllers are required. To satisfy the required profile, they should:

- Be in active operational service at an en-route control centre.
- Have experience with the measured sectors.
- Be familiar with experimental air traffic control simulation environments.

- Be aware of the problems of innovative research and the upstream study of concepts.
- Have knowledge and an understanding of the CATS study.

It is desirable for the controllers to first run the HIL2 experiment in order to avoid CATS training, and improve familiarisation with the concepts and the control working position.

9.3.2.1.2 Aircraft pilots

The experiment will be performed with Air France professional pilots.

Two pilots are required. To satisfy the required profile, they should:

- Be in active operational service at Air France. Pilots who are active or who retired less than one year ago are required.
- Have experience with glass cockpit aircraft in order to be familiar with and favourable to the use of modern aircraft systems and interfaces. In addition, the future of ATM, and the SESAR Project relies on aircraft navigation aids such as the Flight Management System (FMS). The experiment has to take full account of these aspects. The aircraft simulation platform is a glass cockpit aircraft.
- Be aware of the problems of innovative research and the upstream study of concepts.
- Have knowledge and an understanding of the CATS study.

It is desirable for the pilots to first run the HIL2 experiment in order to avoid CATS training, and improve familiarisation with the concepts and the flight simulator.

9.3.2.2 Actors

Three main actors are required for the HIL3 experiment, as described in [26], in order to run the renegotiation process:

- Network Manager (ANSP Flow Management Position), also representing the tower runway controller
- Airline Operation Centre staff (AOC)
- Airport Operation Centre staff (APOC), also representing the Ground Handling Agent

The three actors need to be active and experienced in their service in the same function.

The CATS concepts, HIL3 goals and HMI features will be presented to each actor before the experimental session, in one day of training.

9.3.3 HIL3 Conduct of the experiment

The experiment will take place over five days at the SkySoft premises in Geneva, Switzerland from 14 to 18 June 2010.

These five days will include:

- Pilot and ATCO training relating to the CATS scenario (TW management), the simulation environment and interfaces, and traffic management in the measured sectors.
- Conduct of the measured experimental simulations, including the gaming exercises

9.3.3.1 Scenario design

An overall description of the HIL3 scenarios is given in the Deliverable D1.3.5 CATS HIL3 Operational Scenario description [26].

Each scenario will encompass several renegotiation processes. In each scenario, aircraft piloted by pilots will be involved in the renegotiation processes, except for trigger 5.

The learning effect will be avoided by proposing various traffic patterns, operational situations, and function assignments. Any run designed from the scenario will be to the same as another.

9.3.3.2 Experiment design

The HIL3 experiment has been designed in order to assess the renegotiation process at the actors' level, and the impact of new TWs at the operators' level.

Due to the duration of the HIL3 experiment and the need to manage various independent variables, no statistical analysis of quantitative data will be made. In addition, at this early stage of renegotiation concept maturity, the gaming exercises are by definition exploratory, allowing the validation team to refine the concept, and clarify roles and responsibilities.

As a consequence of this, the experiment has been designed to test several experimental situations in which the independent variables are encompassed without the need to quantitatively compare these situations. The objective is to learn from the runs' qualitative outcomes in order to better understand the mechanisms of the renegotiation process, and illustrate the subjects' feelings using the quantitative data collected.

As mentioned above, the HIL3 experiment covers the V1 validation level of the E-OCVM lifecycle model. Acceptability and scope of the concept are the main purposes.

The experiment is structured around runs. Each run is a sequence of traffic control in which some aircraft are piloted from flight simulators by pilots and others are automatically controlled by orders given by controllers on the control working position. During the traffic control sequence, various triggers occur and require TW renegotiation for aircraft impacted. The renegotiation process is carried out by the actors located together in the same room, but outside the control and flight simulator room. During the renegotiation process, the traffic flies while controllers and pilots are carrying out their tasks. At the end of the renegotiation process, new TWs, validated by all actors, are implemented in the system. At this point, the controllers and pilots have to manage the new TWs.

Each run lasts 80 minutes.

As shown in Figure 17, there are two control positions with an executive and planning controller on each position, and two pilot positions. During a run, each pilot flies consecutively two flights during the en-route phase.

The first day will be dedicated to subject training and familiarisation with the tools, use of the concepts and working environment.

Ten measured runs will be completed; three runs a day. The tenth run will be completed on the morning of the last day, before the final debriefing.

9.3.3.3 Experimental layout

The HIL3 experiment is supported by simulation facilities, and will take place at the Skyguide premises on the ATC training simulator.

Three work places will interact:

- ATCO CWPs for the 2 sectors simulated (one position for the executive controller and one position for the planning controller in each sector);
- Two Airbus A320 flight simulators each piloted by one pilot;
- A CDM module, with three positions for airline, ANSP, and airport actors. Each position will have a dedicated HMI in order to display relevant data and control the process. The TW generator is connected to this module in order to exchange data, calculate TWs, and provide outcomes to the actors for validation.

9.3.3.4 Experiment schedule

The experiment timetable encompasses three main periods:

- Pre-experiment period.
- Experiment period.
- Analysis period.

9.3.3.4.1 Pre-experiment period

The pre-experiment period encompasses the HIL3 experiment overview for the ATCOs, pilots and actors' staff involved.

A presentation of the CATS concepts and HIL3 goals will be scheduled before the HIL3 experiment for the participants.

An experiment handbook will be produced to present the CATS concept, experiment’s purposes, simulation devices, HMI, working methods, and measurements. This document will be delivered before the HIL3 preparation meeting.

9.3.3.4.2 Experiment period

The experiment period will last five days from June 14 to June 18 2010. This period will encompass three phases:

- Familiarisation and training phase;
- Measured runs phase;
- Debriefing phase.

Familiarisation and training phase

The familiarisation and training phase will take place on the first day. The schedule of the day will be:

- CATS concept and HIL3 presentation for all participants;
- Working position presentation for ATCOs, pilots, AOC, APOC and NM.
- Training exercises encompassing interactions between ATCOs, pilots and actors.
- Use of HIL3 experiment measurement techniques
- Familiarisation with exercises.

The familiarisation will be led with tuition scenarios in order to confront the subjects with very typical demonstration situations. The scenarios must illustrate the features of the experiment, i.e.:

- Familiarisation with their working position.
- Learning about the tools and the operational environment.
- Working methods.

Training scenarios will be built like experiment/measured scenarios. Their aim is to develop the controllers’ pilots’ and actors’ habits and skills needed when faced with the measured scenarios. The situations simulated will need to be equivalent in terms of traffic load and complexity to those in the measured scenarios. The subjects should not discover, during the measured sessions, working constraints previously unknown to them and for which they need to learn *in situ*. In such cases, measurements will assess learning more than usage. The controllers, pilots, and actors will clearly not be able to acquire complete and perfect mastery of the use of CATS concepts with only one day's training, but an attempt must be made to control these parameters as tightly as possible in the light of organisational, planning and staff availability constraints.

The content of these training scenarios will be determined so as to be representative of the experiment/measured scenarios in terms of traffic density, renegotiation triggers, and complexity.

Measured runs phase

Each measured run will last 80 minutes. Three runs will be performed in a day. Consequently, 10 runs will be performed during the measured runs phase from June 15 to June 18 2010.

The runs have been designed from the independent variables: traffic load and renegotiation triggers. The "renegotiation process" variable has no impact on the run design, only on the actors' working methods.

Four equivalent traffic samples have been created, to avoid learning effects for the participants. The "traffic load" variable is systematically managed in all runs because the traffic samples are designed with two parts:

- First part for the 2008 traffic;
- Second part for the expected 2010 traffic.

The "renegotiation trigger" variable has 5 modalities:

- Trigger 1: airport closure.
- Trigger 2: airport closure and operational management of two aircraft from the same airline (SWAP).
- Trigger 3: forbidden airspace.
- Trigger 4: aircraft windshield crack.
- Trigger 5: entry in the ECAC area modified.

It is not possible to play all modalities in the same run. As the airport closure may concern several aircraft at the same time (at least four aircraft piloted per run, two in the first part of the traffic sample, two in the second part), the validation team decided to mix the three last triggers (i.e. one forbidden area in the first part of the traffic sample, involving several piloted aircraft, and one aircraft failure (piloted aircraft) in the second part of the traffic sample). It should be noted here that the last trigger will concern only the actors and not the ATCOs and pilots.

In order to achieve the goal of performing 10 runs while avoiding learning effects, various parameters will be given to the actors (e.g. different airport weather, different parking occupancies and different areas forbidden, etc.)

Table 15 below shows the HIL3 experiment schedule using the traffic samples and triggers. Table 16 describes the parameters used in HIL3 as a function of to the corresponding triggers. A more detailed explanation can be found in [26]

Date	June 15			June 16			June 17			18
Traffic sample	**CATS _1**	**CATS _5**	**CATS _6**	**CATS _9**	**CATS _5**	**CATS _1**	**CATS _6**	**CATS _9**	**CATS _1**	**CATS _5**
Trigger/ parameter	**1a**	**2a**	**3a 4a 5a**	**3b 4b 5b**	**1b**	**2b**	**3b 4b 5b**	**3a 4a 5a**	**1c**	**2c**

Table 15: HIL3 schedule

1 a/b/c	LFLL closed + different parameters (parking occupancy, weather, runway capacity, technical limitations, etc.)
2 a/b/c	LFLL closed + SWAP TW arrival + different parameters
3 a/b	Forbidden area over LAPRI TIXIR IXUSA VEROB / over MEDAM KUKEV ADISO VANAS ALBER
4 a/b	Aircraft failure / Aircraft failure + diversion
5 a/b	Modification of out of ECAC departure flight 1/ flight 2

Table 16: HIL3 triggers and parameters

The schedule of each run is:

- Short run presentation and briefing for subjects;
- Run performance (80 minutes);
- Completion of questionnaire and self-evaluation scales (15 minutes)
- Post-run debriefing (20 minutes)

Debriefing period

At the end of the experiment, on the afternoon of the 18th, a post-experiment debriefing will be held with all participants. The debriefing will be preceded by a post-experiment questionnaire.

9.3.4 HIL3 data analysis plan

Data will be processed at the two experiment levels:

- Renegotiation process at the actors' level;
- Implementation of new TWs at the operators' level (controllers and pilots).

The main purpose is a qualitative analysis of data recorded by observation, questionnaires and debriefings. Subjective data gathered will be processed by analysing its content.

Data will then be analysed with quantitative outcomes. Quantitative data will not be processed using statistical methods because the number of measures will not be sufficient in terms of management of the independent variables and the number of runs. The assessment of the CATS concept has already been performed in HIL1 and HIL2; this is not one of the objectives of this experiment. The experiment focuses more on renegotiation refinement than assessment.

The HIL3 experiment will offer outcomes on strategic decision-making processes as well as renegotiation feasibility and guidelines for improvements. Attention will be paid to the requested tools and the information each actor needs to properly perform his/her role during the renegotiation.

At the operators' level, HIL3 will help to build a better understanding of how operators manage the period between the renegotiation initiation and the implementation of a new set of TWs. Acceptability of the proposed process will be assessed and management of the new set of TWs will be analysed in depth.

10 Systemic validation plan

10.1 Benefits assessment

In the context of the CATS systemic assessment, WP2.2.2 develops a Cost-Benefit Analysis (CBA) whose scope is to build an Economic Feasibility Case that serves as an instrument for presenting the economic impact of the CoO to the stakeholders and obtaining their feedback regarding the acceptability of this impact.

The CBA investigates from an economic perspective the potential benefits and limitations of implementing the CoO for airlines, airports and air navigation service providers. The analysis consists of two main steps which reflect the existing Level of Maturity of the CoO operational concept. In particular, this CBA qualitatively weighs positive and negative impacts of the CoO with respect to the current system and presents a quantitative study which highlights the role of TWs as a possible tool to implement a User Driven Prioritisation Process (UDPP).

The aim and scenarios to be used in the development of the Economic Feasibility Case must be stated. First of all, this involves the identification of the scope and boundaries to be considered. In particular, Section 10.1.1 delineates a general framework of the positive and negative impacts (benefits and costs respectively) to be included, together with their generating mechanisms. Chapter 3 of D2.2.2 (Report on Cost-Benefit Analysis) [27]describes in much greater detail the benefits and drawbacks linked to the implementation of the Business Trajectory through the CoO concept of operations. Scenarios are developed and evaluated by representatives for the following stakeholders: airport, ANSP and airline. Each scenario is decomposed into a hierarchy of criteria (or elements) which are likely to have an impact on them. Both flight planning and flight execution phases are considered for each stakeholder.

10.1.1 Scope and Boundaries

This section focuses on the observation of the positive and negative impacts generated by the concept of operations proposed to the three stakeholders identified above. These impacts will be referred to as benefits and costs, respectively.

A benefit is a perceptible, substantiated and tangible improvement for a stakeholder. It follows from technical, operational, organisational, and/or institutional changes. Benefits might have associated costs (for the same or other stakeholders). Benefits and costs can be quantifiable and non-quantifiable (soft benefits). If they are quantifiable, they might be translated into monetary terms.

10.1.1.1 Benefits identified

In order to delineate the benefits to be considered in the analysis, it is useful to point out first the mechanisms involved. A benefit mechanism is a specific change implied by the concept of operations that has a positive impact on one or more stakeholders. Benefit mechanisms can be represented in a causality diagram (Figure 18). This documents the changes introduced in the ATM environment and explains how those changes lead to benefits.

The starting point of a causality diagram is the operational concept, the implementation of which will generate the identified benefits via the mechanisms reported. It is represented as a rectangle with rounded corners. From this node, the stakeholder benefits are listed. They are represented as hexagons. In addition, the benefit mechanisms are depicted. They correspond to the characteristics of the ATM system (measurable or otherwise) that will be affected by the operational concept itself, and KPAs. They are represented as rectangles. An arrow drawn on one of these rectangles indicates an increase or decrease in the given value. Finally, rectangles are linked by arrows indicating the rationale for the causality connection between two mechanisms.

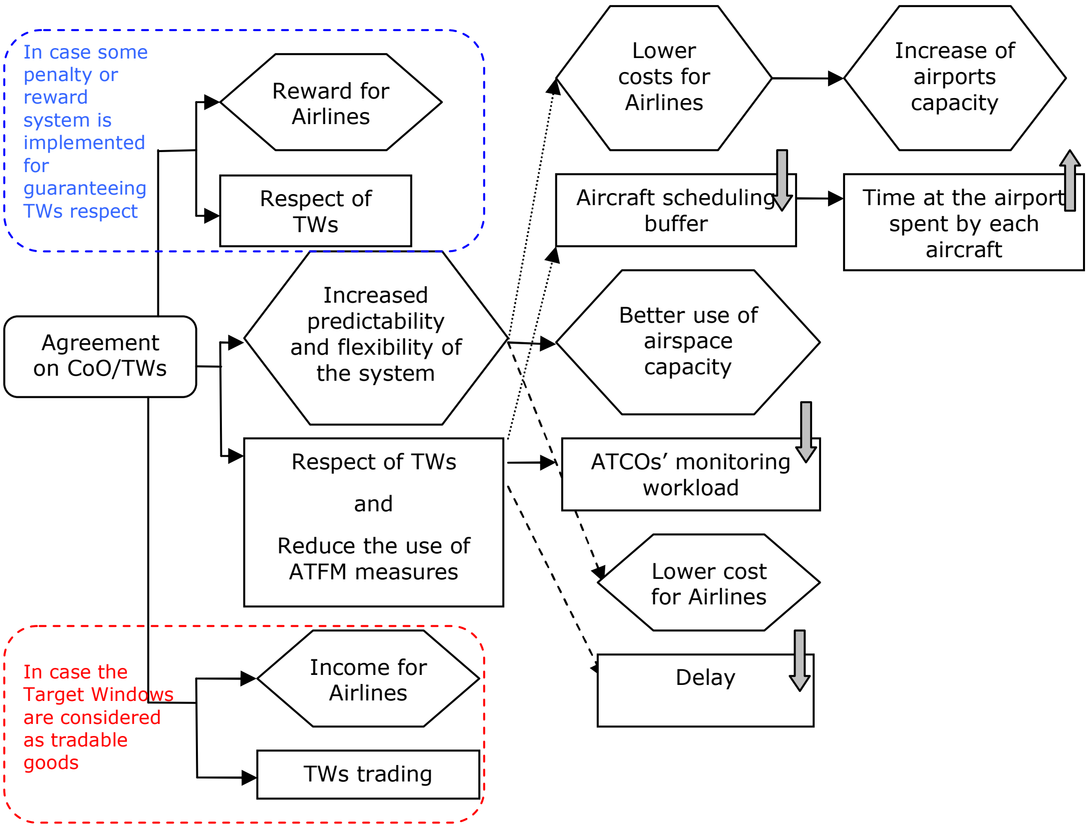

Figure 18: Graphical representation of benefits mechanism

The first element considered is the compliance of the TWs. Such compliance increases the predictability of the system as a whole. This is due to the fact that through the CoO, all the actors involved jointly agree to fulfil the agreements and to "deliver" flights in accordance with them. TWs allow each actor to verify its own operations in relation to the approved contract, and thus to ensure global predictability as soon as it is able to "deliver" the flight in accordance with them. A further point to be noted here is that the departure and arrival airports are included in the broader Gate-to-Gate context in which a CoO is developed, and this integration makes it possible to represent the whole flight more realistically. It contributes significantly to the increase in predictability. Moreover, through the renegotiation process, operators have room for manoeuvre to deal with uncertainty. This element increases the flexibility of the system.

The compliance of TWs leads, first of all, to a reduction in the delay, and, as a consequence, in the cost of delay. Airlines can clearly take advantage of this element.

Moreover, these stakeholders may exploit the increased predictability by reducing the aircraft scheduling buffer, thus achieving a higher rate of usage of their fleets.

Such a reduction in the aircraft scheduling buffer and delay implies a greater airport capacity, since the space availability of each aircraft at the airport is required for a shorter and more predictable time interval. On the other hand, the airport capacity increases also for the introduction of TWs in the departure and arrival airports in the negotiations: In particular, the fact that all partners in the system are working towards compliance of the TW at the destination airport improves on-time operations and the turn-around process.

On the other hand, ANSPs will enjoy major benefits as a result of the increase in predictability, and in general from the implementation of CoOs. First of all, according to SESAR the use of a shared trajectory as a means to detect and reduce potential conflicts implies a decrease in the monitoring workload of ATCOs. Another source of a reduced monitoring workload is the fact that the enhanced coordination achieved through the use of shared trajectories removes the need for explicit coordination between sectors. From a different perspective, by knowing in advance the agreed RBT, and by being part of the negotiation process, ANSPs have the potential to identify demand/capacity imbalances in advance, and to decide whether to allocate capacity or to introduce constraints in the negotiation.

Finally, airlines and ANSPs may obtain income if TWs are considered as tradable goods and they positively enter such a market. Other benefits may be achieved if a penalty and reward mechanism is applied to ensure the enforcement of the CoO and if stakeholders maintain a positive balance in terms of compliance with the agreed constraints.

10.1.1.2 Costs identified

As regards costs, a convenient approach consists in grouping them in various categories. Figure 19 shows the identified cost drivers.

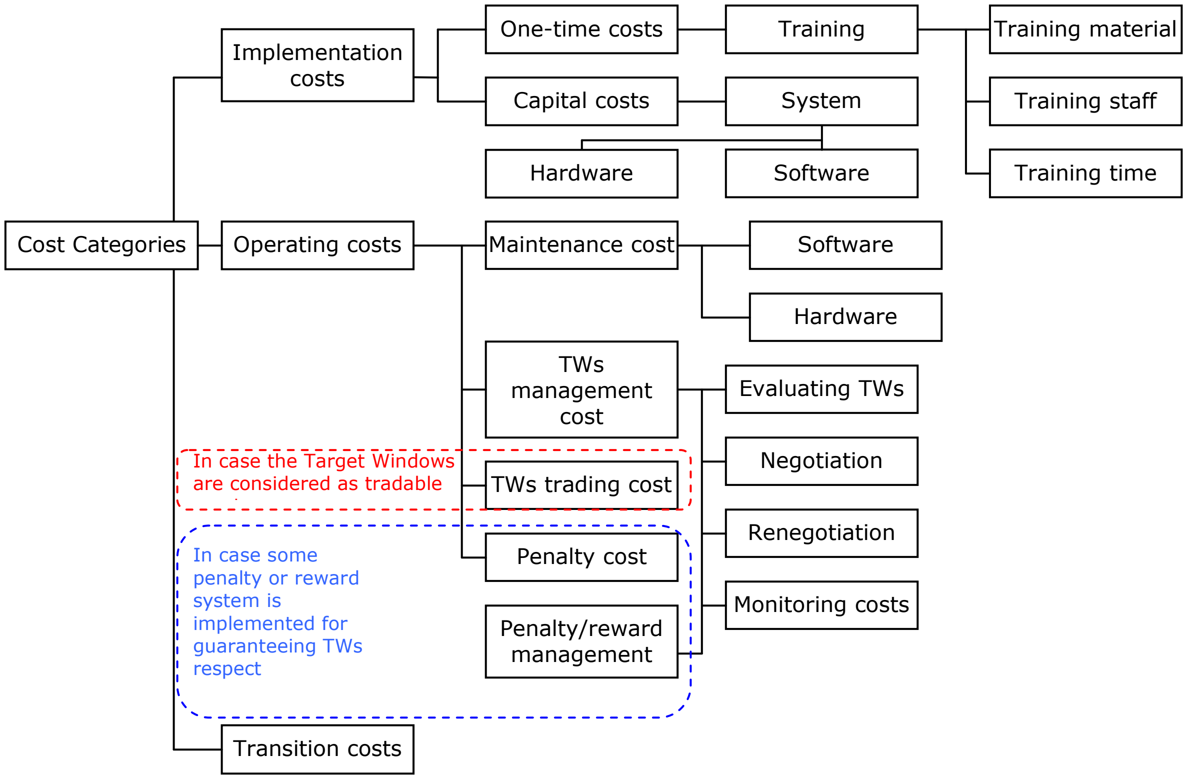

Figure 19 : Graphical representation of cost categories

The first cost category concerns implementation costs, which can be divided into one-off and capital costs. In the current analysis, the former correspond to the human resources training expenditure accounted for by training material, staff and, clearly, time. The latter are mostly system-related. They can be identified with the need for hardware and software investment.

The second category includes operating costs. They can be grouped into maintenance and TW management costs. Maintenance is mainly required for the hardware and software resources, while TW management includes the negotiation and renegotiation of CoOs, the monitoring of TW compliance, and the evaluation of TWs. Such an evaluation, as previously introduced, is necessary in the negotiation and renegotiation phase, in order to be able to take sharp decisions in with the light of system constraints. This is particularly important where no feasible solution exists, i.e. if it is not possible to satisfy all requests without violating certain constraints. In this case, some of the actors need to slightly modify their restrictions. For example, airlines may agree to change the route of some flights, or ANSPs may increase the capacity of certain sectors for a period of time.

If a penalty or reward system is implemented to ensure the compliance of TWs, its management may generate operating costs, which can be seen in the TW management framework. Moreover, such a system would impose penalties on stakeholders responsible for causing the violation of the CoO. Finally, if a market of TWs is allowed, some trading costs may be generated, in terms both of management of the trading process and of commercial losses.

The last category to be taken into account concerns transition costs, when moving from the current situation to the implementation of the concept of operations proposed.

As regards costs, however, it should be noted that most investments will be necessary to comply with SESAR guidelines. Only the costs strictly associated with TWs and the CoO are to be allocated to the CATS Project.

10.1.2 Economic Feasibility Case Scenarios

In this Economic Feasibility Case two alternative scenarios are analysed.

The first scenario is identified as business-as-usual. Even if according to SESAR this will not be a feasible alternative in the future, it is used here as the base case. This choice is motivated by two considerations: First of all, costs and benefits may be quite clearly quantifiable. Moreover, different implementations of the SESAR indications may imply significantly different, and currently not identifiable, impacts.

On the other hand, a scenario in which CoO is implemented is considered.

The two aforementioned alternatives represent a first step analysis. This gives a first indication of the convenience of the concept of operations proposed. After obtaining the first-step results, a more detailed study can be performed. In particular, the role of TWs appears to be crucial to the impact that the CoO has on the system. One possible approach is to use a mathematical programming model for setting TWs. In particular, TWs are identified as an outcome of the collaborative decision-making (CDM) process where all actors display their preferences and constraints, and then a set of feasible TW values, if it exists, is derived. As previously mentioned, several steps need to be considered, i.e. various rounds of negotiation:

- Airport TWs are defined in line with an agreement among airlines and airports.
- Based on the airport TWs, airlines and ANSPs define the sequence of the en-route TWs. Thus it is possible to move from a feasibility problem to an optimisation problem by defining a suitable objective function.
- If the TW set is empty, adjustment policies will be needed. These may be based on priority rules to sequence flights in a regulated sector. Or they may rely on penalty/reward schemes.
- If new events occur, it is necessary to decide whether (and when) to re-optimise the whole system or to introduce robust optimisation techniques to deal with such disruptions.
- The process of updating the TWs is complete at the end of the strategic planning phase.

In each of these phases, an optimisation process needs to be performed. The weightings assigned to each element (e.g. the minutes of delay, the route travelled) may have a strong impact on the solution to the problem, and hence on the associated costs and benefits. The same applies to the consideration of the constraints imposed as more or less strict. A penalty can be imposed if constraints are violated, or unmet restrictions may not be allowed.

Chapter 4 of D2.2.2 [27] introduces a mathematical optimisation model to formulate one possible CoO negotiation mechanism in order to quantify the impact of TW negotiation on the stakeholders, suitable for adoption from the long-term up to the short term phase.

10.1.3 Impact Assessment Methods

The abovementioned costs and benefits associated with the implementation of the concept of operations proposed may be divided into two groups. In the first one, hard costs and benefits are included. These are the impacts that can be quantified in monetary terms. For example, better use of airspace capacity belongs to this group. An increase in the number of aircraft entering the airspace of a country can be translated into a decrease in air navigation service charges. The second group includes soft benefits and costs. These may be quantifiable, but it is very difficult to assign them a monetary value. As an example, minutes of delay may be quantified, but there is no generally accepted cost of delay, although some figures are available from recent studies and some average values have been proposed.

All these elements need to be taken into account in the project assessment, and various methods have been proposed in the literature and used in practice. Following the classification of costs and benefits, these methods can be separated into two groups: quantitative and qualitative methods. As the names themselves indicate, the former are based on elements that can be translated in monetary terms. The latter deal with all the elements involved in the decisions to be taken, both tangible and intangible. Clearly they have the disadvantage of involving a high level of subjectivity in the process.

For the application of all the techniques proposed in the following, input data on costs and benefits are needed. Depending on their nature, these may be monetary amounts (e.g. penalty costs or rewards), quantities of various kinds (e.g. minutes of delay), or even just qualitative judgments (e.g. magnitude of the decrease in the ATCO’s monitoring workload).

Due to confidentiality reasons, the CATS benefits assessment does not include a quantitative study to measure economic and financial metrics. Instead, it looks into the CoO and Business-as-Usual scenarios for each stakeholder (airline, ANSP, and airport) at three hierarchical levels: organizational, flight planning and flight execution

First of all the organisational level is addressed. This involves the modifications to the internal structure of the stakeholder organisations necessitated by the introduction of the new concept of operations. For example, human resources training costs need to be taken into account at this stage. Next, the flight planning level is tackled. All the benefits and costs incurred several months before the day of departure are involved. The first-round negotiation takes place, and the consequent effort needs to be taken into account. Concurrently, the TW assessment process starts, generating overheads. The planning phase goes up to the day of departure, and at this point much of the uncertainty has already been resolved. The second and third rounds of negotiation are completed, and the costs relating to the signature of the CoOs are taken into account. Finally, in the execution phase, late flight demands are incorporated in the existing CoOs. Flights are executed and a CoO renegotiation is implemented when necessary.

The analysis allows each stakeholder to choose the alternative that maximizes its utility, which can be defined as the difference between benefits and drawbacks associated with an implementation of the Business Trajectory. Such utilities are not expressed in monetary terms because some elements of the hierarchies are hard to quantify in economic terms

The Analytic Hierarchic Process methodology (see Sections 5.3.1) is used to perform this study.

11 Conclusions

The results of the assessments performed within the framework of the project have been detailed in the following reports:

- D2.1.1 HIL1 primary results analysis
- D2.1.2 HIL2 primary results analysis
- D2.2.1.2 Report on modelling and quantitative risk analysis
- D2.2.2 Report on cost-benefit analysis
- D2.2.3 Report on legal assessment

A summary of the results is shown in the table below:

Objectives	***Hypothesis***	***Status***	***Results***	***Comments***
Safety	CoO implementation does not affect safety operations.	Confirmed	Safety is not impacted by TW use, even where capacity matched the expected 2020 traffic load	CoO adds benefits in situation awareness, enhancing safety
	Potential safety improvement as a result of the greater predictability afforded by TWs	Confirmed	Safety is maintained and potentially improved due to the room for manoeuvre afforded by each TW	The results are based on FT simulations and complement HILs.
	TW concept provides additional flexibility for safety-critical case studies	Confirmed	Fast-time simulation results indicate the positive effect of the concept on TW-constrained conflict resolution for critical scenarios	Significant reduction in conflict probability is achieved by exploiting the freedom afforded by TWs when issuing resolution manoeuvres
Capacity	CoO is still manageable even with the increase in traffic forecast for 2020.	Confirmed	ATCOs and pilots found the concept manageable even in 2020 conditions, but with the support of dedicated tools.	2020 STATFOR predicted traffic was used during the HILs.
Efficiency	CoO implementation impacts positively on the flight within the sector (flight duration, etc.).	Confirmed	The TWs allow better adherence to the trajectory	CoOs are respected.

	TWs integrate flexibility to cope with uncertainty.	Confirmed	The number of "out of TWs" aircraft was found acceptable by all controllers	TW margins allow optimum navigation inside organisational windows.
Predictability	Implementation of TWs ensures application of schedules.	Confirmed	The TWs allow better adherence to the trajectory	Flight planning is respected.
	TWs integrate flexibility to cope with uncertainty.		The number of "out of TWs" aircraft was found acceptable by all controllers	This should be further studied in HIL3
		Confirmed	The number of "out of TWs" aircraft in FT simulations was kept to a low percentage for various sources of uncertainty	Additional study is required and potential improvement is expected, by implementing the TOA controller (WP2.2.1)
Feasibility & Acceptability	The working methods and tools offered to ATCOs and pilots, as a result of the CoO implementation, are acceptable (task sharing, role and responsibility, as well as the support tools offered).	Confirmed	TW information increases the traffic picture. TWs allow the same level of performance to be achieved. TWs do not modify the way the controllers work. TWs are easy to use and to learn.	Low impact on ATCOs' working methods.
	Implementation of CoO does not impose significant additional workload on ATCOs or pilots.	Confirmed by ATCOs Open for pilots	Target Windows management normally requires additional task. However, the level of workload remains the same because it is balanced by improved traffic picture acquisition.	HIL2 dealt with the en route part of the flight. This issue should be analysed in departure and arrival phases, and in connection with degraded situations.

		Confirmed	Monte Carlo analysis for various sources of uncertainty was used to verify the feasibility of the additional tasks. The results indicate a high TW hitting probability	Feasibility is maintained even by tightening the TW constraints. The latter can be used for the TW redesign process so as to reduce conservativeness
Economic Benefits	CoO implementation provides economic benefits to ATM stakeholders	Confirmed	A qualitative analysis suggests that CoO provides economic benefits in terms of the 'Business as Usual' scenario to airlines and ANSPs, but not to airports.	Results driven by the current economic situation that favours cost containment and limits investments. Results are quite robust for airlines and ANSPs, but could easily change for airports as soon as there is an economic recovery.

CDM implementation	CoO concept supports effective CDM mechanism	Confirmed	A Target Window CDM mechanism for airspace users and the Network Management can be designed for the short-term planning phase Airspace users can achieve delay cost savings of about 15% as a result of the First-Come-First-Served rule	The proposed CDM mechanism is – Individual Rational (each airspace user can benefit from participating in the CDM mechanism) – Budget Balanced (no need for external subsidisation) – Allocative Efficient (airspace users' utilities are maximised) Distributed (no revelation of confidential data from airspace users)

Table 17 : summary of the CATS hypothesis and results

The operational concept described in D1.2.2 has been defined with the right level of detail for the development of:

- an initial legal impact study,
- an initial Business Case,
- and for the identification of the related performance, in terms of users' acceptance and safety issues

Nevertheless the renegotiation process should be further analysed before testing operational and technical viability.

The platform and tools proposed to assess the concept were similar to the operational platform currently used at SkyGuide and allowed confirmation of the potential benefits.

12 References

The available documentation is described below:

[1] EC CATS Contract TREN/07/FP6AE/S07.75348/036889

[2] CATS Description of Work (Annex1 of [1])

[3] CATS Project Management Handbook D0.1 v2.0

[4] SESAR Definition Phase-Deliverable D1- Air Transport Framework - The current situation

[5] SESAR Definition Phase – D2 Air Transport Framework - The performance target

[6] SESAR Definition Phase-Deliverable D3- The ATM Target Concept

[7] SESAR Definition Phase-Deliverable D4- ATM Deployment Sequence

[8] CATS Concept of Operations Definition D1.2.1 v1.0

[9] CATS State of the Art D1.1 v1.2

[10] European Operational Concept Validation Methodology – Version3

[11] EP3 Performance Framework E3-WP2-D2.0-03-RQT

[12] CATS HIL1 Operational Scenario Description D1.3.3 v1.0

[13] *4D Trajectory Management Controller Simulation Report*, 2008, EUROCONTROL, EEC Technical/Scientific Report No. 2008-002

[14] Regulations on the Single European Sky , package 2, http://ec.europa.eu/transport/air_portal/traffic_management/ses2/index_en.htm

[15] Barnhart, C; Belobaba P; Odoni AR (2003) Applications of Operations Research in the Air Transport Industry, Transportation Science, 37*,* 368-391*.*

[16] Kotnyek, B; Richetta, O (2006) Equitable Models for the Stochastic Ground holding Problem under Collaborative Decision Making, Transportation Science, 40, 332-343.

[17] Halloway S (2003) Straight and Level: Practical Airline Economics, Ashgate Publishing Limited, Hampshire, England.

[18] Talluri, KT; Van Ryzin, GJ (2004) The Theory and Practice of Revenue Management, Springer, New York, NY.

[19] CAATS II (Cooperative Approach to Air Traffic Services II) Guidance Document for a Typical Business Case D20.1 v0.1

[20] CAATS II (Cooperative Approach to Air Traffic Services II) Business Good Practices D19.1 v0.8

[21] CATS HIL2 Operational Scenario Description D1.3.4 v1.0

[22] Pinska, E.; Hickling, B. (2009) Measuring En Route Separation Assurance Performance. In proceedings of EUROCONTROL Safety R&D Seminar 2009, Munich, Germany

[23] SESAR Validation and Verification, June 2009

[24] SESAR Concept Story Board, June 2009

[25] CATS Concept of Operations Definition D1.2.2 v1.5

[26] CATS HIL3 Experimental Scenarios Description D1.3.5 V1.0

[27] CATS Report on cost-benefits analysis D2.2.2 v1.0

13 Annex A

The following document is the data analysis plan for HIL1. This chapter brings together the indicators and metrics that will be used to qualitatively and quantitatively assess the HIL1 objectives, and explains how they will be measured during the HIL1 experiment.

(Please refer to the .pdf file attached)

14 Annex B

The following document is the data analysis plan for HIL2. This chapter brings together the indicators and metrics that will be used to qualitatively and quantitatively assess the HIL2 objectives, and explains how they will be measured during the HIL2 experiment.

(Please refer to the .pdf file attached)