	HUMAN Model-based Analysis of Human Errors during Aircraft Cockpit System Design	

Error Types and Error Production Mechanisms

Full Title:	D1.2 - Error types and error production mechanisms in modern cockpits in an ATM environment.
Document ID:	HUMAN/WP1/AIF/D1.2
Nature:	Deliverable
Classification:	Public
Version:	V1.0 Saved: 21/10/2008
Parts & Classifications:	Main document & Annexes (Public)
Work Package:	WP1, WP2, WP3, WP4, WP5
Milestone:	M1
Document Timescale:	Project Start Date: March 1, 2008
Start of Document:	T0+1
Final version due to:	T0+6
Time now:	T0+8
Scope:	The document describes a list of error types, and associated error production mechanisms of particular relevance in modern cockpits in an ATM environment, and how this list has been obtained. These error types and error production mechanisms are those on which HUMAN will now focus, in order to found the cognitive model-based Human Error prediction techniques and their integration in system design methodologies.
Keywords:	Taxonomy, Human error
Issue Date (dd/mm/yyyy):	21/10/2008
Compiled:	Denis Javaux (AIF)
Authors:	Denis Javaux (AIF) Jan-Patrick Osterloh (OFF), Andreas Lüdtke (OFF) Tina Mioch (TNO), Rosemarijn Looije (TNO), Tijmen Muller (TNO)
Technical Approval:	Denis Javaux (AIF)
Issue Authorisation:	Andreas Lüdtke (OFF)

DISTRIBUTION LIST		
Copy type[1]	Company and Location	Recipient
C, E	Stephanie Stoltz-Douchet European Commission DG RTD - Directorate H - Transport CDMA 4/138 1049 Brussels	Stephanie Stoltz-Douchet
T	All HUMAN partners	All HUMAN partners

[1] Copy types:
M = Master copy,
E = Email,
C = Controlled copy (paper),
D = Electronic copy on disk,
T = TeamSite (Sharepoint)

RECORD OF REVISION			
Version	Date (dd/mm/yyyy)	Status Description	Author
0.1	14/04/08	• Initial version. • Definition of open issues (actions list) • Definition of document structure. • Definition of objectives for each section. • Authoring of : - § 1 : Objectives - § 4 (including 4.1, 4.2 & 4.3): Method - § 5.2 (including 5.2.1 & 5.2.2): Requirements for an error taxonomy relevant to HUMAN	DJ
0.2	28/04/08	• Modification of the main objectives of the deliverable: 1) the production of a general taxonomy of human errors in a scope to be defined (e.g., glass-cockpit aircraft), 2) the selection of a subset of this taxonomy to define the error categories to be specifically studied by the HUMAN project. • Subsequent update to the document structure. • Subsequent update to the objectives of each section. • Limited localized input to various sections	DJ
0.3	13/10/08	• Modification of the main structure, and associated objectives, to reflect the work done in May, June and July • Authoring of sections or instructions for other authors	DJ
0.5	20/10/2008	Insertion of material in allocated subsections	DJ, TiM, JPO
0.6	20/10/2008	Cleaning and proof-reading	JPO
1.0	20/10/2008	Final checks, proof-reading and cleaning	DJ

RECORD OF OPEN ISSUES				
Issue No.	Description	Added by	Assigned to	Status

1 Table of Contents

2 Abbreviations and Definitions

Abbreviation	Definition
EPM	Error Production Mechanism
ET	Error Type
HF	Human Factors
PSP	Physical Simulation Platform
SA	Situational Awareness
VSP	Virtual Simulation Platform

3 Introduction

The main objective of this deliverable is to define:

- The set of **Error Types** (ETs) that HUMAN is going to investigate. These ETs have to comply with a series of requirements, such as being relevant in the HUMAN target scope, interesting for HUMANs industrial partners (AIF and ALA), possible to be investigated in HUMAN (given the resources available), etc.
- The set of **Error Production Mechanisms** (EPMs) that HUMAN is going to investigate. Obviously these EPMs have to contribute to the occurrence of the selected Error Types (ETs), cf. above. They also have to comply with other requirements such as being investigable on the Physical Simulation Platform (PSP) or implementable on the Virtual Simulation Platform (VSP), be sufficiently mature (from the point of view of the scientific literature available) to be understandable and able to be modelled, etc.

3.1 Main concepts

A series of concepts, rather particular to HUMAN, must be introduced before proceeding further to the core objective of the deliverable:

- **Error Type (ET)**: An Error Type is a type of Human Error. The definition of what constitutes a Human Error is controversial and the object of many discussions or articles. Given the reliance of HUMAN on the notion of normative activities (T1.5), and its prominent role in the development of the cognitive model (T3.1) or the comparison between actual and predicted activities (T4.5), we tend to favour a definition that relies on the notion of deviation from prescribed or acceptable activities. Examples of Error Types are omission errors (when a prescribed step in a task is omitted) or fixation errors (when a human operator focus his or her attention on something – a piece of data, an external stimulus, a cognitive process – in a way that is detrimental to some task, e.g., maintaining situation awareness, or perceiving potentially safety threatening events). Determining the Error Types on which HUMAN will focus is one of the main goals of this deliverable, with the determination of the Error Production Mechanisms. Their selection is described in section 4.7.

- **Error Production Mechanism (EPM)**: They are simply the mechanisms by which Error Types (i.e., human errors) occur. The mechanisms are related to normal cognitive processes that do not perform optimally (for a large variety of exogenous and endogenous causes, such as high workload, fatigue, inappropriate design of information display, etc) and contribute to the occurrence of the Error Type. Each Error Type can therefore be produced by one or more Error Production Mechanisms. The distinction between Error Type and Error Production Mechanism also relate to the dissociation by E. Hollnagel between the phenotype of Human Error (Error Type) and the genotype of Human Error (Error Production Mechanism). See Hollnagel (1993). Examples of Error Production Mechanisms are biases (usually related to heuristics working outside of their domain of competence), loss of information in working memory (which may lead to many different Error Types), etc. Determining the Error Production Mechanisms on which HUMAN will focus is one of the main goals of this deliverable, with the determination of the Error Types. Their selection is described in section 4.7 of this document.

- **HUMAN target scope**: HUMAN aims at investigating Human Errors, by modelling and predicting them. This however cannot be achieved in a fully general way: there are simply too many Error Types (ETs), and Error Production Mechanisms (EPMs) contributing to them. Because the resources of HUMAN are limited (time, personnel, facilities, etc), we must focus on an operational domain, in which we will investigate, model, and predict a subset of Error Types and Error Production Mechanisms (those most relevant to the domain in question). This operational domain in which we want to invest our efforts is the target scope of HUMAN, and it must be delineated clearly, so that the output of HUMAN is interesting and relevant for the industry, but, on the other hand, still manageable within the framework of the project. Selecting a too large target scope would threaten our capability to correctly address the interesting Error Types and Error Production Mechanisms that may be involved in its operation. The determination of the target scope of HUMAN is described in section 4.1.

One notes that the dissociation between ETs and EPMs is not fully clear – nor easy. EPMs themselves can typically be decomposed into simpler EPMs, which could be therefore considered as mechanisms that explain the parent EPMs. The relation is the same than with ETs, and the parent EPMs could therefore be considered as ETs themselves. In the scope of HUMAN, and this document, we will simply consider an EPM as a mechanism that contributes to a deviation from a prescribed or acceptable way of performing, be it an ET or an EPM itself.

The determination of interesting and relevant Error Types and Error Production Mechanisms is absolutely central to HUMAN: the main goal of HUMAN is to investigate human errors and implement a cognitive model of a crew that allows predicting them. ETs and EPMs are critical steps towards this endeavour.

Determining the ETs and EPMs on which HUMAN is going to focus has to be done early in the project, since most work-packages (WP1 to WP5), with the exception of WP6 (Dissemination and Exploitation) and WP7 (Project Management), will have to tune their effort specifically towards them, once they are known.

3.2 Main methodological approach

A general methodological principle has been used for the determination of the Error Types (ETs) and Error Production Mechanisms (EPMs) on which HUMAN is going to focus.

Each time a selection has to be made (for example within a set of possible Human Error taxonomies, section 4.4, or within a set of possible ETs or EPMs, section 4.7), we have to

- determine the requirements we want the selected objects (e.g., taxonomies, or ETs and EPMs) to comply with,
- determine the set of potential objects where the selection will be performed (e.g., a set of taxonomies or a set of ETs and EPMs),
- select (or filter) the final objects (or solutions) by considering each potential object (e.g., taxonomies, or ETs and EPMs) and determine if it matches – or not – the requirements. The subsets of objects that match the requirement constitute the set of selected objects.

This principle has thus been used for the selection of HUMAN source Human Error taxonomies (section 4.4) and the selection of HUMAN final Error Types and Error Production Mechanisms (section 4.7). Its main benefit is:

- to consider the set of possible solutions in a very **exhaustive way**.
- to select the solutions, based on a series of **explicit requirements**.

4 Method and results

To reach the goal of determining the most relevant Error Types and Error Production Mechanisms for HUMAN, we have proceeded in a series of steps that progressively narrow the focus of interest, with decisions made at each stage based on explicit rationale considerations.

We will first determine the target scope in which we want HUMAN to investigate Error Types and Error Production Mechanisms (section 4.1). We will then determine in section 4.3 a series of Human Error taxonomies considered as potentially interesting for HUMAN. We will select a subset of them (see section 4.4), the

HUMAN source taxonomies, by exploiting a series of previously defined requirements (section 4.2). These source taxonomies will then be used as sources (hence their name) of potentially interesting (or relevant to HUMAN) Error Types and Error Production Mechanisms. The selection of the ETs and EPMS on which HUMAN will focus (HUMAN final ETs and EPMs) will be performed within these taxonomies (section 4.7), based on another series of dedicated requirements for ETs and EPMs (section 4.5). We will then describe these "final" ETs and EPMs, to make them fully understandable and usable in the framework of all HUMAN WPs and tasks where they have to be considered, e.g., defining generic experimental scenarios in task T1.4 that allows the investigation of these ETs and EPMs, on both the Physical Simulation Platform and the Virtual Simulation Platform.

4.1 Determination of HUMAN target scope

The selection of the scope in which we want to investigate Human Error taxonomies, Error Types and Error Production Mechanisms is particularly important. The Error Types and Error Production Mechanisms in peculiar will then become the central focus of HUMAN. These selections will therefore strongly influence the final output of HUMAN, and the range of ETs and EPMs the error prediction methodology will be able to deal with.

Several possible scopes have been considered, ranging from very large to very focused:

- Dynamic situations (i.e., any situation where one or more human agent is in charge of controlling a dynamic system)
- Complex dynamic situations (e.g., nuclear power plant, electric network, …)
- Vehicle control (e.g., aircraft, car, ship, …)
- Aircraft control (e.g., single engine aircraft, airliner, …)
- Modern glass-cockpit aircraft control (e.g., control of 4D flight management system)
- More restricted scopes linked to HUMAN application domain and technical competencies (i.e., in the aviation domain)

The final scope selected for the error taxonomies is "**The modern cockpit in a future ATM environment**", since it obviously relates to the overall scope of HUMAN itself, and since aiming at something larger (e.g., other vehicle types than aircraft) could be too ambitious and yield selected Error Types and Error Production Mechanisms which, despite being very interesting, are beyond the modelling and investigation capabilities of HUMAN.

4.2 Definition of requirements for interesting Human Error taxonomies in HUMAN target scope

Having determined the scope where to search for potentially interesting Human Error taxonomies, one now has to identify requirements for selecting those that are indeed relevant for HUMAN.

The following requirements have been defined:

- the taxonomies have to be relevant within the defined scope,
- the taxonomies should cover the most relevant error types (ETs) within the defined scope,
- the taxonomies should include observable error characteristics (phenotypes) for each Error Type (ET) or the taxonomies should focus on understanding the cognitive process involved in the production of human error (genotypes) and the associated Error Production Mechanisms (EPM),
- the taxonomies should refer to ETs and EPMs that have either a significant frequency of occurrence, or whose occurrence is particularly safety threatening (i.e., covering the most frequent and the most dangerous ones),
- taxonomies with strong theoretical or methodological foundations should be preferred to those that do not display these characteristics,
- well-established and well-tested (i.e., already successfully applied) taxonomies should be preferred to new ones, especially if the later do not display strong theoretical and/or methodological foundations.

4.3 Determination of interesting Human Error taxonomies in HUMAN target scope

To determine interesting Human Error taxonomies, among which we will select a series of specific "source" taxonomies (section 4.4) more dedicated to HUMAN, we will proceed in two steps:

4.3.1 Investigation of literature on Human Error and Human Error taxonomies in HUMANs target scope

The first step in investigating human error and human error taxonomies was collecting and evaluating existing literature on these topics. We took several weeks searching and collecting papers over human error taxonomies and human errors in general, to get a better acquaintance of these subjects in our newly identified scope (section 4.1).

To make sure that we were as exhaustive as possible in the available time we

- did not limit our search to the aviation domain, but also included other safety critical domains, such as the medical domain.
- collected literature about the relevancy of errors to the aviation domain, e.g. accident reports.

- collected literature about the frequency of occurrence of errors in the aviation domain, e.g. literature reviews of past accidents.

The literature was collected on Sharepoint to make is accessible to each partner. In addition, for the literature that was deemed by the partners to be the most relevant and interesting, small summaries were written to make it even more accessible to everyone.

4.3.2 Identification of potentially interesting Human Error taxonomies in HUMAN target scope

The second step has been to list all human error taxonomies found in the material previously gathered (section 4.3.1). These taxonomies constituted our main starting point. They can indeed massively contribute to reveal interesting Error Types already investigated by the Human Factors' research community or encountered during operational life in past and contemporary cockpits, and hence help us identifying the Error Types (and associated Error Production Mechanisms) HUMAN will investigate in the future.

The literature (books, reports, scientific papers, etc.) on these taxonomies has therefore been collected and discussed altogether. The main source used has been a very interesting and exhaustive error taxonomy literature survey, conducted by EUROCONTROL during the HERA project (Isaac, A.; Shorrock, S. T.; Kennedy, R.; Kirwan, B.; Andersen, H. & Bove, T.; 2002). The authors of this document categorised the taxonomies by their different foundations:

- Task-based Taxonomies mostly describe lists of "External Error Modes" (EEMs) which refer to the structure and elements of the external human task and classify the overt characteristics of the error. Examples for this taxonomies are (Swain, A.D.; 1982) and (Swain, A.D. & Guttmann, H.E.; 1983).
- Communication System Models and Taxonomies have been developed since the forties. Most of these models deal with mass communication, and are not primarily models of cognition. However, some of the models can be used to model communication within HUMAN, e.g. the communication model of (Shannon, C.; Weaver, W.; 1949).
- Information Processing Models and Taxonomies examining human performance by attempting to trace the information flow through several processing stages from information input to response output, e.g. (Wickens, C.D.; 1991).
- Symbolic Processing Models and Taxonomies regard humans and computers as general purpose symbol manipulating systems, and are closely related to the information processing tradition, and related to artificial intelligence and cognitive science. Taxonomies in this class are e.g. (Norman, D.A.; 1983), (Reason, J.; 1990) and (Rasmussen, J.; 1986).

- Other Models and Taxonomies bundles other taxonomies, not fitting in the above schema, e.g. the Situation Awareness Error Taxonomy of (Endsley, M.R.; 1999).

In addition to the taxonomies described in the EUROCONTROL document, additional taxonomies have been collected from other sources, e.g.:

- (Orasanu, J. & Martin, L.; 1998) describes errors in decision making.
- (Hollnagel, E.; 1998) and his CREAM method.
- (Shappell, S. A. & Wiegmann, D. A.; 2000)
- (Zhang, J., Patel, V. L., Johnson, T. R. & Shortliffe, E. H.; 2004)

The taxonomies above constitute the set of potentially interesting taxonomies in which we intended to select a subset of "source" taxonomies (section 4.4) of particular interest for HUMAN, based on the requirements defined in section 4.2.

The intention is to investigate these source taxonomies in great detail, and fully document the Error Types and Error Production Mechanisms they involve (section 4.6). These documented Error Types and Error Production Mechanisms will then be used as the source material to perform a final selection, based on a series of requirements defined in section 4.5, and identify the final ETs and EPMs HUMAN will focus on in the future. The selection process is described in section 4.7.1 and the resulting ETs and EPMs in section 4.7.2.

4.4 Selection of source Human Error taxonomies

The selection of the source Human Error Taxonomies particularly relevant to HUMAN, within its target scope, has been achieved based on our expertise in the field of Human Errors, on the documentation of the taxonomies in the EUROCONTROl document (Isaac, A.; Shorrock, S. T.; Kennedy, R.; Kirwan, B.; Andersen, H. & Bove, T.; 2002), and on the requirements specified in section 4.2.

Each partner involved in this task (OFF, AIF* and TNO) identified, independently of each other, the most relevant Human Error taxonomies, in accordance with the requirements mentioned in section 4.2. By doing this independently, we attempted to ensure that the selection process was exhaustive and that no interesting candidate was missed. Noticeably, most of the taxonomies finally selected were selected by all partners. We discussed the proposed taxonomies and their associated papers during several telephone conferences, taking the requirements explicitly into account. At the end, we decided to add to the commonly chosen taxonomies some of the taxonomies chosen by only a single partner, as this was in line with the requirement of being able to cover most of the relevant Error Types. Adding 'extra' taxonomies to the selection could have no negative effect on our final objective (the determination of the "final" ETs and EPMs), on the contrary, except in slightly increasing our effort, something we were ready to do.

After these discussions, we all agreed on the following "source" error taxonomies:

- (Swain, A.D.; 1982) and (Swain, A.D. & Guttmann, H.E.; 1983). This taxonomy is very relevant within the defined scope, and includes observable error characteristics (phenotype).
- (Rasmussen, J.; 1986). This taxonomy is very relevant within the defined scope, and focuses on understanding the cognitive process involved in the production of human error (genotypes).
- (Reason, J.; 1990). This taxonomy is very relevant within the defined scope, and focuses on understanding the cognitive process involved in the production of human error (genotypes).
- (Norman, D.A.; 1983). This taxonomy is very relevant within the defined scope, and focuses on understanding the cognitive process involved in the production of human error (genotypes).
- (Endsley, M.R.; 1999). This taxonomy is very relevant within the defined scope, and focuses on understanding the cognitive process involved in the production of human error (genotypes). Processes involved in creating and maintaining Situation Awareness (SA) have been described by Endsley in her conceptual SA model. Based on her model she defined potential errors in cognitive processing that may lead to an inadequate or even incorrect SA.
- (Orasanu, J. & Martin, L.; 1998). This taxonomy is very relevant within the defined scope, and focuses on understanding the cognitive process involved in the production of human error (genotypes).

The list above not only covers most of the categories of taxonomies in the EUROCONTROL document, but also additional ones, mentioned in separate literature. The only category we decided to leave out is the Communication models and taxonomies, since we had no communication model for HUMAN at that stage. This model is only to be chosen during WP3, for the matter of crew communication. The early development of a communication model in the framework of this deliverable would have required too many resources and would not have been possible in WP1.

To reduce the risk of ignoring or even rejecting important ETs and EPMs, it was decided to conduct a survey with other HUMAN experts, namely human factor experts from AIF and ALA, as well as pilots. They were asked to comment on the completeness of the selection and whether the Error Types in the source taxonomies occured frequently in the cockpit. The experts approved the selection and did not propose any additional taxonomy (in the pool of potentialy interesting taxonomies determined in section 4.3.2), nor Error Types. They also confirmed that the frequency of occurence of the Error Types was indeed variable, and dependent on the Error Type itself.

Cross checks with the FAA error list (FDAI Database[2]) were also performed at this stage to confirm that the list of source taxonomies selected for further investigation was indeed fully appropriate.

4.5 Definition of requirements for the selection of HUMAN final Error Types and Error Production Mechanisms

The main objective of this deliverable is to determine the sets of Error Types (ET) and Error Production Mechanisms (EPM) on which HUMAN will specifically focus, and on which the observations on the Physical Simulation Platform (PSP) and predictions on the Virtual Simulation Platform (VSP) will be achieved.

These particular ETs and EPMs are central to HUMAN and to the conduct of most of its work-packages and tasks. To denote them, we will use the term **"final ETs and EPMs"**, since their determination is the final goal of this document (the impatient reader will find them in section 4.7.2).

The source Human Error taxonomies selected (section 4.4) as particularly relevant in the scope of HUMAN (section 4.1) contain a series of potentially interesting ETs and EPMs. The identification of these taxonomies precisely aimed at providing this raw material: each ET and EPMs contained or referred to in these taxonomies must be considered in great detail, to decide if it has to belong to the final ETs and EPMs.

To perform this selection, we have defined a series of requirements that each final ET or EPM must satisfy. The definition has been performed in two steps: determining a series of elementary requirements derived from the relations between the final ETs and EPMs and HUMAN's work-packages and tasks (section 4.5.1), and then aggregating them into a list of final requirements (section 4.5.2).

4.5.1 Relations to work-packages and tasks

To determine the requirements the final ETs and EPMs have to satisfy, we have considered each work-package – and associated tasks – to see how they relates to these ETs and EPMs, and possibly impose specific needs or constraints on them.

---

[2] http://www.flightdeckautomation.com/qs-issues.aspx

4.5.1.1 WP1. Requirement Analysis

Task 1.1: To define a target system (FMS 4D)

The target system must be prone to some or all of the final ETs and EPMs.

The final ETs and EPMs must particularly cover the ETs and EPMs prone to occur on the target system.

Task 1.2: To propose evolutions of system functionalities and specifications for their integration in the simulator

Same as above (task 1.1).

Task 1.3: To write design requirements and standards related to the AHMI

The final ETs and EPMs have no interaction with this task.

Task 1.4: To produce generic experimental scenarios

The generic experimental scenarios must allow – and possibly increase the probability of occurrence – of some or all of the final ETs and EPMs.

The final ETs and EPMs must be likely to occur in the kind of envisioned generic experimental scenarios (accessible in the scope of the resources allocated to HUMAN for the development and implementation of experimental and virtual sessions on the physical and virtual simulation platforms).

Task 1.5: To specify the normative pilot activities, related to the chosen generic scenarios

It will be interesting to document the normative pilot activities produced during this task with the different (final) ETs and EPMs that may occur during their performance, and possibly when/where and why (providing hints for T4.1 on the definition of working hypotheses in WP4).

Task 1.6: To define basic capabilities for the cognitive architecture

The basic capabilities of the cognitive architecture must obviously be compatible with the final ETs and EPMs we intend to study: they are the result or the expression of the non optimal functioning of some cognitive mechanisms and these mechanisms must therefore be implementable in the cognitive architecture.

4.5.1.2 WP2. Preparation of the Physical Simulation Platform

Task 2.1: To implement the target system according to WP1 requirements in form of a physical prototype

Same as task 1.1 and 1.2 above.

Task 2.2: To integrate the target system prototype in the GECO simulator

Same as task 1.1 and 1.2 above.

Task 2.3: To define and implement data analysis techniques

The data analysis techniques will have to be able to detect, interpret and understand the final ETs and EPMs occurring either physically (on the physical simulation platform) or virtually (on the virtual simulation platform).

The error taxonomy should therefore mostly target final ETs and EPMs that can be detected, interpreted and understood, given the observational, data gathering and analysis capabilities accessible to the HUMAN project, in the scope of its allocated resources.

Task 2.4: To design generic pilot experiments

The generic pilot experiments must be likely to allow the detection, interpretation and understanding of the final ETs and EPMs, and possibly increase the probability of their occurrence.

The final ETs and EPMs must mostly target ETs and EPMs that can be detected, interpreted and understood in the kind of generic pilot experiments accessible to the HUMAN project, in the scope of its allocated resources.

Task 2.5: To prepare test schedule for recruited pilots

The error taxonomy has no interaction with this task.

Task 2.6: To check the acceptability of the physical simulation platform

The *physical* simulation platform will have to be capable of allowing the *actual* production, detection, interpretation and understanding of the final ETs and EPMs.

The final ETs and EPMs must mostly target ETs and EPMs that can *actually* be produced, detected, interpreted and understood on the *physical* simulation platform.

4.5.1.3 WP3. Preparation of the Virtual Simulation Platform

Task 3.1: To prepare the cognitive model for the virtual simulation

The development of the cognitive model will of course be strongly influenced by the final ETs and EPMs: the cognitive model intends to be able to predict their occurrence.

The final ETs and EPMs must mostly focus on ETs and EPMs that are considered predictable in the scope of the cognitive model developed in HUMAN, given the resources allocated to the project, and in particular the development of the cognitive model.

Task 3.2: To prepare a/c, environment, target system and scenario models for virtual simulation

Same as task 3.1.

Task 3.3: To implement the virtual simulation platform infrastructure

The error taxonomy has no interaction with this task.

Task 3.4: To check the acceptability of the virtual simulation platform

The *virtual* simulation platform will have to be capable of allowing the simulated (predicted) production, detection, interpretation and understanding of the final ETs and EPMs.

The final ETs and EPMs must mostly target ETs and EPMs that can *virtually* be produced, detected, interpreted and understood on the *virtual* simulation platform accessible to the HUMAN project, in the scope of its allocated resources.

4.5.1.4 WP4. Development and Validation of the Cognitive Model

Task 4.1: To define working hypotheses

The working hypotheses produced during this task will of course be strongly influenced by the final ETs and EPMs: the hypotheses will be related to their production (notably the condition of their occurrence).

Task 4.2: To define experimental scenarios

The experimental scenarios will be instantiation of the generic experimental scenarios. They will therefore exhibit the same relation to the final ETs and EPMs than the generic experimental scenarios. See task 1.4.

Task 4.3: To produce the actual activities of flight crews on the physical simulation platform

The actual activities produced and observed on the physical simulation platform will of course be related to the final ETs and EPMs, since the objective of the whole physical simulation platform is to be able to produce, detect, interpret and understand ETs and EPMs belonging to the final ETs and EPMs.

The final ETs and EPMs on the other hand must mostly address ETs and EPMs that can be produced, detected, interpreted and understood on the physical simulation platform.

Task 4.5: To compare actual and predicted activities

The techniques or tools used or developed for comparing actual and predicted activities will need to be able to detect the occurrence of the final ETs and EPMs, be they actual (on the physical platform) or predicted (on the virtual platform).

As already mentioned, the final ETs and EPMs themselves will have to mostly address ETs and EPMs that can be produced, detected, interpreted and understood on both the physical and virtual platforms, and then compared, in the scope of the resources available in HUMAN for buying or developing the techniques or tools used to do so.

Task 4.6: To derive requirements for the cognitive model

The requirements for the cognitive model will be directly linked to the final ETs and EPMs, since the main goal of the cognitive model will be to reproduce and predict some of the ETs and EPMs committed by humans. Most of the requirements for change will be directed to improving the predictive capabilities of the model (this is the principle behind the main development loop in WP4).

The final ETs and EPMs will mostly have to deal with ETs and EPMs that are within the reach of the development resources available in HUMAN, that is for which realistic requirements will be produced during this task.

Task 4.7: To improve and extend the cognitive model based on the requirements

Same as task 4.6 above. The improvements are a direct implementation of the requirements.

4.5.1.5 WP5. System Design Methodology

Task 5.1: To capture the current design process methodologies

The capture of the current design process methodologies, at both AIF and ALA, must pay special attention to the final ETs and EPMs: are the system designers aware of these error possibilities, do they take them into account when designing the systems, how, do they verify at some point that these error possibilities are indeed covered by the system, its user interfaces or the whole cockpit system (safety nets).

The final ETs and EPMs on the other hand have to include ETs and EPMs that are relevant to the type of systems designed by the manufacturers (with their current design process methodologies), and found in today or future cockpits (i.e., in HUMAN's scope, cf. section 4.1 above).

Task 5.2: To define a cognitive-model-based methodology for analysis of human error

The very objective of the cognitive-model-based methodology is to analyze – and if possible – predict the ETs and EPMs likely to be observed on an existing or new system, when used in general or specific operational circumstances. The cognitive-model-based methodology will therefore have to target the final ETs and EPMs, and in particular those considered as the most critical ones (e.g., in terms of frequency or potential impacts on the safety of the flight).

The final ETs and EPMs on the other hand have to include ETs and EPMs that are within the reach of the cognitive-model-based methodologies that can be developed during HUMAN, given the scope of resources available (and this applies in particular to the underlying cognitive model).

Task 5.3: To improve the current design process by integrating the methodology for analysis of human error

The improvements to existing system design methodologies brought by integrating the cognitive-model-based methodology for analysis of human error must mostly target the (final) ETs and EPMs that are the most relevant (notably in terms of safety) to the system designers (e.g. those with which they do not currently cope well) or that can be realistically integrated in the current system design methodologies (i.e., it would not be useful to address final ETs and EPMs, whose consideration during the design process would require so much resources from the system designers that it would not be realistically applicable to deal with them in the framework of an industrial system design process and manufacture).

The error taxonomy has to mostly address the ETs and EPMs which are the most relevant for the system designers, and in particular those with which they do not currently cope well.

4.5.1.6 WP6. Dissemination and Exploitation

The ETs and EPMs can be used as a basis for the preparation of a paper.

4.5.1.7 WP7. Project Management

The final ETs and EPMs have no impact on project management

4.5.2 Final requirements

The previous section (section 4.5.1) has allowed the determination of a series of elementary requirements the final ETs and EPMs must satisfy, by considering each HUMAN work package and task. Many of these requirements are redundant (i.e., the same requirement is obtained by considering different work packages or task) or semantically related. We will now aggregate them into a series of final requirements, that will be used to perform the selection (cf. section 4.7) of HUMAN's final ETs and EPMs, within the set of potential ETs and EPMs found in the HUMAN source taxonomies (section 4.4).

HUMAN's final ETs and EPMs must be:

- likely to occur within our target application, AFMS and AHMI, once it has been fully developed (T2.1)
- likely to occur on the Physical Simulation Platform (PSP), after integration of the target application (T2.2)
- observable (detectable, understandable) on the Physical Simulation Platform in the experimental settings that we intend to develop, given the amount of resources available in HUMAN for their development:
    - generic experimental scenarios (T1.4)
    - dedicated, instantiated experimental scenarios (T5.2)
    - observation and data gathering capabilities in the GECO (T2.3)
    - post analysis of observations and data gathered (T2.3)
- likely to be predictable on the Virtual Simulation Platform (VSP) by the cognitive architecture and error prediction methodology, given the amount of resources available in HUMAN for their development (WP3 and T5.3)
- relevant for the aviation domain and aeronautic industry, in particular the manufacturers of cockpit systems.

4.6 Documentation of all Error Types and Error Production Mechanisms in the source Human Error taxonomies

4.6.1 Dimensions for documenting the Error Types and the Error Production Mechanisms

In order to determine HUMAN's final ETs and EPM, all ETs identified in the source error taxonomies have to be described in detail according to a series of key dimensions. The dimensions are directly derived from the requirements described in section 4.5. By documenting each ET along these dimensions, we will be in position to determine if it complies – or not – with the above requirements and select HUMAN's final ETs.

In addition, for all ETs in the source taxonomies, we have identified the associated EPMs (i.e., the EPMs that may lead to the Error Type being considered). Key dimensions for documenting the EPMs have also been defined, based on the requirements of section 4.5. They will allow selecting HUMAN's final EPMs.

Error Types and Error Production Mechanisms are thus documented along a series of dimensions for:

- Defining and understanding them
- Allowing the selection of the final HUMAN Error Types and Error Production Mechanisms
- Deeply understanding the final HUMAN Error Types and Error Production Mechanisms, in particular to drive further choices (e.g., generic scenarios, cognitive model development, etc.) or informing the other partners of their specificities.

The following dimensions have been identified for documenting the Error Types in the source taxonomies (the list within brackets indicates the partners involved in the production of the corresponding data for the Error Types):

General description

- Description (AIF*, OFF, TNO)
- Example (AIF*, OFF, TNO)
- Theoretical affiliation(s), if any (AIF*, OFF, TNO)
- References, which includes articles, papers and books where the error type is described (AIF*, OFF, TNO)

Error type specification

- Family of error production mechanism(s) (optional) (AIF*, OFF, TNO)
- Error production mechanism(s) (AIF*, OFF, TNO)

Additional documentation for cognitive modelling

- Cognitive process(es) affected by the error type (AIF*, OFF, TNO)
- Cognitive step(s) in Action Production Model affected by the error type (AIF*, OFF, TNO)
- Factors increasing the likelihood of occurrence (AIF, AIF*, ALA, DLR, OFF, TNO)
- Factors decreasing the likelihood of occurrence (AIF, AIF*, ALA, DLR, OFF, TNO)

Documentation for the selection of HUMAN ETs and EPMs

(These five dimensions are described in more detail in the section 4.7.1 below on the selection process)

- Plausible frequency of the ET's occurrence with the AFMS/AHMI (in the GECO) (AIF, ALA, DLR)
- Relevance of the ET for cockpit crew in the ATM environment (AIF, ALA, DLR)
- Observability: detection of ETs (AIF*, OFF, TNO)
- Observability: understanding of the identified EPMs (AIF*, OFF, TNO)
- Predictability with the cognitive model of the identified EPMs (OFF, TNO)

The documentation of these dimensions for each error type in the source error taxonomies can be found in the annexes. The general description, specification and additional documentation is given in Annex I, the plausible frequency and relevance of the ETs can be found in Annex V, and the observability detection, observability understanding and predictability within the cognitive model in Annex II, Annex III and Annex IV respectively.

4.6.2 Example of documentation of an Error Type and its Error Production Mechanisms: "Entire task omitted"

An example of documentation of these dimensions for one of the error types, namely the ET *Entire task omitted* is given below. This error type belongs to the phenotypical, task-based taxonomy by Swain & Guttman (1983).

The tables describe the raw data, as they have been obtained from each contributing partners. The contributing partners had to provide ratings, and optional comments, in accordance with their field of expertise.

These raw data have then been aggregated for further processing (see section 4.7.1 for the detail).

General description of ET *Entire task omitted*

Description	AIF*, OFF, TNO
Occurs when the entire task to be achieved is omitted	
Example	AIF*, OFF, TNO
The crew do not execute the after takeoff checklist	
Theoretical affiliation(s), if any	AIF*, OFF, TNO
None	
References	AIF*, OFF, TNO
Swain (1982), Swain & Guttman (1983)	

In the following table, the Error Production Mechanisms that might induce the ET are identified.

Error type: Entire task omitted

Family of error production mechanism(s) (optional)	AIF*, OFF, TNO
Error production mechanism(s)	AIF*, OFF, TNO
• inadequate planning • learned carelessness/frequential simplification • loss of information in working memory • ineffective prospective memory • routine capture	

Additional documentation for cognitive modelling

Cognitive process(es) affected by the error type	AIF*, OFF, TNO
Plan execution, procedure execution	
Cognitive step(s) in APM affected by the error type	AIF*, OFF, TNO
Not defined	
Factors increasing the likelihood of occurrence	AIF, AIF*, ALA, DLR, OFF, TNO
Not defined	
Factors decreasing the likelihood of occurrence	AIF, AIF*, ALA, DLR, OFF, TNO
Not defined	

Documentation for selection of final HUMAN error types

Plausible frequency of occurrence with AHMI (in GECO)	
DLR	rare 1 **2** 3 4 5 frequent
ALA1	rare **1** 2 3 4 5 frequent
ALA2	rare 1 **2** 3 4 5 frequent
ALA3	rare **1** 2 3 4 5 frequent
ALA4	rare **1** 2 3 4 5 frequent
AIF1	rare 1 2 3 **4** 5 frequent
AIF2	rare 1 2 3 **4** 5 frequent

Relevance for cockpit crew in ATM environment	
DLR	Not relevant 1 2 3 **4** 5 relevant
ALA1	Not relevant 1 2 3 **4** 5 relevant
ALA2	Not relevant 1 2 3 **4** 5 relevant
ALA3	Not relevant 1 **2** 3 4 5 relevant
ALA4	Not relevant **1** 2 3 4 5 relevant
AIF1	Not relevant 1 2 3 4 **5** relevant
AIF2	Not relevant 1 2 3 4 **5** relevant
Comments (optional)	DLR
	The relevance is derived from the fact that there are some tasks like transitional checklists that cover items that had been performed in the past. Such as the gear retraction during take-off. The associated after take-off checklist is read considerably later. As tasks like gear retraction are so obvious, any checklist containing mostly items of similar obviousness bears the risk of omission.

Observability: detection	
AIF*	easy **1** 2 3 4 5 hard
	It obviously very easy: the whole task is missing (e.g. no take-off briefing)
OFF	easy 1 2 **3** 4 5 hard
	- Based on behaviour data - Eye movements (as part of behaviour data) - Comparison with normative behaviour
TNO	easy **1** 2 3 4 5 hard
	Seems to be easy to detect

For each EPM identified (see the Error Production Mechanism(s) field, there are five potential EPMs in the case of the Entire task omitted example), the two predictability and understanding dimensions are evaluated. For the sake of readability we will only present data on one of them in this report: learned carelessness/frequential simplification. The documentation for the four other EPMs for Entire task omitted can be found in Annex III and Annex IV.

Predictability with the cognitive model for learned carelessness/frequential simplification:

	SUM	**4**	**MEAN**	**2**	**FINAL RATING**	**2**
OFF	easy 1 **2** 3 4 5 hard					
TNO	easy 1 **2** 3 4 5 hard					
	Is partially implemented					

Observability understanding for learned carelessness/frequential simplification:

	SUM	13	MEAN	4,3	FINAL RATING	4,3
AIF*	easy 1 2 3 **4** 5 hard					
	It's hard (4), if not very hard (5) to prove that a simplified action pattern observed in the simulator is the by production of learned carelessness/frequential simplification. This can only be done with some confidence when the predictive capabilities we have regarding these implicit learning phenomena exactly predict the simplified pattern observed. In this case we may have good reasons to believe they are at play.					
OFF	easy 1 2 3 **4** 5 hard					
	• video based post-interviews • based on behaviour data (including historical data over several episodes) • eye movements • rule identification based on behaviour data					
TNO	easy 1 2 3 4 **5** hard					

4.7 Selection of final HUMAN Error Types and Error Production Mechanisms

4.7.1 Description of the selection process

As described above in section 4.6, all Error Types and Error Production Mechanisms involved in the source taxonomies have been documented along a series of dimensions, with integrated contributions from all HUMAN partners (except BCH) and experts.

The corresponding raw documentation data for the ETs and EPMs (for all source taxonomies), see section 4.6.2, have been aggregated and inserted in a single Excel file (cf. annex VI) to help with the selection process.

Five of the documentation dimensions reflect – or summarise - the requirements defined for ETs and EPMs (section 4.5), and are central to the selection process.

The first three dimensions concern Error Types:

- *frequency*: that is the frequency with which this ET is likely to occur in HUMAN target scope, according to AIF, ALA and DLR specialists. Evaluation of frequency has been obtained via a dedicated questionnaire (cf. section 4.6 above and appendix V). Values for frequency range from 1 (very rare) to 5 (very frequent).

- *relevance*: that is the relevance of considering this ET for HUMAN target scope, once again according to AIF, ALA and DLR specialists. The manufacturers are indeed in position to provide such a rating (e.g., an ET can be frequent, but there may be strong safety nets in today's cockpits to capture them, hence a low relevance. On the other hand, some ETs may be rare, but potentially detrimental to flight safety, and the safety nets in place considered less efficient. In this case, relevance is high). This data has also been obtained via a questionnaire (cf. section 4.6 and appendix V). Values for relevance range from 1 (not at all relevant) to 5 (very relevant).

- *detection*: that is the ability of the HUMAN consortium to detect this ET if it occurs on the Physical Simulation Platform (PSP). Some ETs are easy to spot (e.g., a required action has not been performed), but others are notably more difficult to detect (e.g., an error in the reasoning that leads to action, which is by a definition a covert, non observable behaviour, which must therefore be assess after the simulation runs, with a possibly low level of confidence). Values for the detection field range from 1 (very easy to detect) to 5 (very hard to detect), see section 4.6 (Observability: detection) and appendix II.

In a similar vein, two fields have been used to characterize the EPMs in the Excel file:

- *understandability*: this is the difficulty of understanding (or proving the existence) of this EPM when it occurs. EPMs resort from covert behaviours and they can only be detected by indirect means in most cases. Some EPMs, with strong overt manifestations, are not too difficult to spot (e.g., attentional capture, where the pilot or crew focuses its attention on something which is not fully relevant for the circumstance at hand: by considering the current circumstances and the appropriate reactions to it, and observing on the other hand that the crew persists at doing something else, it’s easy to state that attentional capture is occurring). Other are notably more difficult, if not impossible, to assess (e.g., ineffective prospective memory. Prospective memory is memory of things that has to be done in the future. Failure of prospective memory implies that something to be done later is not achieved when the time comes. This leads to an omission error (ET) but it’s very hard to prove that its underlying cause is indeed a failure of prospective memory. Many other causes could be involved. Hence the poor understandability of the ineffective prospective memory EPM). Ratings for understandability range from 1 (very easy to understand, i.e., to prove that it was indeed this EPM that was behind the observed ET) to 5 (very hard, if not impossible, to understand). The ratings for understandability have been provided by the HF experts involved in the project (AIF*, OFF and TNO). See also section 4.6 and appendix III.

- *predictability*: this is the difficulty (or importance of the effort) of implementing the EPM in the cognitive model, in order to accurately predict its occurrence on the Virtual Simulation Platform (VSP). Implementing an EPM in the cognitive model typically requires implementing specific basic features in the cognitive model and Virtual Simulation Platform (e.g., activation levels) and then programming the cognitive model so that it will fail in accordance with the EPM (e.g., attentional capture, routine capture). These features may pre-exist in the cognitive model (e.g., they may have already been there in the ISAAC cognitive model, the predecessor to HUMAN), be involved in already foreseen developments (e.g., concurrency) or on the other hand require completely specific and new developments (e.g., specific heuristics in the cognitive layer that lead to specific biases). These developments cost effort and predictability just evaluates that. Values for the predictability field range from 1 (very easy to predict = little or no development) to 5 (very hard to predict = a lot of developments, possibly beyond the resources available in HUMAN). The ratings for predictability have been provided by the partners with HF expertise or involved in the developments (TNO and OFF). See section 4.6 and annex IV.

Having all ETs and EPMs explicitly documented in the Excel file (cf. Annex VI), with associated numerical values allows to have a better grasp of the interest for HUMAN of selecting some of them, but also of the associated cost (in terms of development effort), difficulty and risk (of failure).

To help with the selection process, we have developed a series of formulas in the Excel file (cf. Annex VI) that precisely attempt to capture these two aspects: ET interestingness and EPM interestingness.

- An ET will be considered interesting if it is frequent, relevant for HUMAN and easy to detect on the PSP. A formula computes a value derived from the three associated ratings and considers the ET as "interesting" if the value obtained is above a given threshold.
- An EPM will be considered interesting if it is easy to understand and to predict. Another formula therefore computes a value derived from the two associated ratings and considers the EPM as "interesting" if the value obtained is above another given threshold.

A third formula is then used to recommend ETs for selection. An ET is recommended for selection if:

- It is "interesting"
- It has a least one of its contributing EPMs considered as "interesting". It wouldn't indeed be worth trying to investigate a given ET if its underlying EPMs are of little interest or not within HUMAN's reach.

These formulas therefore provide a first set of candidate ETs, recommended for selection, based on the ratings provided by the different partners in section 4.6 and its associated annexes, and our capability of addressing their underlying EPMs.

One should not consider however that these formulas perfectly indicate what should be selected or not. They do just provide suggestions. The final decision has been "man made", and involved all HUMAN partners, since the list of final ETs and EPMs is so central to HUMAN's further steps.

To help with this selection, two additional formulas have been defined:

- One of them computes the "Return on Investment" (ROI) associated with a given ET (and its underlying EPMs). ETs with a high level of interestingness and whose EPMs are interesting and easy to implement have a high ROI.

- The ROI value obtained for each ET is then processed to determine if the ET should be considered "priority 1" (i.e., it should be addressed in priority, typically during HUMAN cycle 1) or "priority 2" (i.e., it should be addressed a later stage, typically during HUMAN cycle 2, or even not addressed at all).

Once the Excel file had been filed with all the required data for the ETs and EPMs in the source taxonomies, and the associated formulas implemented and executed, it has been presented to the whole consortium, during the technical and management meeting held at TNO on July 8-9, 2008. Each ET and EPM has been presented and explained to the whole group, along with the recommendations contained in the Excel file (selection, ROI, priority). Discussions where then been held, in particular to determine what the selection of specific ETs and their associated EPMs implied for the project, in terms of developments on the VSP (cognitive model) and PSP (observation of crew behaviour). Many ETs and EPMs indeed require the same developments (e.g., implementing activation levels in the cognitive model) and performing these developments once therefore means that all of them will be addressable within HUMAN. Other ETs or EPMs on the other hand require very specific developments (i.e., peculiar to them only) and would therefore be very costly to chose.

The final list of ETs and EPMs was then selected during those discussions. It is described in the section below.

4.7.2 Final Error Types (ETs) and Error Production Mechanisms (EPMs)

In this section, the final error types and error production mechanisms are described.

4.7.2.1 Final Error Types (ETs)

The final error types have been selected, as described in section 4.7.1, according to several dimensions. In the following, the selected **eight** error types are given, with the aggregated ratings obtained for the three dimensions associated with the selection of error types (and involved in the computation of their interestingness). See Annex V.

In addition, the tables below also provide optional comments on the detection field, as well as the requirements that are imposed on the PSP and VSP, in terms of developments, features or devices, by the selection of the specific ET. These are important inputs for WP2 and WP3.

4.7.2.1.1 Entire task omitted

Description	Occurs when the entire task to be achieved is omitted

Frequency		**2.2**
Relevance		**3.6**
Detection		**1.7**
	Comment	This is obviously very easy to detect. The whole task performance is missing. This will be directly detectable from the data logs and eye tracking data obtained from the GECO after each run.
	Requirements	Data logs from the GECO Eye tracking data (for both pilots)

4.7.2.1.2 Step(s) in task omitted

Description	Occurs when one or more sub-steps in the task to be achieved are omitted

Frequency		**3.3**
Relevance		**3.6**
Detection		**2.3**
	Comment	This is slightly harder to detect than the omission of the whole task, because of the smaller granularity of what is missing. It should not be too difficult to detect provided **a normative description of the task** is available and confronted to the data logs and eye tracking data obtained from the GECO after each run.
	Requirements	Normative and formalized description of the task Data logs from the GECO Eye tracking data (for both pilots)

4.7.2.1.3 Attentional Slips

Description	Attentional slips occur when we fail to monitor the progress of routine actions at some critical point, usually when our situation or intention has just changed. Then, actions of habit for the old situation or intention override the new actions.

Frequency	**3.7**
Relevance	**4.0**
Detection	**3.3**
Comment	The occurrence of the error can be detected if a normative description of the task exists, by comparing the logs of actual behaviour with the prescribed norm. Eye tracking data can also be useful, to show that indeed the pilot or crew was not monitoring their own actions when the error occurred (though other types of monitoring exist than purely visual, e.g., tactile or haptic).
Requirements	Normative and formalized description of the task Data logs from the GECO Eye tracking data (for both pilots)

4.7.2.1.4 *Application of bad rules*

Description	Occurs when a rule inappropriate for the present circumstances is applied.

Frequency	**3.6**
Relevance	**3.9**
Detection	**3.0**
Comment	It's easy to detect that something has gone wrong, but here the observer has to prove that another rule than the one supposed to apply has been applied. So to detect this error, one has to be able to infer the wrong rule and recognize it among a set of possible candidates. This also means that the observer knows the whole repertory of rules available. The use of a wrong rule can also be confirmed in some cases by evaluating what the pilot or crew is attending to (e.g., if the wrong rules requires specific perceptions, not associated with other rules applicable in the current context).
Requirements	Normative and formalized description of the task Data logs from the GECO Eye tracking data (for both pilots) Repertory of applicable rules

4.7.2.1.5 *Biases*

Description	Biases describe ways in which humans deviate from rational choice. A rational agent exhaustively collects information necessary to assess a set of options and selects one options based on a complete assessment of decision criteria. Contrary to this humans tend to systematically use mental short-cuts or "heuristics". The biases considered in HUMAN are presented below, in the error production mechanisms section.

Frequency	**3.2**
Relevance	**3.7**
Detection	**4.3**
Comment	Typically very difficult to detect (i.e., to discover the error and prove that it resorts from biases). This cannot be done without offline analyses, after the simulation runs. Interviews, and video-based debriefings sessions are

	required for this, and may not be sufficient for some type of biases.
Requirements	Normative and formalized description of the task (to prove an error has occurred) Data logs from the GECO (to prove an error has occurred) Video-based recording of the runs (to prove this error is a bias) Video-based debriefings and analysis (to prove this error is a bias)

4.7.2.1.6 External activation errors

Description	Occurs when external events cause the activation of inappropriate schemas

Frequency	**3.2**
Relevance	**4.1**
Detection	**3.3**
Comment	Of variable difficulty. May be easy when the source of external activation is obvious (e.g., an incoming aircraft), but more difficult in other cases. For the first cases, eye-tracking data and data and behaviour logging are sufficient. For the latter ones, video-based debriefing sessions are needed, to investigate with the pilot or crew what external event prompted the error.
Requirements	Normative and formalized description of the task Data logs from the GECO Behaviour logs Eye-tracking data Video-based debriefings and analysis

4.7.2.1.7 Misordering of action sequences errors (Optional)

Description	Occurs when steps of an action sequence are misordered, skipped or repeated.

Frequency	**3.0**
Relevance	**3.4**
Detection	**2.3**
Comment	This is normally easy to achieve provided a definition of normative behaviour is available, alongside clear observations of actual behaviour. This allows to determine when a series of actions are performed in the wrong order.
Requirements	Normative and formalized description of the task Data logs from the GECO Behaviour logs

4.7.2.1.8 Failure to monitor or observe data

Description	Occurs when the human operator fails to monitor or observe data, although the information is directly available.

Frequency	**3.3**

Relevance		**3.7**
Detection		**3.0**
	Comment	Can be mildly difficult but not too hard, especially if the observer relies on an eye-tracking system, which allows him or her to prove that the pilot or crew has not perceived a piece of data which was there.
	Requirements	Normative and formalized description of the task Data logs from the GECO Behaviour logs Eye-tracking system

4.7.2.2 Final Error Production Mechanisms (EPMs)

The final error production mechanisms have been selected, as described in section 4.7.1, according to several dimensions. In the following, the selected **nine** EPMs are given.

Each EPM is first described in plain English.

The aggregated ratings obtained for the two dimensions involved in the selection of EPM (and in the computation of their interestingness, see 4.7.1 above) are then given:

- difficulty of understanding (or proving the existence) of this EPM when it occurs, together with the requirements that are necessary to make this possible
- difficulty of predicting this EPM (i.e., importance of the implementation effort in the cognitive model). Additional information is also provided on the nature of the implementation effort itself: at the level of the global architecture, the cognitive layer and the associative layer. This helps better pinpoint what will be needed in terms of development during WP3.

4.7.2.2.1 Learned carelessness/frequential simplification

Description	Learned carelessness is a procedural knowledge simplification mechanism, active on procedures in the associative layer. Frequential simplification is both a procedural and declarative knowledge simplification mechanism, active in both associative and cognitive layers. The two mechanisms are implicit learning mechanisms. In both cases, knowledge is simplified based on the frequency with which actions or facts are experienced, with no detrimental consequences. It can occur both in the associative and cognitive layers.

Understanding		**4.3**
	Comment	This can be considered very hard to achieve, because learned carelessness resorts from implicit learning, and applies on procedural knowledge in the associative layer, which is mostly unconscious. It is therefore difficult for the pilots to verbalize on the errors issued from learned carelessness and the mechanism by which they have been

	produced. Video-based analyses and after the fact interviews may only barely touch on them. The best indicator of the presence of learned carelessness as an EPM in an observed error will be the match between the predictions (of simplification) and the observations (predicted vs actual)
Requirements	Debriefing sessions with the pilots Video-based analyses during the debriefing sessions Formalized models of actual and predicted activities Computer-based tools for the comparison between actual and predicted activities.
Predictability	**2.0**
General architecture (framework)	
Comment	
Requirements	
Associative layer	
Comment	Learned carelessness is already implemented and will be improved.
Requirements	
Cognitive layer	
Comment	Frequential simplification in the cognitive layer is not considered at this stage (but could probably be implemented too).
Requirements	

4.7.2.2.2 Loss of information in working memory

Description	Occurs when information contained in working memory is lost and therefore not available for further use. It can occur both in relation with the associative and cognitive layers.

Understanding	**4.3**
Comment	It is usually extremely hard to prove that this EPM is behind an observed error. The problem is that pilots may be more or less aware of the loss of information in their working memory during the few seconds that follow the event (~"yes, I had that in mind, but completely forgot it") but will very certainly not be able to do this during after the fact debriefing sessions. There will be however some cases, where it is clear, for the audio and video recordings, that the pilot indeed considered an information, and then did not made use of it (which may not prove that it has been lost, it may have been deliberately discarded, but in this case this may be investigable during the debriefing sessions). An other option to investigate this type of EPM could be to resort to the "simulation freeze" technique, where the observers ask specific questions to the crew. This may allow to capture the limited instants where the pilot is aware of the information loss.
Requirements	Video-based analysis of the crew activities (without the crew) Video-based analyses during the debriefing sessions Interruptive "simulation freeze" technique during the runs
Predictability	**2.0**
General architecture (framework)	
Comment	This is not too difficult to achieve. It requires implementing a working

		memory, accessible by the cognitive layer and the associated layer, and of limited capacity. Also see discussion on working memories in Rasmussen's architecture).
	Requirements	A working memory accessible from the cognitive and associative layers. Loss of information could be implemented based on activation. Theory of activation in memory is needed. Could be hard to find formulas, tuning needed.
Associative layer		
	Comment	
	Requirements	access (input/output) to the working memory activation theory and formulas
Cognitive layer		
	Comment	Loss of information can occur because of limited working memory (+-7)
	Requirements	access (input/output) to the working memory activation theory and formulas)

4.7.2.2.3 Routine capture

Description	Occurs when the actions or cognitive processes performed by the pilot resort from frequently performed and well learned activities ("routine" activities"), in place of the actions or cognitive processes appropriate for the current circumstances. The routine sequence "captures" the control of ongoing actions or cognitive processes. It can occur both in the associative and cognitive layers.

Understanding		**4.0**
	Comment	This is mildly hard to achieve. The observer has to prove that crew activities deviated from the prescribed normative activities, and that the deviation resembles a routine pattern already performed by the crew. This can be achieved by the mere analysis of the actual activities, and possibly completed by a dedicated debriefing session (video-based) with the crew after the fact.
	Requirements	Formalized model of normative activities Formalized model of actual activities Statistical analysis of actual activities on a large number of runs, in order to extract routine activities, and formally model them Computer-based tool for comparing actual activities and normative activities, detecting the deviations, and trying to match them with pre-stored formalized models of routine activities
Predictability		**2.0**
General architecture (framework)		
	Comment	It could be interesting to implement features dedicated to the management of priorities or activation levels directly in the general framework and make them accessible to the associative and cognitive layers. See layers below.
	Requirements	Features dedicated to the management of priorities or activation levels in the general framework, e.g. in memory
Associative layer		
	Comment	Can typically be addressed by implementing priorities or activation levels associated with currently active procedures.

	Requirements	Priorities or activation levels associated with currently active procedures
Cognitive layer		
	Comment	Can typically be addressed by implementing priorities or activation levels associated with currently active plans.
	Requirements	Priorities or activation levels associated with currently active plans

4.7.2.2.4 Action priming (optional)

Description	Occurs when the likelihood of execution of an action is increased by some past event. Action priming is usually explained – and implemented – in term of activation level: the level of activation of the action is increased by an internal (e.g., a thought, an intention) or external (e.g., a perceptive event) event. It can also be faked by means of priorities. Action priming can occur in both the associative and cognitive layers.

Understanding		**4.3**
	Comment	This is very hard to investigate. Proving that action priming is the EPM behind an observed error is not easy. This is an EPM which acts at a completely implicit level, and of which most pilots are not aware. After the fact video-based debriefing sessions will barely be able to address this. We will only have hints (suspicions) on the involvement of this EPM, for example when the cues that primed the action are obvious (e.g., an external event to which the crew dedicated some attention).
	Requirements	Video-based analysis of actual activities Eye-tracking system allowing to determine what the crew is attending to
Predictability		**3.0**
General architecture (framework)		
	Comment	It could be interesting to implement features dedicated to the management of priorities or activation levels directly in the general framework and make them accessible to the associative and cognitive layers. See layers below.
	Requirements	Features dedicated to the management of priorities or activation levels in the general framework, e.g. in memory
Associative layer		
	Comment	Can typically be addressed by implementing priorities or activation levels associated with currently active procedures.
	Requirements	Priorities or activation levels associated with currently active procedures
Cognitive layer		
	Comment	Can typically be addressed by implementing priorities or activation levels associated with currently active plans.
	Requirements	Priorities or activation levels associated with currently active plans.

4.7.2.2.5 Attentional capture/selective attention

Description	Occurs when the pilot's attention is focused on specific tasks, in a way that is detrimental to the performance and monitoring of other ongoing tasks (possibly more important than the one the pilot focuses on). It

	can occur both in the associative and cognitive layers (but is most critical for the cognitive layer, since the associative layer requires little attention).

Understanding		**2.7**
	Comment	Proving that this EPM is involved is most of the times easy. Attentional capture is generally obvious for the mere (video-based) observation of the crew activities and of the things the crew is attending to (eye tracking system). This will only allow to deal with attentional capture on external events though. Attentional capture by high workload cognitive processes is more difficult to tackle (but since these are usually associated with spontaneous verbalization, they may not be too difficult to detect too).
	Requirements	Video-based analysis of the crew actual activities Eye-tracking system allowing to determine what the crew is attending to
Predictability		**3.0**
General architecture (framework)		
	Comment	It could be interesting to implement features dedicated to the management of priorities or activation levels directly in the general framework and make them accessible to the associative and cognitive layers. See layers below.
	Requirements	Features dedicated to the management of priorities or activation levels in the general framework.
Associative layer		
	Comment	Can typically be addressed by implementing priorities or activation levels associated with currently active procedures.
	Requirements	Priorities or activation levels associated with currently active procedures
Cognitive layer		
	Comment	Can typically be addressed by implementing priorities or activation levels associated with currently active plans.
	Requirements	Priorities or activation levels associated with currently active plans.

4.7.2.2.6 Salience bias

Description	Occurs when attention is focused on the most prominent object. It only occurs in the cognitive layer.

Understanding		**2.7**
	Comment	This can be achieved by determining what the crew is attending to, and possibly by confirming the hypothesis by means of video-based debriefing sessions
	Requirements	Eye-tracking system Video-based debriefing sessions
Predictability		**2.0**
General architecture (framework)		
	Comment	
	Requirements	Trigger of external events.
Associative layer		

Comment	
Requirements	
Cognitive layer	
Comment	
Requirements	Some knowledge is necessary of what prominent objects in the environment are, e.g. red objects, blinking objects, in addition to an according goal prioritization.

4.7.2.2.7 Recency bias (optional)

Description	Occurs when more recently acquired information is seen as more valuable. It only occurs in the cognitive layer.

Understanding	**4.7**
Comment	Extremely difficult to investigate. This involves information in working memory, in a conscious or nearly conscious way, in typically covert (i.e., not observable) cognitive processes, and of which little traces are available when the runs are analyzed with the pilots, after they have ended. Due to the possibly nearly conscious character of the "recently acquired information", the rather implicit (i.e., unconscious) character of this EPM, pilots do not remember after the fact what happened and why they erred.
Requirements	Eye-tracking system Video-based debriefing sessions
Predictability	**2.0**
General architecture (framework)	
Comment	
Requirements	Timestamp for each perception/belief that is held in memory, activation value of memory items accessible for components
Associative layer	
Comment	
Requirements	
Cognitive layer	
Comment	This bias can be seen on different levels: first, it is possible that information in memory cannot be contradictory. In this case, this bias might only have consequences for action determination (e.g. in deciding which action has priority). This can be done by either having fixed timestamps of memory items or that activation of beliefs in memory deteriorate (and thus implement this bias). It is also possible that contradictory information is kept in memory.
Requirements	

4.7.2.2.8 Simple omission

Description	Occurs when one or more actions are omitted from an action or cognitive processes sequence. It can occur in both the associative and cognitive layers.

Understanding		**3.3**
	Comment	It's relatively easy to prove that this "EPM" has been active (it is however more on the side of a phenotype than a true genotype or EPM). One simply has to have a specification of the normative activities and compared them with what is observed in the simulator.
	Requirements	Formalized model of normative activities Formalized model of actual activities Computer-based comparison of normative and actual activities, in order to determine when actions are omitted.
Predictability		**2.5**
General architecture (framework)		
	Comment	It could be interesting to implement a random failure mechanism for the percept and motor components, which then can be used as a mechanism for manual failure injection by the user. As an extension it could also happen in both layers, for all kinds of actions or processes of the layer.
	Requirements	Random failure mechanism for actions in the percept and motor component.
Associative layer		
	Comment	
	Requirements	
Cognitive layer		
	Comment	This can be implemented by having a random failure mechanism in the cognitive layer which makes all actions or processes in the cognitive layer prone to fail with a given (very low, e.g., <.01) frequency (without any identified cause).
	Requirements	Random failure mechanism for actions or processes

4.7.2.2.9 Distraction due to external events

Description	Occurs when some external event(s) prompt(s) the pilot to attend the event, and possibly produce a response. This distracts the pilot from the currently ongoing tasks he or she is involved in. This can occur in both the associative or cognitive layers.

Understanding		**2.7**
	Comment	It's normally easy to determine that this EPM has been active: the pilot or crew makes an error in reaction to an external event to which they pay attention. All these aspects are easy to observe.
	Requirements	Video-based analysis of crew activities Eye-tracking system for determining what the crew is attending to Data logging of "external event" (possibly scenario based) Formalized model of normative activities Formalized model of actual activities (including eye tracking data) Computer-based analysis of formalized models of normative and actual activities to determine: - that an error has occurred (deviation from normative activities) - that some external event occurred - that the crew paid attention to the external event - that the error occurred immediately, or not longer after, the external event occurred and the crew paid attention to it.
Predictability		**3.0**
General architecture (framework)		
	Comment	
	Requirements	Sophisticated perception (e.g. bottom-up attention + auditive component) as trigger of external events, then, reactive behaviour in associative layer or cognitive layer needed, plus priorisation or activation of goals
Associative layer		
	Comment	
	Requirements	Reactive behaviour (rules) Multi-tasking (task change) Task interruption Goal priorisation
Cognitive layer		
	Comment	
	Requirements	Goal prioritization

5 Summary

The objective of this deliverable was to define a set of ETs and EPMs that are relevant for HUMAN, and will thus be investigated during the project. The final ETs and EPMs obtained in this deliverable are of utmost significance, as they are used in nearly all work packages of HUMAN. We have defined a method that enabled us to make a selection without missing relevant ETs and EPMs, in a very systematic way, by relying on explicit procedures and decision criteria, as well as on input and judgement from all types of expertise in HUMAN.

The result is the list of ETs and EPMs that we are going to investigate in HUMAN:

- Error Types (n=8):
  - Entire task omitted
  - Step(s) in task omitted
  - Attentional Slips
  - Application of bad rules
  - Biases
  - External activation errors
  - Misordering of action sequences errors (Optional)
  - Failure to monitor or observe data
- Error Production Mechanisms (n=9):
  - Learned Carelessness/Frequential Simplification
  - Loss of information in Working Memory
  - Routine Capture
  - Action priming (Optional)
  - Attentional Capture/Selective Attention
  - Salience bias
  - Recency bias (Optional)
  - Simple omission
  - Distraction due to external events

6 References

- Amalberti, R. 1998. *Automation in Aviation: A Human Factors Perspective*.
- *Aviation Safety Reporting System Database (ASRS)*. 2008.
- Begosso, L. C. & Filgueiras, L. L. L. 2006. Human error simulation as an aid to HCI design for critical systems. Paper read at IHC '06: Proceedings of VII Brazilian symposium on Human factors in computing systems at New York, NY, USA.
- Endsley, M. R. 1999. Situation Awareness and Human Error: Designing to Support Human Performance. Paper read at Proceedings of the High Consequence System Surety Conference at Albuquerque, NM.
- Hollnagel, E. (1993) *Human Reliability Analysis: Context and Control,* London: Academic Press (CREAM) Elsevier Ltd.
- Hollnagel, E. (1998) *Cognitive Reliability and Error Analysis Method (CREAM)* Elsevier Ltd.
- Isaac, A., Shorrock, S. T., Kennedy, R., Kirwan, B., Andersen, H. & Bove, T. (2002). *Technical Review of Human Performance Models and*

Taxonomies of Human Error in ATM (HERA) (No. 1.0). Brussels, Belgium: EUROCONTROL (DIS/HUM).

- Isaac, A., Shorrock, S. T., Kennedy, R., Kirwan, B., Andersen, H. & Bove, T. (2002). *Short Report on Human Performance Models and Taxonomies of Human Error in ATM (HERA)* (No. 1.0). Brussels, Belgium: EUROCONTROL (DIS/HUM).
- Norman, D. A. (1983) Design rules based on analyses of human error. IN *Commun. ACM*, 26. 254— 258.
- Orasanu, J. & Martin, L. 1998. Errors in Aviation Decision Making: A Factor in Accidents and Incidents. Paper read at Proceedings of HESSD 98: 2nd Workshop on Human Error, Safety, and System Development.
- Rasmussen, J. (1982) Human errors: A taxonomy for describing human malfunction in industrial installations. IN *Journal of Occupational Accidents*, 4. 311— 333.
- Rasmussen, J. (1986) *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering* New York, NY, USA, Elsevier Science Inc.
- Rasmussen, J. (1987) Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IN *System design for human interaction*, . 291— 300.
- Rasmussen, J. (1997) Risk management in a dynamic society: a modelling problem. IN *Safety Science*, 27. 183— 213.
- Reason, J. (2000) Human error: models and management. IN *BMJ*, 320. 768— 770.
- Shannon, C. & Weaver, W. (1949) The Mathematical Theory of Communication. IN *Urbana, USA, University of Illinois Press*.
- Shappell, S. A. & Wiegmann, D. A. (2000). *The Human Factors Analysis and Classification System–HFACS* (Final Report No. DOT/FAA/AM-00/7). Washington, DC 20591: FAA - Office of Aviation Medicine.
- Shorrock, S. T., Kirwan, B. & Smith, E. 2003. Human error prediction in Air Traffic Management: A comparison of two approaches. Paper read at IBC Conference on Preventing Human Errors and Violations.
- Swain, A.D. (1982). Modelling of response to nuclear power plant transients for probabilistic risk assessment. Proceedings of the 8th Congress of the International Ergonomics Association, Tokyo, August, 1982.
- Swain, A.D. and Guttmann, H.E. (1983). A handbook of human reliability analysis with emphasis on nuclear power plant applications. NUREG/CR-1278, USNRC, Washington, DC 20555.

- Wickens, C. D. (1991) *Engineering Psychology and Human Performance* Harper-Collins College Div.
- Wiegmann, D. A. & Shappell, S. A. (2001). *A Human Error Analysis of Commercial Aviation Accidents Using the Human Factors Analysis and Classification System (HFACS)* (Final Report No. DOT/FAA/AM-01/3). Washington, D.C. 20591: FAA - Office of Aviation Medicine.
- Wiegmann, D. A., Faaborg, T., Boquet, A., Detwiler, C., Holcomb, K. & Shappell, S. A. (2005). *Human Error and General Aviation Accidents: A Comprehensive, Fine-Grained Analysis Using HFACS* (Final Report). Washington, DC 20591: FAA - Office of Aerospace Medicine.
- Zhang, J., Patel, V. L., Johnson, T. R. & Shortliffe, E. H. (2004) A cognitive taxonomy of medical errors. IN *J. of Biomedical Informatics*, 37. 193— 204.