International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

# Human Cognitive Bias Identification for Generating Safety Requirements in Safety Critical Systems

**Salah Ali, Aekavute Sujarae**

***Abstract – Safety critical systems are systems whose failure could result loss of life, economic damage, incidents, accidents or undesirable outcome, but it is not doubt that critical system safety has improved greatly under the development of the technology as the number of hardware and software induced accidents has been definitely reduced, but number of human deviations in their decision making found in each accident range remains more. We deeply reviewed traditional human error approaches and their limitations and propose approach of Human Cognitive Bias Identification Technique (H-CoBIT) that identifies, mitigates human potential cognitive biases and generates safety requirements during the initial phase of system Design. This proposed method, analyses the design of safety critical systems from a human factors perspective. It contributes system analyst, designers, and software engineers to identify potential cognitive biases (metal deviations in operator's decision-making process) during the system use. To ensure the validity of the proposed method, we conducted an empirical experiment to validate the method for accuracy and reliability comparing different experimental outcomes using signal detection theorem.***

***Keywords – Keyword: safety critical systems, cognitive bias, Human Reliability Analysis.***

## I. INTRODUCTION

Development of safety critical systems require more attention and analysis than any other information systems design. In safety critical systems, when the system fails, it leaves unforgettable outrages, such as death, environmental damages, and extensive economy loss [1]. Therefore, to prevent such these losses, it is necessary to build and conduct human cognitive bias analysis when designing such these systems (safety critical systems). Typically, failures of computer systems are contributed by many factors including hardware, software, and human operators. Many techniques and frameworks have been designed and proposed to analyze and prevent the errors triggered by above mentioned factors, and improved much more in the past decades, but still there is a need for analyzing and mitigating human triggered errors.

Manuscript received on March 15, 2020.
Revised Manuscript received on March 24, 2020.
Manuscript published on March 30, 2020.
* Correspondence Author
**Salah Ali***, Department of Science in Information Technology, Shinawatra University, Thailand.
**Aekavute Sujarae,** Instructor, Department of Science in Information Technology, Shinawatra University, Thailand.



Contemporaneous technology is considered by complexity, changing rapidly, and growing fast in size of technical systems that caused increasing concerns with the human involvement in safety critical systems. Undoubtedly that critical system safety has improved greatly under the development of the technology as the number of hardware and software induced accidents has been definitely reduced, but the number of human deviations in their decision making found in each accident range remains more as analyses of the major safety critical accidents during recent decades have concluded that human errors on part of system operators, managers, and designers have played a major role [2]. In literature there are many safety critical system failures contributed by cognitive biases, which left a massive tragedy. These accidents of safety critical systems include KLM flight 408 accidents, Three Mile Island Nuclear Power Plant incident, Air France 447 crash accident and many more safety critical system accidents, which have all been blamed on human errors[2]. In safety critical systems, it should be well analyzed potential risks to prevent future operators' failure and this needs to concentrate on not only external human errors but also psychological perspective specifically cognitive biases to stop errors stemmed from thoughts and beliefs that finally lead to poor and incorrect decisions. As there are many human error identification techniques in the field of current researches, still there is a need to be identified the root cause of human errors. Therefore, in this paper, we propose human cognitive bias identification technique (H-CoBIT) that helps system analysts, designers, engineers and all system stakeholders to identify potential cognitive biases and generate safety requirements. This approach will be conducted in the early phase of the system design to prevents operators' deviations stemmed from cognitive biases (mental deviation from the rational decisions to irrational decisions).

## II. SAFETY CRITICAL SYSTEMS

Safety critical system (SCS) is any system that will leave an extensive tragedy if it fails. For instance, failures of avionic systems may contribute loss of lives and economic damages, similarly, failures of nuclear or chemical power plants may also trigger life and environmental devastations, medical systems such as chemotherapy, and insulin pump systems may also cause undesirable outcome if the operators misuse them. Therefore, to prevent such failures we need to set and establish safety barriers in the first stage of development life cycle, by conducting a strong risk analysis on all safety facets of the systems.

Human Factor in safety critical system development Human factors in the development of safety critical systems are to study more about the system operators, for their cognitive capabilities and limitations to apply and interact with the systems. Human factor is a multidisciplinary approach that studies mental information processing in a psychology perspective[3]. The fundamental objectives of human factors are to prevent and reduce human centric errors that lead to undesirable consequences. the main goal of human factors also includes to increase the efficacy and safety constraints during operator interaction with critical systems. *Human Errors:* According to many definitions [3], [4], human errors are deviation from the required procedure and committing errors that is not what has been expected. There are many types of human errors based on the errors committed. For instance, skill-based errors, rule-based errors and knowledge-based errors are a good example of human error classifications and each reflects from specific position in mental information processing. For skill-based errors, operators have more knowledge about the task they are doing, and it is routine task, but still they make mistakes with slips or memory lapses. Therefore, such these errors are considered to be skill-based errors. In rule bases-errors, the operators do not follow the rules by disregarding the sequence of procedures or some other rules to be fulfilled[4]. According to[5], fundamental theory that human error types are categorized into three different ways (SRK) is to be understood how human cognitive deals with each error type. For instance, the skill-based errors are associated with human sensorimotor execution that execute lack of conscious control, effortless, and automatic. The second category, Rule bases error is representing a behavior that associates human perception pattern as it related how stimuli (rules, procedures) is interpreted and execute an appropriate action. The third category which is knowledge-based error type is associated in a condition that a person doesn't have any past experience or knowledge that he can apply for response. Therefore, human error classifications SRK has rigorous correlation with cognitive functions as each error type associates specific cognitive function. Doing one type of error, reflects a failure of specific cognitive function such as perception, memory and thinking.

## A. Cognitive Bias

Human cognitive bias are mental deviations from the rational decision to incorrect decisions [6]. As human brain has capability to interpret and process information from the outside world, then it has the limitations to perform such information processing. It may have misinterpretation data, retrieving data incorrectly, and distorting from the correct decision to incorrect decisions trapping "a cognitive bias". On January 1989, well known air crash accident occurred known as "Kegworth disaster". The aircraft fell down to the ground after one of its engines being shutdown incorrectly as pilots trying to make an emergency landing following an excessive engine vibration. These errors are characterized one of human cognitive biases (confirmation bias) to a poor cockpit design. [6] To understand cognitive biases, [7] proposed neural networks framework that enlightens the reason human brain systematically gets default to heurist decision making. This framework involves four fundamental principles, which are biological neural network characteristics. These characteristics of neural wetware are essential to all neural networks that occur throughout human brain, which consists of large number of interconnected neurons. These principles are *associative*, *compatibility*, *retainment* and *focus*.

"*Associative principles* are defined as the brain tends to seek associatively for the link, coherence, and patterns in the available information" [8].

"*Compatibility Principle* are also defined that associations are highly determined by their consistency and conformity with the momentary state and connectionist properties of the neural networks. i.e. we see, recognize, accept or prefer information according to its consistency with what we already know, understand, expect, and value" [9].

*"The Retainment Principle* states that when misleading information is associatively integrated, it is captured in the brain's neural circuitry, such that this cannot be simply made undone, erased, denied or ignored and thus will (associatively) affect a following judgment or decision" [8].

*"The Focus Principle* tell us when the brain gives full attention to and concentrates associatively on dominant information, i.e., dominant 'known knowns' that easily pop up in the forming of judgments, ideas, and decisions (availability heuristic biases). The fact that other possible relevant information may exist beyond is insufficiently recognized or ignored [8]."

There are more than hundred cognitive biases listed in the literature, but we focus on most common cognitive biases influence in safety critical systems.

*Confirmation Bias* **–** Confirmation bias is a tendency that people mostly seek information that confirms to their thoughts and beliefs [10]. Confirmation bias in well known in aviation domains as pilots form their own mental models based on their past experience and it has been associated with triggering many aviation accidents reports such as Kegworth, and air France 447.

*Attentional Bias* – attentional bias is a tendency the people focus on their attention on to specific aspect of their activity [11]. This bias has an effect on aviation domains as pilots mostly focus their attention into specific thing that they keep in mind. Such this example is air France 447 crash the pilot flying gave his attention to weather ahead.

*Attentional Tunneling* – this bias looks similar to attentional bias but slightly different, as it is defined that people tend to focus and allocate their attention to specific activities ignoring other channels of information and failing to perform the required activities[12].

*Optimism bias* – is the definition of the people who miscalculate the outcome of the situation, turning into positive [13]. These people overestimate the result will be okay and positively. This bias is popular in aviation where pilots turn sometimes the bad situation into good perceptually.

*Overconfidence* – as the name implies, it is the tendency people overestimate the situation on the environment. They tend to overestimate their ability to do something, their objectives to perform actions and so on. This bias takes part accidents in aviation domain [14].

*Anchoring bias* – This is a situation where people tend to rely on previously perceived information when making decisions because their judgement and decision making is affected by anchoring effects of information, which has been given and processed before the decision[15].

### B. Human Reliability Analysis

The notion of human reliability approaches was introduced in 1960, but the considerable majority of the methods for human factor evaluation, in terms of tendency to fail have been developed since the mid-80s. HRA approaches can be fundamentally categorized into two namely, first and second generation. Presently Human Reliability Assessment techniques or methods of the third generation understood as an evolution of the previous generations [4], are subject to research.

The first generation of HRA methods utilizes a simple error taxonomy. The techniques within this class were the initial tool developed to help risk assessors predict and quantify the probability of human error; they encourage the assessor to break down tasks into components and consider the potential influence of modifying factors such as; time pressure, equipment design, and stress. These compound analyses allow us to determine the potential for a nominal human error [16].

Second-generation in human reliability assessment methods such as, Cognitive Reliability and Error Analysis Method (CREAM) and TRACEr (technique for the retrospective and predictive analysis of cognitive errors), are based on human behavior models. So, they emphasize the qualitative description of human errors, relating to cognitive roots and human cognitive functions involved [17].

*Thea Approach:* The technique for human error assessment is one of HRA methodologies, that easier to apply
than the other existing HRA. This technique was developed for aim of system designers and engineers to use in the development of each cycle [18].

The main goal for developing this approach was that to have tools that can be used by non-human factor designers and system engineers. This approach engages cognitive error analysis using Norman's (1998) model of how humans execute actions. THEA apply sequence of questions in a checklist based on Norman's Model (goal, actions, plans, and interpretations or evaluations).

This approach is also applying scenarios that enable system analysts to describe the system being analyzed before any assessment is carried out. Therefore, the scenario makes the approach strong enough helping the analysts to describe potential errors as the scenario identifies, actions, environment and situation and the task being performed.

Since this approach is well structured approach, using series questions prompt in a checklist, it doesn't have error classifications or error modes, and this leads the system analysts to confuse which error type might occur in a specific scenario. The other drawbacks of this approach are that claiming it can be used by non-human experts but still it has not rigorous validation evidence of this approach [18].

*CREAM Approach:* Cognitive Reliability and Error Analysis Method (CREAM) is human error
Identification method that was developed by Hollnagel in response to the analysis of existing human reliability assessment approaches. This method can be used both prospectively, to predict potential human errors, and retrospectively, to analyze and quantify occurred human errors [19].

Cognitive Reliability and Error Analysis technique consists of a method, classification scheme and a model. It uses the model of cognition, which is Contextual Control Model (COCOM). This model focuses on how actions are chosen and assumes that the degree of controls that operators have over their actions are variable and also that the degree of controls to which operators hold determine the reliability of their performance.

The classification system of CREAM approach is contained of both phenotypes (error modes) and genotypes (error causes). Phenotypes and genotypes are further divided into detailed classification groups that are described in terms of general and specific consequents. CREAM technique also uses a set of common performance conditions (CPC) that are used by the analyst to describe the context in the scenario or the task, which is under analysis. Common performance conditions are like Performance Shaping Factor's (PSF) used by other HRA techniques [18].

CREAM has a cognitive model called CoCoM (Contextual Control Model), which is lacking to identify human cognitive bias. As it does not have clear decision-making framework, it is difficult to identify human cognitive biases which mostly stemmed form system 1 thinking or mental short cuts.

***TRACEr:*** TRACEr is human error identification
technique developed specifically for use in air traffic control (ATC). The TRACEr technique is represented in a sequence of decision-flow plans that contains eight error classification schemes, which are Task Error, Information, Performance Shaping Factors, External Error Mode, Internal Error Mode, Psychological Error Mechanism, Error detection, and error correction [19].

TRACEr technique can be applied both prospective and retrospective and it is based on a review of literatures of several domains, including experimental and applied psychology, human factors literature and communication theory. Existing human error identification methods were reviewed and research within Air Traffic Management (ATM) was conducted in the development of the method. TRACEr is not generic and developed specific domain (air traffic control).

In this method, it concentrates to extract internal error mechanisms using above mentioned error classification schemes and it appears unnecessarily over-complicated for what it is, a taxonomy-based error analysis tool. This approach claims to identify psychological error mechanisms but has not specific techniques to identify mental heuristics – mental shortcuts that make people to solve problems or/and make judgement quickly using system 1 thinking processing.

## III. H-CoBIT METHOD

In this paper, it is conducted an experiment, which is intended to validate the reliability and the consistence of the H-CoBIT (Human Cognitive Identification Technique) method. The experiment participated twenty-four graduate students with different academic backgrounds. They were taught how the method is applied, what the cognitive bias is and how it affects human decisions and also provided all materials needed to use for the method. The participants are divided into 3 groups (8 students in each) and provided specific design scenario on avionic system to be analyzed and requested to predict and identify potential human cognitive biases.

As shown below section IV, participants were asked to analyze aircraft Go-Around system design, which contains 3 scenarios and each group was asked to analyze all the 3 scenarios and identify the potential human cognitive biases using H-CoBIT method. This experiment is not only useful for the method reliability and consistence, but also validates developed H-CoBIT guidewords.

Materials given to the experiment participants consist of written scenarios, hierarchical task analysis drawn from scenarios, human error classifications, list of cognitive biases and its generic meaning (since they are not human factor experts), and H-CoBIT record table to document their identified results. To validate reliability and consistency of the proposed method, it is compared the prediction of the participants with the previous result obtained by the authors and used signal detection theorem to test and validate the correctness of H-CoBIT method. SDT is useful for testing error predictions using with comparisons [20]. Below is the formula of sensitivity index.

$$SI = \frac{\left(\frac{Hit}{Hit + Miss}\right) + \left(1 - \frac{\text{False Alarm}}{Falase\ Alarm + Correct\ RFejection}\right)}{2}$$

The formula contains hit rate, misses, false alarm and correct rejections. Hit rate is the cognitive bias identified by experiment participants that was also reported in the early prediction. Misses are defined as "failures to identify" potential cognitive biases, which already have in the comparison report. False alarm is to identify cognitive biases that was not reported in the comparison report. Finally, the correct rejections are cognitive biases that correctly rejected by both early report and experiment participants.

As it is name implies H-CoBIT is human cognitive bias identification technique, which is intended to analyze and discover potential human cognitive biases that may cause undesirable outcome during system use. This method is based on qualitative and is a team-based approach that will analyze and extract the potential human cognitive biases from task analysis models such as hierarchical task analysis in the first phase of development life cycle. Below diagram is conceptual framework of H-CoBIT method.

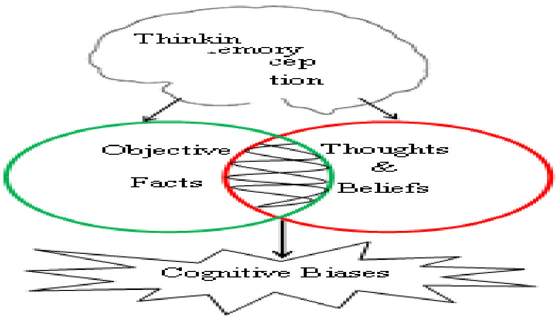



**Fig.1. Conceptual Framework of H-CoBIT**

This figure shows the conceptual framework of our method and how it identified human cognitive biases during analyzing safety critical systems. As cognitive bias is the deviation from the rational, the objective facts and the operators' thoughts and beliefs attribute some nonconformities from the norm, which then lead to cognitive biases. This diagram encompasses three main parts namely, Cognitive domain, Objective facts and thoughts and beliefs.

Cognitive domain – is a term referred to the mental processes that involved in perceiving knowledge, recognizes, remembering and thinking. It consists of the main information processing functional units, such as memory, perception and thinking. Objective facts – these facts are concepts of being something true independently from persons' subjectivity caused by perception, emotions, or imaginations. For instance, in Kegworth disaster, pilot misperceived information (objective facts) that he thought the problem is right side engine while it is not, and his thoughts and belief lead cognitive bias named 'confirmation bias'. Thought and beliefs – can be defined when things/ideas that you have in your mind accept as it is true other than the facts and the operator deviated from the objective facts on the environment and systematically created cognitive biases. This proposed method engages a systematic way of analyzing and extracting potential cognitive biases based on how human/operators interact with the system designs. The process model of H-CoBIT consists of three main activities, namely, system task Analysis, cognitive bias identification process, and cognitive bias Documentation as shown below in figure 2.

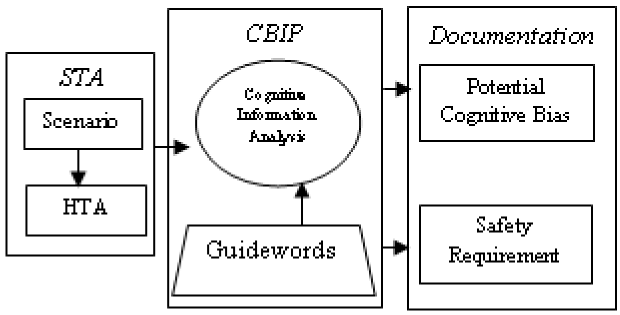



**Fig.2. H-CoBIT Process Model**

The above diagram process steps of H-CoBIT is intended to simplify the tasks and apply the proposed method in a consistent manner to be simple, consistent and reliable method. This process steps consists of three main parts, namely, System task analysis, cognitive bias identification, and documentations (both identified cognitive biases and generated safety recommendations).

A. System Task Analysis

First phase of the process model is system task analysis, which is the main part of the process model. It contains two sub tasks, which are scenario and hierarchical task analysis. In the scenario part, we modified and used [12] template scenario for structuring and understanding task being analyzed by the H-CoBIT team. This scenario contains agent, situation and environment and the action being executed by the system operators. The second task in the system task analysis phase is a hierarchical task analysis, which is intended to restructure the scenario in a graphical notation. Classic hierarchical task analysis lacks formal and suitable notations for safety critical systems. Therefore, we extended the HTA in line with corresponding to safety critical systems as we detailed in below.

*1. Scenario:* The scenario will be based on, who is responsible for performing specific task and what is his/her role. It is also included actions (task), and the environment and situation of the scenario.

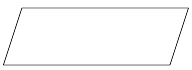
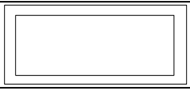
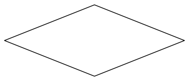
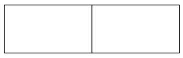
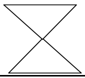
This scenario will help H-CoBIT team to easily draw hierarchical task analysis and understand every agent's roles and responsibility. The scenario will shape the overall analysis by providing information and task being analyzed. A template form proposed by [12] for describing scenarios, which has a space for recording the contextual information will be used in our method.

**Table-I: Scenario Template**

| |
|---|
| **Agents:** Human operators involved the task and their responsibility. |
| **Situation and Environment:**<br>The situation in which the scenario reflects.<br>What design technology involved the scenario and the usability issue may come across.<br>What problems can operators have on the system design. |
| **Action:** Sequence of the operational task in the scenario How the system activity interrelates. |

*Hierarchical Task Analysis:* Hierarchical task analysis is used to perform breakdown tasks in a form of goals or tasks and subtasks. The HTA will help the team to identify how the system is being analyzed will interact with the users of the system. In our newly extended HTA is not only provided the team how the system tasks are broken down, but also presents classification of the task such as iteration task, critical task, non-critical task, choices and concatenation task to be easily understood the description of the task. Below is the extended hierarchical task analysis and its descriptions.

**Table-II: Modified HTA Symbols**

| Symbol name | Notation | Annotation |
|---|---|---|
| Strategy |  | Is the plan which specific task will be performed. |
| Iteration |  | Task that iterates, which is being performed every specific time |
| Critical Task |  | Critical task that needed to be focused and analyzed |
| Choice |  | Selecting one task when faced two possibility tasks that needed to be performed |
| Concatenation |  | Tasks that needs to be performed simultaneously |
| Waiting Time |  | Period of time when the task is requested and when it performed |

B. Cognitive Bias dentification Process

The second phase of this process model is cognitive bias identification process, which is the main parts of the method. In this phase, the team will analyze the system. After being nominated H-CoBTI team build scenario and structure hierarchical task analysis, they start discussing potential human cognitive biases in the system being analyzed. The team take bottom most of each subtask from the HTA and start discussion figuring out potential cognitive biases by using H-CoBIT Guidewords as detailed below section.

*1) H-CoBIT Guidewords:* H-CoBIT guide-words work by contributing a systematic and consistent way of Brainstorming the potential cognitive biases to the analysis. The role of H-CoBIT guide-words is to motivate imaginative thinking to focus and bring together the analysis, discussions and extract the ideas.

in line with the principles of neural network on cognitive biases, we developed a set of guidewords that help identify potential cognitive biases. These guidewords originated from the four fundamental principles of neural network framework and based on cognitive functions such as perception. memory, and thinking.

As human brain seeks available knowledge and information related to the stimuli (perception) without paying too much thinking (fast and effortless), we established an ***EQUATE*** guideword to help identify potential cognitive biases related to associative principles such as control of illusion, overconfidence and optimism biases. Similarly, human brain has a characteristic of compatibility that is to be consistent with what already known, experienced or have done leading to a tendency to ignore the relevant information as it does not match with the currently mindset. Therefore, ***EQUITE*** guideword will also help generate biases related with thoughts and beliefs such as confirmation base.

We also developed ***ENCODE*** guideword that will help extract potential cognitive biases such as anchoring and hindsight biases as the brain sometimes captures irrelevant information and retains once the information is processed.

This retainment attributes that incorrect information kept on the wetware and ignores to be erased easily. So, the guideword ***ENCODE*** with parameter (the incorrect information) provide cognitive biases.

**FOCUS** and **DISRACT** are other two H-CoBIT guidewords that identify cognitive biases such as availability heurists overconfidence, attentional biases and attentional tunneling as human brain draw to conclusions based on its limited amount of information that pops up when making decisions and ignoring the other information that might be useful.Thus, each guideword represents cognitive failure that help analysts to find out easily which cognitive function is failed. Below table-III shows guideword and its generic meaning.

**Table-III: H-CoBIT Guidewords**

| #No | Guidewords | Generic Meaning |
|---|---|---|
| 1 | Equate | To associate stimuli with thoughts, beliefs and experiences. |
| 2 | Encode | To encode information structurally in wetware which cannot be erased or ignored. |
| 3 | Focus | Concentrate on specific information while ignoring the other parts (what you see is all there is WYSIATI). |
| 4 | Distract | Prevent someone from giving full attention to something/Unable to concentrate because of preoccupied mind. |

C. Documentation

In stage three of the process, the H-CoBIT team is wrapping up and record the potential cognitive biases identified during the analysis. The team records the data on a H-CoBIT table which consisted of six columns namely, task type, external error mode, cognitive failure (guideword with parameters), failure descriptions, and identified cognitive biases as shown below. As the main goal of conducting design analysis is to prevent human triggered deviations (cognitive biases) that may result in incidents, H-CoBIT method suggest safety constraint, design recommendation, and safety requirements at the final stage, when the team correctly identified the potential cognitive biases and its causes to mitigate those identified deviations (cognitive biases).

## IV. VALIDATING H-COBIT WITH ERROR ANALYSIS

In this section, we present example of safety critical system application notably aviation systems, and then we identify potential human cognitive biases using our proposed method (H-CoBIT). We used safety analysis for avionic systems as an example of validating and demonstrating how this proposed method works and analyzes potential cognitive biases. The main objective of this scenario is to evaluate the reliability and consistency of H-CoBIT method. The reason we apply this example is that Aviation domain provides us several perfect candidates as its communication complexity and designs which need more vigilant than the other system domain was chosen because of its safety critical nature and its inherent complexity. in this section, we demonstrate only scenario 1 as the rest of the scenrios go the same analysis.

A. System Description

Avionic system architecture is comprised of many individual subsystems. In this scenario, we concentrate on analyzing one of avionic subsystem activities such as operating on Go- around procedures by the crew.
Go-around (GA) Switches are designed for activating the Auto-throttle system quickly in an emergency. During the go-around the PF (pilot flying) should press the GA switches or advance the thrust levers manually. PM (pilot monitoring) verifies Auto-throttle system has activated during the go-around and monitors the thrust lever position to check that it has advanced. Failing to press the GA switch will not activate go-around thrust and the flight director will also display wrong pitch guidance confusing the pilots following their decision to go-around with serious consequences.

*Scenario 1:* When activating go around procedures pilots should either press go around switch or advance thrust lever manually within 30 seconds from the time being informed to perform GA to activate the GA procedure by following the standard. In this design of aircraft X, it should be pressed the switch and hold for 5 seconds to activate. At first push the primary flight director shows a signal of go around symbol being pushed even if it is released before 5 second, and GA is not activated, and this may cause a serious consequence by failing go around activation on the required time. Because when pilots see GA signal on the display, think it is activated and confirms their thought.

- *Agent***:** in this scenario, the pilot (flying pilot) and the co-

pilots (pilot not flying) are responsible performing GA procedure. Below table explains their responsibility of performing overall standard procedure of GA by considering responsibility relationship rules.

**Table-IV: Roles and responsibilities of the aircraft X crew**

| *AGENT* | *TASK* |
|---|---|
| Pilot Flying PF | Activate GA switch |
| Pilot Flying PF | Advance thrust levers |

- *Responsibility Relationships* – responsibility

relationships are defined the direct relation between the pilots and their target (tasks being accomplished). Therefore, in this scenario, pilot flying is the responsible holders, who performs the target (tasks), while pilot not flying is a responsible principle as below responsibility relationship diagrams shows.

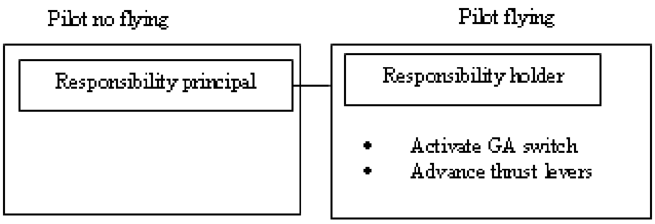



**Fig.3. Scenario 1, responsibility relationships**

Retrieval Number: F9598038620/2020©BEIESP
DOI: 10.35940/ijrte.F9598.038620
Journal Website: www.ijrte.org



Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

*Situation and Environment* – In this scenario, the precondition is that the aircraft preparing to land on a specific airport, as it is approaching; the ATC informed that there is an obstruction on the runway and need to conduct go-around procedure.

The ATC informs that there is another aircraft on the runway and avoids collision occurred in the airfield. If the landing crew failed to follow instructions from ATC, it is dangerous and may cause catastrophic accident.

- *Actions* – The flight crew operational actions are clear as

we have defined in table-IV above. It is physical activities such as actions (i.e. input commands), communications (i.e. order to do specific action loudly), checking (i.e. verifying and checking flight indicators). Both pilot flying and pilot not flying are responsible to maintain go around procedure safely. In this scenario, pilot flying is the actor who is responsible for pressing go-around button, and both pilots should verify that go around procedure activated successfully.

B. Hierarchical Task Analysis (Task decomposition)

In this section, we present HTA to break the task down into detailed sub goals, which together form the main tasks needed to be accomplished. we decomposed into sub-goals, for instance, task 1.1 push GA switch must detail sub-goals namely, 1.1.1 press G-A switch, 1.1.2 advance thrust levers. Therefore, this process goes until appropriate operations are achieved. The bottom level of any task in hierarchical task analysis should always be actions done by operators. Whilst everything above the actions specifies operational goals, which is what needs to be performed. Below figure 6. demonstrates go around procedure hierarchical task analysis (HTA).

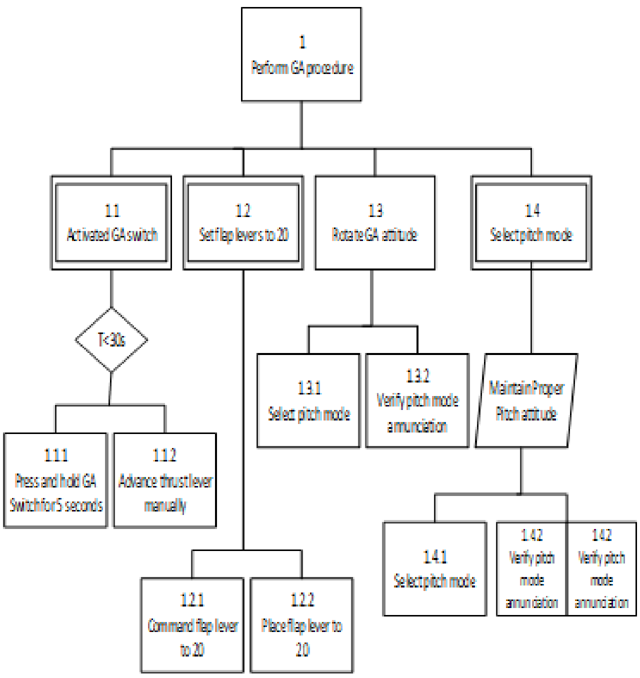



**Fig.4. Hierarchical task analysis of aircraft X go around procedure.**

C. Cognitive Bias Identification

In Scenario 1, after being constructed hierarchical task analysis, we will compute task complexity of the HTA. and if it is complicated the team will consider the complexity and take into account that the complex task may take part triggering human cognitive errors.

Therefore, above mentioned HTA is not complex and we start to analyze step by step by taking bottom most tasks (subtasks) and analyzing. We distinguish the HTA into critical and non-critical tasks. Then we take critical tasks into consideration and focus on subtasks with critical issues by analyzing them. For instance, task type 1.1 press GA switch has two subtasks, which are press GA switch and advance thrust lever manually. So, if pilots do not press and hold the button at

required time, a critical issue will happen that finally leads to accidents. Then before using guidewords, we classify human failures using human external error taxonomies [4]that will help describe observational errors done by crews, which originally stemmed from psychological error mechanisms (cognitive biases).

So, analyzing potential cognitive bias from the HTA, we chose one human error classification. For example, the team selects "*Action error*", and then it is discussed, which action error may involve in that specific task type 1.1.1. and what problem in design cockpit may cause to happen that human error. then we consider that the complexity of the instruments in aircraft design that displays system information may result in human error charges (i.e. design complexity). Then H-CoBIT team used guidewords finding the potential cognitive biases that may trigger. Applying ***FOCUS*** guidewords, it is analyzed and considered as pilot flying ***focused*** on Primary Flight Display (PFD) to see 'GA' Activated and ignores to hold *the button 5 seconds*. Therefore, this guideword generates 'Attentional bias', and 'attentional tunneling bias'.

In scenario 2 analysis we perform same process by taking bottom most hierarchical task analysis. Once they choose task type from the HTA they choose human error classifications or in other words external error modes (E.E.M.) such as "*selection error*", then team will discuss deeply paying close attention which selection error may involve subtask 1.4.1. and what kind of problem may the design GA have? That will lead to human cognitive bias.

**Table-V: Identified potential cognitive biases**

| Task type | E.E.M | Guideword | Parameter | Failure Description | Cognitive Bias |
|---|---|---|---|---|---|
| 1.1.1 | Action error | Focused | PFD to check GA activation | PF presses GA button and releases immediately when PFD shows activated by not holding 5 second | Attentional tunneling |
| 1.4.1 | Selection Error | Distracted | To read flight director messages correctly. | PF mistakenly pushes wrong button and mis interpret the visual display on the flight director. | Attentional bias |
| 1.2.2 | Action error | Equate | His thoughts with auditory alert | Pilot not flying insert wrong input and listens auditory alert that informs his input to conform his actions. | Confirmation bias |

Then, it is found that design similarity may trigger most human cognitive errors. For instance, in the pitch mode selection, there are different switches on it, such as, VNA, LNAV, VS, and FLCH speed. Therefore, the similarity of LNAV and VNAV may cause human errors during emergency situations and time pressure. then we use ***DISTRACT*** guideword, as pilot flying was unable to concentrates on reading flight director messages and pushes mistakenly a wrong button.In scenario 3, we conduct same processes as above scenarios 1 and 2, so taking the subtask of place flap lever to 20 will be selected 1.2.2. we normally find possible external error modes of human error classification. Then we find that "*action error" may* occur. then we should analyse what internanal error mechanisms may trigger this action error? we discuss and brainstorm using ***EQUITE*** guideword and consider that the pilot flygin may fall in trap cogntive biases as he compares his thoughs and belief with the actual situation.

D. Cognitive Bias Documentation and Generating Safety Requirements

This final stage, we document potential cognitive biases that is identified from the system description during H-COBIT team discussion. Above table-VII is recoded all identified cognitive biases.

E. Safety Requirement and cognitive bias Mitigation

When it comes to allocate risk reduction and mitigation, safety requirements should be generated as table below presents safety constraints, design recommendation and safety requirements. When potential cognitive biases are identified and recognized, the final step will be document safety documentations to mitigate potential cognitive biases that might eventually lead to undesirable outcome during system use.

**Table-VI: Safety Requirement**

| Safety constraints | Design Recommendations | Safety Requirements |
|---|---|---|
| 1.Pilots should pay close attention to perform a full thrust before they do any other task | DR1. Pressing Go-around switch and related steps should be simple and ergonomic (to have sequence steps) | SR1a. GA switch button should be designed to press once instead of holding 5 seconds.<br>SR1b.Auditory signals should be alerted starting from the first push until establishing full thrust go-around. |
| 2. Pilot should give their full attention to the cockpit design and eliminate any barriers may cause misinterpreting on cockpit design | DR2. Cockpit design should be designed ergonomically, especially reducing design similarity of the switches on the FD that may confuse crew during emergency or time pressure. | SR2a. it must be designed switches with different style and symbols so as to not be chosen incorrectly.<br>SR2b.When pitch mode is selected, aural signals should be designed that will inform to the crew which button is pressed to draw their attention and let them know what they have executed. |
| 3. Both pilot flying and pilot not flying should verify that flap position is put in place correctly. | DR3. Aural design on the system should be improved to draw the attention of the flight crews to prevent incorrect input. | SR4a. There should be designed a flap gate that will prevent to retract incidentally the flap position more than 20 when GA mode is activated<br>SR4b. Auditory system should be developed on PDF with long lasting beep until the pilots respond to the system. |

## V. RESULTS

As human cognitive biases are one of the major human mental deviations that trigger human errors during system use by operators, it is indispensable to reduce these cognitive distortions which lead to undesirable outcomes. To achieve this, we developed human cognitive bias identification technique along with guidewords that help identify potential cognitive biases.

Comparing the results, it is used signal detection theorem formula to validates and obtain sensitivity index score of group participants. Typically, the values of the sensitivity index start from 0 to 1. Whenever the score closes to 1 the method is considered to be highly reliable. Below table shows the result of the experiment participants.

**Table-VII: Mean score of H-CoBIT sensitivity index**

| | Team 1 | Team 2 | Team 3 | Mean Score |
|---|---|---|---|---|
| Hit Rate | 0.66 | 1 | 0.33 | 0.66 |
| Misses | 0.33 | 0 | 0.66 | 0.33 |
| False Alarm | 0.16 | 0.16 | 0.33 | 0.21 |
| Sensitivity Index | **0.66** | **0.83** | **0.33** | **0.6** |

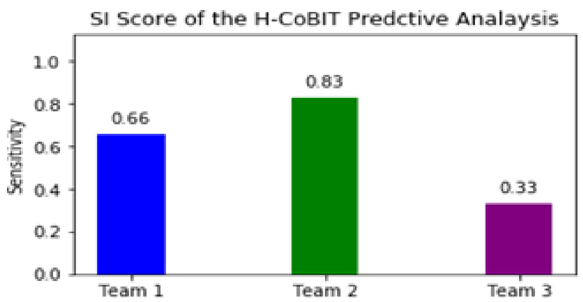



**Fig**.5. **Sensitivity index score the participants**

This diagram shows the sensitivity index of each team participants. The participants of team 3 obtained least score of the SI, as they identified only one cognitive bias from the scenario, and this shows how H-CoBIT needs a lot of practice and much time to understand the H- CoBIT. Some experiment participants have participated more than 3 times previous experiments that is why some results are getting high.

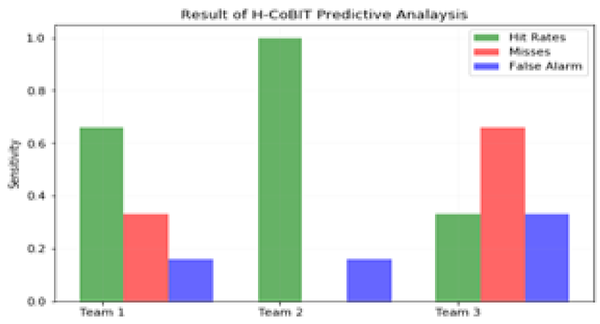



**Fig**.6. F**frequency of hits, misses and false alarms**

This diagram shows, the result of all participants scores, such as hit rates, misses, and false alarms. As we can see the diagram team 1 obtained 66% of hit rates, missed 33% and identified incorrectly 16% according to signal detection theorem. The second team identified correctly all hypothesized cognitive biases in the given scenarios, while they marked 16% as false alarm and missed nothing. On the other hand, team 3 obtained 33% of hit rate, while they missed the greatest number of hypothesized cognitive biases, so they missed 66% while they identified cognitive biases at the rate 33% incorrectly.

## VI. CONCLUSION

Identifying potential cognitive biases are paramount for understanding the root cause of human errors during system use. Huma errors are one of major factors that contribute system failure and traditional human error approaches underline 'human error' without further looking at beyond the external errors (psychological perspective). This paper proposes predictive analysis method for extracting human cognitive biases during safety critical system design analysis. The method has been developed for reducing human mental deviation named cognitive biases and generate safety requirement. To achieve this, we have reviewed and studied human information processing models and established a set of systematic guidewords that help extract potential human cognitive biases. To validate our work, we conducted experiment to evaluate the efficient, reliability, and the consistency of the proposed method. Finally, this proposed method will contribute system analysts, designers and engineers to analyze and prevent human potential cognitive biases that lead to undesirable outcome during the system use.

## REFERENCE

1. J. C. Knight, "Safety critical systems: challenges and directions," in Proceedings - International Conference on Software Engineering, 2002.
2. A. Murata, T. Nakamura, and W. Karwowski, "Influence of Cognitive Biases in Distorting Decision Making and Leading to Critical Unfavorable Incidents," Safety, 2015.
3. L. T. Kohn, J. M. Corrigan, and M. S, "To Err Is Human: Building a Safer Health System Free Executive Summary," 2000.
4. J. Reason, "Review. Human error.," Human error., 1990.
5. K. J. Vicente and J. Rasmussen, "On Applying the Skills, Rules, Knowledge Framework to Interface Design," Proceedings of the Human Factors Society Annual Meeting, 1988.
6. D. Groome, An Introduction to Applied Cognitive Psychology. 2016.
7. D. Kahneman, "Thinking, Fast and Slow, Chs. 1-8," in Thinking, Fast and Slow, 2011.
8. J. E. Korteling, A. M. Brouwer, and A. Toet, "A neural network framework for cognitive bias," Frontiers in Psychology, 2018.
9. M. Bar, "The proactive brain: using analogies and associations to generate predictions," Trends in Cognitive Sciences, vol. 11, no. 7, pp. 280–289, Jul. 2007.
10. A. M. Thomson, "Facilitation, augmentation and potentiation at central synapses," Trends in Neurosciences, vol. 23, no. 7. Elsevier Current Trends, pp. 305–312, 01-Jul-2000.
11. P. C. Wason, "Reasoning about a rule.," The Quarterly journal of experimental psychology, 1968.
12. C. D. Wickens and A. L. Alexander, "Attentional Tunneling and Task Management in Synthetic Vision Displays," The International Journal of Aviation Psychology, vol. 19, no. 2, pp. 182–199, 2009.
13. D. M. DeJoy, "The optimism bias and traffic accident risk perception," Accident Analysis and Prevention, vol. 21, no. 4, pp. 333–340, 1989.
14. D. Dunning and A. L. Story, "Depression, Realism, and the Overconfidence Effect: Are the Sadder Wiser When Predicting Future Actions and Events?," Journal of Personality and Social Psychology, vol. 61, no. 4, pp. 521–532, 1991.
15. A. Furnham, H. B.-T. journal of socio-economics, and undefined 2011, "A literature review of the anchoring effect," Elsevier.

16. P. A. Baziuk, S. S. Rivera, and J. Núñez Mc Leod, "Fuzzy Human Reliability Analysis: Applications and Contributions Review," Advances in Fuzzy Systems, 2016.
17. P. Salmon, N. Stanton, and G. Walker, "Human Factors Design Methods Review," HFIDTC/WP1, no. November 2003.
18. B. Fields, M. Harrison, and P. Wright, "THEA: Human error analysis for requirements definition," Report-University of York Department of Computer Science YCS, no. {YCS} 294, 1997.
19. S. T. Shorrock and B. Kirwan, "Development and application of a human error identification tool for air traffic control," Applied Ergonomics, 2002.
20. N. A. Stanton and S. v. Stevenage, "Learning to predict human error: Issues of acceptability, reliability and validity," Ergonomics, vol. 41, no. 11, pp. 1737–1756, 1998.

## AUTHOR PROFILE

**Salah Ali** received Bachelor of Science in Information Technology from SIMAD University Somali, and Master of Science in Information Technology from Shinawatra University, Thailand. Currently, he is doing his PhD in Information Technology from Shinawatra University. He has more than 10 years of teaching experience in the area of Software Engineering. His research interests are in the areas of requirement engineering in safety critical systems, and Human factors in the development of safety systems.


**Aekavute Sujarae** is working as instructor in the school of Science and Technology at Shinawatra University, Thailand. He obtained his Doctor of Engineering in Computer Science and Master of Engineering in Computer Science at Asian institute of Technology, Thailand. He completed his Bachelor of Engineering in Computer Engineering from Florida Institute of Technology, USA in the year of 1989. His area of interest includes learning Technology Ubiquitous IoT and Pervasive Computing Technology and Multimedia Streaming Technology.

*Retrieval Number: F9598038620/2020©BEIESP*
*DOI: 10.35940/ijrte.F9598.038620*
*Journal Website: www.ijrte.org*



*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*