

---

# **Comparison of EACAC and AGIE metrics using CARE/ASAS #2 framework**

# Project overview

---

## ◆ Context:

- To help the Agency in progressing with the Eurocontrol/FAA Action Plan 5 activities related to metrics comparison

## ◆ Objectives:

- To compare AGIE and EACAC'2000 experiments and metrics, including INTEGRA metrics applied on EACAC
- To help in getting a common understanding of US and European metrics

## ◆ Participants:

- CENA, EEC, Sofréavia (Leader)
- Eurocontrol, FAA HQ and Technical Center (Reviewers)

# Scope of the experiments

---

- ◆ **Air-Ground Integration Experiment (AGIE):**
  - Investigation of “shared-separation” concept (i.e. four conditions)
  - Human-in-the-loop experiments with both pilots/controllers
  
- ◆ **EACAC’2000 experiments:**
  - Initial evaluation of concept related to “limited delegation of separation tasks”
  - Mainly, from controllers’ perspective
  
- ◆ **INTEGRA study on EACAC:**
  - To conclude on effectiveness of the INTEGRA metrics under development, as well as on expected benefits from EACAC operational concept

# CEAMCA approach overview

---

- ◆ **Framework development for review of experiments (and metrics):**
  - Inputs from CARE/ASAS #2, MAEVA and VALSUP
  - Metrics classification per validation objectives, and per elements of the ATM system
  
- ◆ **Independent analysis of the experiments, then comparative analysis with respect to:**
  - Validation approaches, objectives and technique used
  - Metrics framework, characteristics and relevance
  
- ◆ **Consolidation (and lessons learnt)**

# Framework for experiment review

---

## ◆ Elements for metrics comparison:

- Operational concept under validation
- Validation approach and objectives (high-level)
- Experiment scope and objectives (low-level)

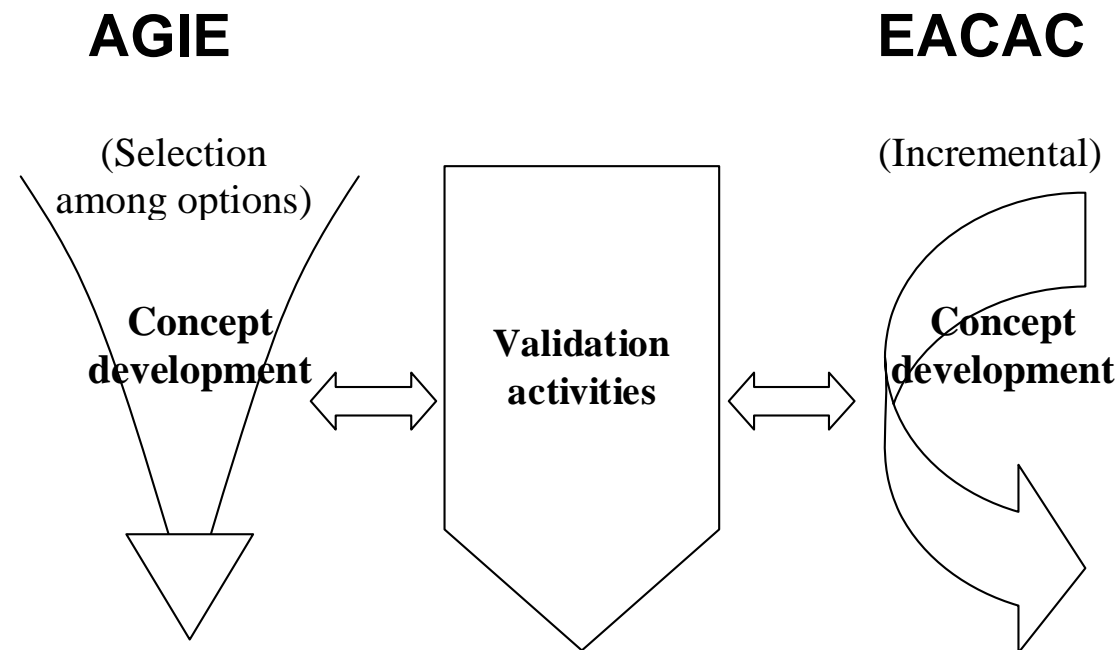
## ◆ Metrics classification and description:

- Type (absolute/relative), name and attributes (e.g. subjective/objective)
- Related experiment objective(s) under assessment and decision criteria to decide whether or not the objective is met
- Means of calculation including possible aggregation of “lower level” metrics/measurements (if any)

# Concept development vs. validation

---

- ◆ Both experiments had different concept development histories, which impacted their respective validation approach and experimental design



# External vs. Internal validation

---

- ◆ Both experiments were more dedicated to the evaluation of the new ATM system (i.e. “internal” validation), than actual evaluation of expected benefits and constraints (i.e. “external” validation)

Experiment objectives	AGIE experiment	EACAC’2000 experiment
(Related to validation)	<p>To identify <b>operational issues</b> (e.g., communications and procedures) that affect shared-separation operations</p> <p>To evaluate impact on controller/ pilot <b>workload</b> and <b>situation awareness</b></p> <p>To evaluate impact on <b>flexibility and efficiency</b> of the Airspace System</p>	<p>To assess <b>users</b> (i.e., controllers and pilots) <b>acceptance</b></p> <p>To assess impact on <b>controllers’ activity</b></p> <p>To assess impact on <b>flight efficiency</b></p> <p>To assess impact on safety (i.e. <b>human contribution</b> to safety, and <b>safe separations</b>)</p>
(Related to design)	<p>To provide recommendations for <b>information requirements</b> and <b>procedures</b></p>	<p>To assess usability of new <b>controllers’ working position interface</b></p> <p>To improve <b>procedures</b> and phraseology</p>

# Experiment validation objectives

---

- ◆ Both AGIE and EACAC experiments are missing hypotheses about impact of the concept on the ATM system, as well as its outcomes and outputs

- Only, judgement-based comparison of metrics depending on conditions

- Almost, no combined analysis of metrics

*Expectations from ATM community*

*WHAT (aims)*

*Expected benefits (and constraints)*

*HOW*

Operational concept

WHAT

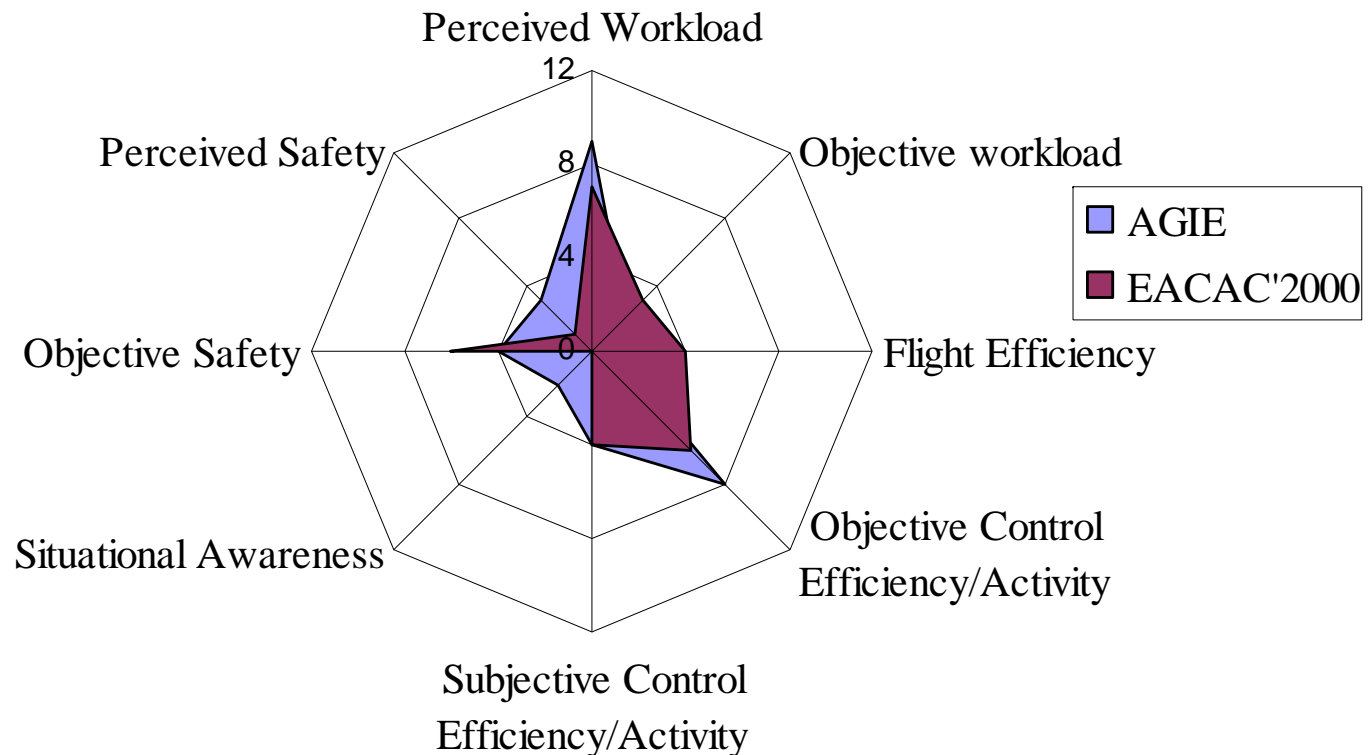
New ATM system

HOW

# AGIE & EACAC performance metrics

---

- ◆ Both AGIE and EACAC experiments used a majority of human-related performance metrics, as well as efficiency and safety metrics



# AGIE & EACAC metrics characteristics

- ◆ Both AGIE and EACAC experiments used a majority of subjective metrics for both design and validation objectives
  - Greater use of objective metrics within EACAC than in AGIE
  - Use of statistical metrics in AGIE, but not within EACAC

Metrics Attributes	AGIE experiment	EACAC'2000 experiment
Objective Quantitative Intrusive		Controller's physiological parameters: - Pupil diameter, Dwell time and Heart rate
Objective Quantitative Not intrusive	- Frequency and duration of Controller Ground→Air and Land Line Push-to-Talk Transmissions	
Subjective Qualitative Non intrusive 5 / 4 level of ratings	Controller Ratings for Physical, Mental, and Overall Workload	- Controller ratings of possible workload reduction - Controller ratings of perceived overall workload - Controller ratings of perceived workload (NASA TLX)
Subjective Qualitative Intrusive	Controller Interval Workload Ratings (WAK)	Temporal distribution of perceived overall workload per controller position (ISA)

# INTEGRA metrics characteristics

- ◆ **Absolute indicators based on weighted aggregation of objective measurements**
  - To be applied for evaluation of performance of any ATM system
  - Weights and scaling factors under development

INTEGRA Metrics on EACAC'2000		EACAC'2000 experiment	
Metrics	Attributes	Metrics	Attributes
- Information Processing Load (overall IPL combining elementary IPLs)	Objective Quantitative Scalar function (of time)	- Controller ratings of possible reduction in ATCO's workload	Subjective Qualitative Non intrusive 4 levels of ratings
- IPL for Determining the forecast interactions, - IPL for Planning resolutions, - IPL for Implementing the planned resolutions, - IPL for Other changes to trajectory	Idem	- Number of (type of) ATC instructions over time	Objective Quantitative Not binary
		- Controller ratings of factors (i.e., tasks) contributing to workload	Subjective Qualitative Non intrusive 4 levels of rating

# Consolidation of results (1/3)

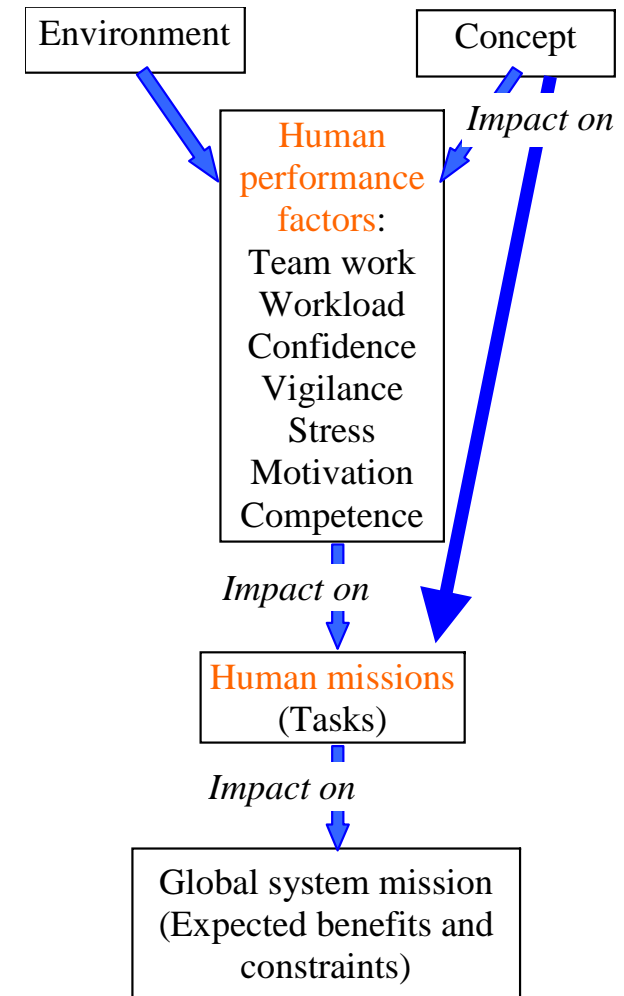
---

- ◆ **Need for clear identification of validation objectives of experiment, and way to conclude on these objectives, to be able to:**
  - **Derive metrics from the validation objectives (top-down approach)**
  - **Go back from data collected to the validation objectives (bottom-up approach)**
  - **Possibly, relate some metrics to different validation objectives**
  - **Combine some metrics to get information on complex validation objectives**

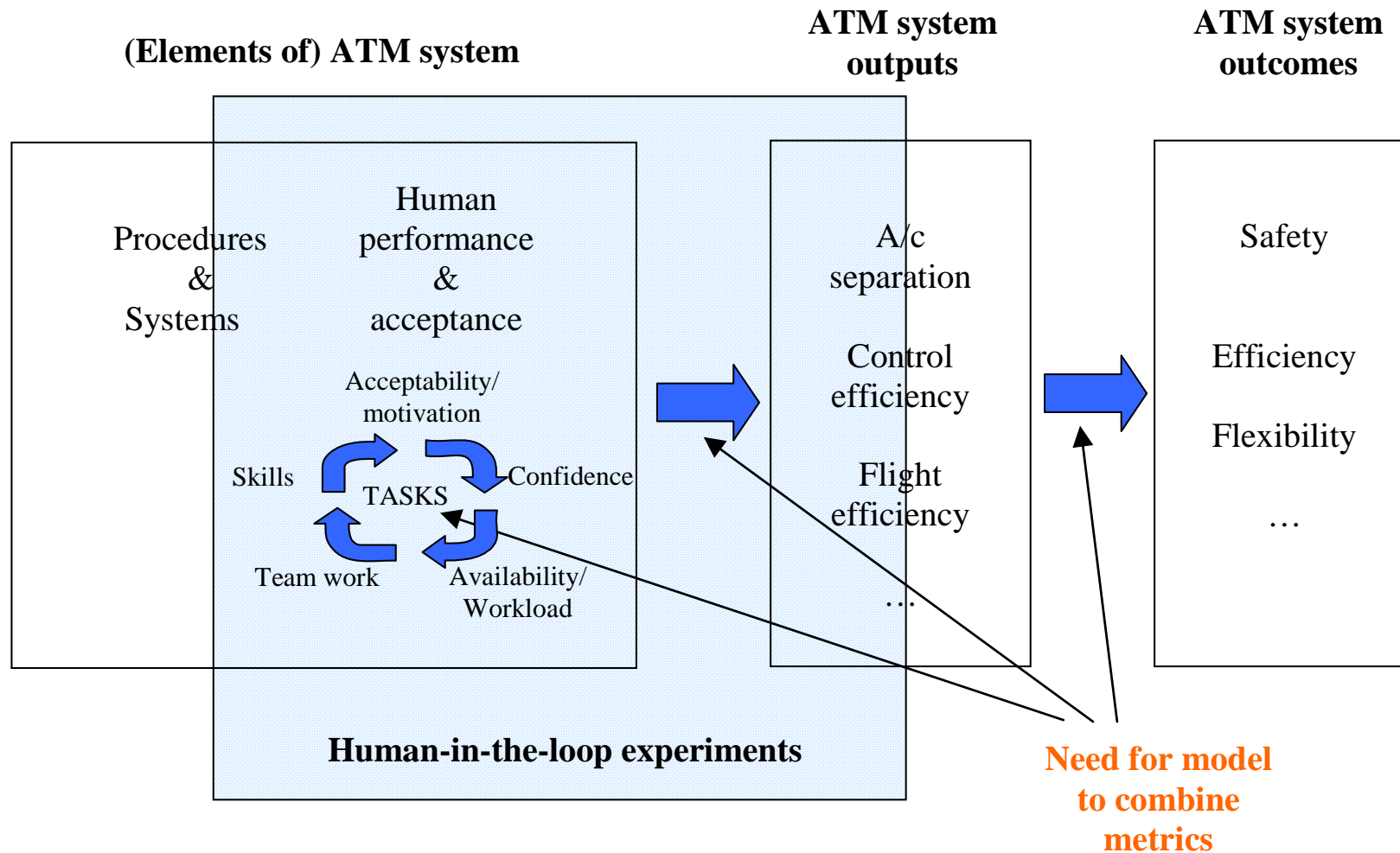
# Consolidation of results (2/3)

- ◆ **To support identification of relevant metrics and their analysis:**

- **Need for identification of expected benefits from concept**
- **Need for a model of expected functions and performances of (human and technical) components of the new ATM system**
- **Need for a model of human performance within the new ATM system**



# Consolidation of results (3/3)



# Conclusions

---

- ◆ **A first step towards definition of a common framework for metrics:**
  - **Some general guidelines for experiment design**
  - **Elements for consideration when defining validation objectives and metrics for experiments**
  - **Elements of discussion about relevance of metrics related to:**
    - **Human performance (and acceptance)**
    - **Efficiency (either control or flight perspective)**
    - **Safety**
  
- ◆ **Need for further work:**
  - **To refine the set of relevant performance areas and associated metrics measurable through human-in-the-loop experiments**
  - **To investigate the need for models (at various levels) to support consolidation of experiments design and results**

# References

---

- ◆ **‘Air-Ground Integration Experiment’, DOT/FAA/CT-TN02/06, January 2002**
- ◆ **‘EACAC 2000 Real-Time Experiments’, EEC Report version 5.4, October 2001**
- ◆ **‘INTEGRA Metrics in EACAC’, Final Report – EATMP edition 0.A, February 2001**
  
- ◆ **Comparison of EACAC and AGIE metrics using CARE/ASAS #2 framework, Eurocontrol SCS, CEAMCA final report, version 1.0, September 2002**