Procedure Design and Validation by Cognitive Task Model Simulations

Tina Mioch  
TNO Human Factors  
Kampweg 5  
3796 DE Soesterberg  
The Netherlands  
Tina.Mioch@tno.nl

Tomasz Mistrzyk  
OFFIS Institute for Information Technology  
Escherweg 2  
26121 Oldenburg, Germany  
+49 441 97 22 560  
Tomasz.Mistrzyk@offis.de

Frank Rister  
Hapag Lloyd Airways  
Alphonsstrasse 10  
22043 Hamburg, Germany  
+49 40 760 80 797  
Frank.Rister@hamburg.de

Keywords:  
Task Models, Human Behavior, Pilot Procedures Design, Communication

ABSTRACT: Currently, the main means of communication between air traffic control and the cockpit is the voice. However, non-auditive datalink communication via the flight management system is increasingly applied for air-ground communication. In this paper, we show that the procedure to handle voice communication with air traffic control is not adequate for datalink communication, as it would lead to less feedback in the cockpit and less active monitoring. The procedure is analyzed by visualizing it through the semi-formal task model AMBOSS, which also makes it possible to simulate the procedure step by step to evaluate safety-critical tasks, e.g. tasks for which there does not exist a safety net within the procedure, such as active monitoring by the other pilot. We argue that the current procedure needs to be adjusted to the changed communication in the cockpit, and we suggest and evaluate a new procedure.

1. Introduction

Human error plays an important role in aviation accidents. The Federal Aviation Administration (FAA) estimates that human error contributes to 60-80% of all airline incidents and accidents, with communication, the governing factor for multi-crew cooperation, being its foundation (Wiegmann & Shappell, 2003).

As research and practice reveal, auditory and visual perception in the cockpit is in imbalance (Gordon et al., 2004). The perception of an auditory channel in a working environment that greatly relies on visual cues, such as the flight deck, is of considerable saliency (Wickens, 2003), whilst the long term working memory cannot store this information (Bredenkamp, 1998). Apart from lacking saliency, visual communication bears the advantage to be longer retainable and that it can be stored by technical means which make this information readily recallable at any time (Lee et al., 1999). This is one reason why the implementation of datalink air-ground communication, embedded into flight management systems is assessed since the Mid-Nineties (Parasuraman, 2001).

The translation into practice of the datalink air-ground communication in the flight management system is still at its beginning: modern aircraft enable controller-pilot-datalink communication (CPDLC), a derivative of the aircraft communication, addressing and reporting system. This technology is currently tested in a trial-phase in Eurocontrol - upper airspace and is already applied for the reception of ground clearances at larger airports as well as in the North Atlantic Track (NAT-track) scheme. (Eurocontrol, 2007).

Typically, the pilot flying (PF) has direct access to aircraft control, including the auto flight system and the flight management system (FMS). According to the standards for workload management, manifested in most procedural standards documentations of the airlines, the areas of responsibility of the pilot...
monitoring (PM) include systems control, such as hydraulics, fuel and pneumatics; and he is the one to communicate with air traffic control (Rister, 2005). As a consequence of the datalink air-ground communication being embedded in the flight management system, the responsibility of the PF and the PM would change according to the above mentioned standards. The communication with air traffic control, before a task of the PM, is done via the flight management system, which is part of aircraft control and is thus the responsibility of the PF.

1.1. Problem description

Datalink communication is on its way of becoming the standard way of communicating with air traffic control in the cockpit. This has direct consequences on the execution of procedures, as we will show by means of an analysis of a particular air-ground communication in section 2. However, the procedure that was in place for auditory communication, when applied in this new situation without substantial modifications, leads to safety critical problems. Neither the CPDLC-operators, nor the aircraft manufacturers have developed flight deck procedures yet which could solve these problems.

In the following, we argue that not adapting the procedure to the changed circumstances in communication leads to less redundancy in the handling of the situation and thus is less probable to withstand errors. We suggest a modification of the procedure, which combines the advantages of both the auditory procedure and the communication via datalink to minimize (unrecognized) errors in the cockpit and to re-establish the monitoring function as an active involvement in the task with a higher potential for shared SA (Endsley et al., 2003, Sarter & Woods, 1995). This new procedure is then validated by simulation to show that the redundancy is back in place and errors are less easily possible.

2. Analyses of Procedures

In this section, the different procedures and communication types are analysed. First, the current procedure to handle auditory communication is described. Second, the current procedure as it would be used for datalink communication if applied without modification is depicted. In addition, it is shown that the different mode of communication leads to a less safe handling of the communication by the procedure. At the end, a modified procedure is described that combines the safety of the first handling of the communication with a datalink communication.

2.1. Auditive Communication

The main means of current communication between air traffic control and pilots is voice transmission (radio). In Figure 1, a schema that depicts the communication between the different communicational partners is given. An uplinked ATC voice message is received by both pilots via headphones. The message that is radioed to an airplane is controlled and read back by the PM. Only if the PF receives the same message and only if the PF agrees with its contents and the PM’s readback, this message will lead to its execution. If the PF does not agree with the message or with the PM’s readback (which would mean that the two pilots have different mental models that inhibit shared SA), the proceduralized task distribution acts as a safety net. The PF simply only executes any clearance if he receives an ATC voice message and a PM’s readback he both agrees with.

In the following, we are looking into the procedure in more detail to evaluate for which reasons errors could occur and how these errors are foreseen and intercepted by the procedure. There are three communicational partners involved, and the procedure is described for each of the partners.

PM: The PM receives the voice uplink. Voice has a high saliency (Wickens, 2003), so that an error that comes forth from not hearing the uplink is not very likely. In addition, as the PF also receives the uplink, he can counteract this unlikely error of the PM. The PM does a readback to PF, who also received the uplink, this possible error will be intercepted by the PF. The PM then monitors the execution of the clearance by the PF. As the PM has been actively involved in the task (i.e. through the readback and decision-making whether the uplink is
acceptable or not), the likelihood of consciously and actively monitoring the actions of the PF is high.

**PF:** The PF receives the voice uplink and hears the readback of the PM. The PF might have understood the uplink differently (either through interpreting it differently or through actually hearing something different). This error is intercepted in this step. If both pilots did understand the air controller wrongly, but both in the same way, this will not directly be caught by the PF, but by the air controller, who is also listening to the readback. The PF actively has to compare his own mental model with the readback of the PM, and makes the decision whether to execute the clearance. If the clearance is acceptable, he executes it.

**ATC:** The air traffic controller initiates the voice uplink. He hears the readback of the PM, and in the case of the readback being wrong, the controller can directly intervene and repeat the uplink.

The errors that can occur in the communication, monitoring or execution tasks of other steps in the procedure are all intercepted by a safety net that is implicit to the procedure. Every possible error is foreseen (or very unlikely) and is recognized either by the person making the error or by one of the other conversational partners.

This safety net also works when either the PF does not perceive or understand the message, or if the PF misses the PM’s readback (absence of active monitoring, lower dotted arrow in Figure 1). Should the PM fail to perceive or understand the message (absence of the active, solid arrow between the PM and ATC), the PF would also refrain from executing any FMS changes, as he would lack the readback for proper comparison with the message (absence of upper dotted active monitoring arrow).

### 2.2. Non-auditive Communication

If the voice-messages are replaced by CPDLC, the received message is stored in the FMS. Using datalink has several advantages compared to voice communication. First, the pilots do not need to memorize the information provided by air traffic control. The information is set in the system, and is available at all times during task execution. If there is uncertainty about the uplink information, the pilot can just check the message again. Second, as the pilots do not need to memorize the information (and recall it when executing the procedure), the pilots experience less workload. If there is less workload, there is less probability of errors in retrieving the information (Wickens, 2003).

The FMS, in which the datalink messages are stored, is the same system with which the PF typically flies the airplane. For that reason, it is the PF who processes and executes the incoming messages, which then would have a direct effect on the airplane’s trajectory.

In the following, we are looking in more detail into the procedure to evaluate for which reasons errors could occur and whether the errors are foreseen and intercepted by the procedure.

**PM:** The PM monitors the FMS and receives the data uplink. No action is involved for the PM when receiving the uplink. He (passively) monitors the execution of the uplink by the PF. If an error occurs at this point of the procedure, e.g., omission of the monitoring task, there is no safety net for intercepting this omission.

**PF:** The PF monitors the FMS and when receiving the data uplink, he has to decide whether to execute the clearance. Execution of a clearance is done by pressing the WILCO button, which represents compliance to the ATC’s request. There are several errors that might occur. First, it is possible that the PF does not see the uplink. However, the likelihood of this error is not higher than for the current procedure, as all datalinks are additionally accompanied by an aural signal. As the PM is also monitoring the FMS, the probability of none of them seeing the uplink is small. Also, it is possible that the PF has a wrong interpretation of the uplink or that he makes an error in the decision-making process. Here, we can differentiate between the following possible consequences:

1. The PF makes a wrong decision. This only will be recognized by the PM if he is actively monitoring the execution of the uplink. If the PM is not monitoring the execution of the task (either not at all or only superficially), there is no safety net in this procedure to intercept a wrong decision of the PF. The PF does not know whether the PM is actively and reliably monitoring the PF’s task execution.

2. The PF’s wrong interpretation or decision-making of the uplink leads to the right decision. The wrong mental model is not recognized by the PM. This does not directly lead to a problem, as the action is correctly implemented by the PF, but it also does not lead to the recognition of the wrong mental model, which might lead to errors later on.

Note that it is solely the PF who has to exercise active, cognitive processing of the uplink. He is the only one involved in the clearance execution process. The readback, which should be understood as the acknowledgement of the uplink whether silent or aloud
as in the first procedure, is a task that rests solely by the PF. The PM’s role becomes passive. Even though he still has the monitoring function, his possibilities to e.g. deliver his mental model for shared SA-building to the PF is restricted. The safety net becomes leaky. Neither does an active communicative action link the PM with ATC anymore (for reception and readback), nor does the PF have an opportunity for synchronization. A modification of the procedure which could allow the PM to operate the FMS would not help, as feedback would still be missing. The situation would be mirrored and the PM would involuntarily take over duties of PF which contradicts task distribution principles as laid down in the Standard Operating Procedures (SOP).

That means that even though there are some advantages of using datalink communication (e.g. that the information is available during task execution without having to memorize it), the procedure such as it is less safe, as just one pilot needs to make an active decision. As decision-making is an error-prone activity (it costs a lot of effort and is susceptible for shortcuts), there should be a safety net in place that includes active involvement of both pilots.

3. Procedure Design

In this section, the existing procedure is modified to account for the new technological circumstances and to close safety gaps. The resulting modifications are validated by simulation, producing a new flight deck procedure. But first of all, the purpose of task modelling in this context is discussed.

Task models are an elementary part of human-machine interaction. Models show which logical steps are necessary in a task to achieve a defined goal. Existing modelling approaches (e.g. K-MADe -- Cafiau et al., 2008, VTMB -- Biere et al., 1999, CTTE -- Mori et al., 2002, Task-Architect-- Stuart & Penn, 2004) allow for task and subtask specifications as well as for their relative timeframes to be set. The task hierarchy displays a detailed description of task allocations by one or more users in a complex environment. Hierarchical task models relate formally defined structures, such as hierarchy and temporal relations, with informal elements, such as additional description of a task.

For our procedure, we decided to use the freeware modelling environment AMBOSS (AMBOSS, 2009). Due to its enhanced concepts and flexible vantage points, AMBOSS represents a useful tool for task modelling in socio-technical and safety-critical systems (Giese et al., 2008). The modelling environment has been specially expanded for the specification of tasks in safety-critical systems and now allows for inspection of relevant aspects, first of all communication (Mistrzyk & Szwillus, 2008). In AMBOSS, it is possible to model communication between non-neighbouring tasks and to implement message objects. Message objects reveal how, why, by whom and for whom an information is being generated. Similar to other modelling tools (e.g. Cafiau et al., 2008, Biere et al., 1999, Mori et al., 2002), it enables to specify the roles of actors within a hierarchy. This allows for more transparency of the task-role-communication relationship than with any other modelling approach.

3.1. Task modelling

AMBOSS allows to determine whether a communication event is classified as critical. Critical
communication events can be optically augmented. Furthermore, it can be determined whether a communication event serves as a trigger for a subtask. Additionally, it is possible to specify the necessity of feedback and to fill each event with detailed text.

Just as the approaches of K-MADe (Cafau et al., 2008), CTTE (Mori et al., 2002) or VTMB (Biere et al., 1999), AMBOSS provides its own simulator which enables an interactive validation of contexts in a task model. Flow of information, triggers, as well as the task hierarchy and its temporal relations are considered by the simulation. The AMBOSS simulator is based on the concept of ‘Enabled Task Sets’ (Mori et al., 2002). This concept provides a presentation of executable tasks. The ability of AMBOSS to simulate task models enables the analysis of pilot interaction in a socio-technical safety-critical system step by step. Thereby, experts are able to simulate various scenarios of task models and to compare them. This kind of validation helps to check the correctness of a task model and to find weak points. In situations in which several tasks are ready to get activated, the user can determine the sequencing of tasks. This enables the modeller to thoroughly examine chosen sequences of the task model for potential problems. Such shortfalls occur, as model simulations reveal, due to incorrect task-sequencing, lack of information transfer, non-observability of problematic instances but also due to unreflected workload distribution amongst the actors as well as due to tense scheduling of the task processing.

3.2. Modelling of non-auditive communication

Figure 2 shows the graphical representation of a Task Model in a tree like format which depicts a procedure for non-auditive communication. One of the challenges related to modelling socio-technical systems is to introduce communication and its parameters in a model. In the model the communication is depicted as ovals. The red ovals symbolize critical communication, whereas white oval represent regular communication.

Transferring the communication models into task models, the non-auditive model’s simulation results do not get influenced by the omission of redundant tasks and messages, such as the simulation task of the PM (subtask: PM RECEIVES CLEARANCE). No matter which irregularities cause the disturbance of the PM’s subtasks, the overall task (Handling an ATC clearance) will be executed anyway – the temporal relations as well as the trigger messages between the PM-subtasks do not necessarily guarantee the utmost necessity of the PM functions for this overall task (Figure 2). For example, if only the PF processes the uplink message, he is not be restricted by the PM at all, as there is no need to act for the PM. The reception has an alternative temporal relation, allowing just one subtasks of several alternatives to be executed. The necessity of processing as well as the readback monitoring becomes obsolete. The stage is set for a PF solo. If both pilots perceive the received message, the PF processes the message in the FMS. The PM lacks the non-auditive means to monitor or intervene in the PF’s performance. The task PM MONITORS READBACK comes with an alternative temporal relation, which is no prerequisite for completion of the entire task.

3.3. Overview of auditive communication

If one of the subtask branches of auditive communication is being destroyed, such as the reception of the uplink by the PM, the overall task, the handling of the uplink, remains incomplete. Both pilots, the PF and the PM, are dependent on reception before the PM is able to initiate a task-relevant readback. This requires that both subtasks, the reception of the uplink by both the PM and PF, have to be fulfilled before it can be proceeded; in an AMBOSS model, this would be reflected by a temporal parallel relation. Furthermore, trigger-messages that couple the subtasks of the reception of the uplink with the readback are necessary prior to initiation of the execution by the PF. Trigger messages represent the conscious processing of a received uplink. Without such cognitive processing, the subtask receiving the trigger message cannot be executed.

3.4. Description of the developed procedure

The simulation as well as the comparison of the previous two models leads to the conclusion that a new procedure shall actively re-insert the PM into the subtasks RECEPTION and READBACK. The new procedure developed by the authors focuses on dual access to the FMS by both pilots (Figure 3). We argue that this new procedure combines the advantages of both the other two procedures, and is thus safer than the datalink procedure that is currently implemented. The idea is to re-establish the monitoring function of the PM as an active involvement in the task.

In the following, we are looking in more detail into the new procedure, which is given in Figure 3, to evaluate for which reasons errors could occur and whether the errors are foreseen and intercepted by the procedure.

PM: The PM monitors the FMS. When an uplink is sent, The PM needs to act on this uplink. He needs to make a decision whether to accept the uplink, and consequently accept it. An error might occur because the PM does not see the uplink, e.g. because of focusing his attention elsewhere. This error-probability is minimized through introducing an aural signal when receiving an uplink, so that the saliency does not differ
from the other two procedures. Additionally, because the PF also receives the uplink and has to act on it, he will, after some time, point out to the PM that there is an uplink waiting for evaluation. Another error that might occur is that the PM interprets the uplink incorrectly or makes a wrong decision. In this case, again two different consequences can be identified:

1. The incorrect interpretation or decision leads to an error (either because the uplink is erroneously accepted or rejected). For this error, the PF is the safety net, as he executes the same task, and if he makes the correct decision, the difference will be found by the cross check of the system. This will lead to additional communication between the pilots.

2. The incorrect interpretation or decision does not lead to an error. The PM has a wrong mental model or makes the decision for the wrong reasons. As this does not lead to an error, it cannot be intercepted by the PM.

The PM has to actively decide whether the clearance should be executed. Here, the PM might make the wrong decision because of a wrong mental model or a bias in his decision-making process.

**PF:** The PF monitors the FMS. The procedure for the PF is the same as for the PM, and might lead to the same errors and has the same safety net. The actions are mirrored.

**System:** The task of the system is to cross check whether the two pilots have accepted (or not accepted) the uplink. This cross check intercepts possible errors that might occur in the actions of (one of) the pilots before. If both pilots make an error in the decision-making of whether to accept the uplink, and the uplink is accepted even though it should not been accepted, this is not caught with this cross check. However, the probability of both pilots making an error in the same step is small, as both pilots are actively and likely cognitively, involved in executing the uplink.

By executing an uplink in the FMS, the PF automatically delivers the task-relevant area of his mental model to the PM. As both pilots need to check, acknowledge and execute the uplink, it is assured that their mental models about this uplink do not contradict each other.

This procedure has the advantages of datalink communication and that both pilots are actively involved in the decision-making of accepting the uplink. The probability of errors decreases, as both need to come independently to a conclusion.

The new task model is safeguarded against inadvertent solos of the PF as the parallel relation of the two RECEPTION subtasks requires both pilots to receive the clearance in order to release trigger messages which are necessary for a successful completion of the sequence’s subtasks, here READBACK. Without such, the last task, EXECUTION will miss in the overall sequence. The received message will not gain access to aircraft control.

The new procedure does not impair the PF’s controllability of the airplane: the acknowledgement by the PM to execute a certain action, normally received verbally by the PF, remains silent; but as the PM needs to also press the WILCO-BUTTON and with it acknowledge and accept the uplink, the PF knows that the acknowledgement has been given.

---

**Figure 3: Non-auditive communication in a task model with active PM Feedback**

1. The incorrect interpretation or decision leads to an error (either because the uplink is erroneously accepted or rejected). For this error, the PF is the safety net, as he executes the same task, and if he makes the correct decision, the difference will be found by the cross check of the system. This will lead to additional communication between the pilots.

2. The incorrect interpretation or decision does not lead to an error. The PM has a wrong mental model or makes the decision for the wrong reasons. As this does not lead to an error, it cannot be intercepted by the PM.

The PM has to actively decide whether the clearance should be executed. Here, the PM might make the wrong decision because of a wrong mental model or a bias in his decision-making process.
Figure 4 provides a procedural visualization of the developed model. Here it becomes obvious that both pilots are required to show active monitoring as both need to accomplish an acknowledgement task.

Non-normal cases such as one pilot being either incapacitated, or simply not present on the flight deck, are covered by this procedure. For such a situation, the FMS has to be programmed to allow for dual execution out of the same seat (with a special reconfirmation bug to be programmed). This enables the PF and, regardless of his role, finally the commander to gain full and if needed sole authority over the aircraft whenever deemed necessary. The models in Figure 3 and Figure 4 remain unchanged as the PF in this special situation would simply take action in lieu of the PM which will complete the entire sequence of subtasks and finally the overall task.

4. Discussion

We have shown that the current procedure for handling datalink communication is not sufficient to guarantee safety. We suggested modifications to the procedure, and showed that these suggestions lead to a safer procedure.

Our developed procedure can be operated independently of the accessible hardware and independently of the FMS’s embedding grade. It requires no structural work, just software adjustments will become necessary and it complies with the Rules of Good Airmanship.

As described above, for several safety reasons, a dual access to trigger the WILCO BUTTON from either seat needs to be possible. This can be regarded as a shortfall, as only daily operation can reveal whether this feature will exclusively restrict to single pilot operations.

5. Acknowledgement

The work described in this paper is funded by the European Commission in the 7th Framework Programme, Transportation under the number FP7 – 211988.

6. References

AMBOSS. (2009). Homepage of AMBOSS Project, University of Paderborn, Faculty of Engineering, Computer Science and Mathematics. (viewed December 1, 2009) http://mci.cs.uni-paderborn.de/pg/amboss/


Transactions on Software Engineering vol. 28 (8), 797-813


Author Biographies
TINA MIOCH is a researcher at TNO, department Human Factors. She received a M.Sc. in Agents and Computational Intelligence from Utrecht University. Her research is on cognitive modeling and artificial intelligence. More specifically, she is interested in human error modeling, cognitive architectures and agent modeling.
TOMASZ MISTRZYK is a Researcher at OFFIS, in the R & D division transportation. He received a Diploma in Business Information Systems from University of Paderborn. His research is on task analysis and communication in safety-critical computer systems. Particularly he focuses on task modelling and human error modelling.
FRANK RISTER is a training and acceptance test pilot at Hapag – Lloyd. He flies Boeing B737 and Airbus A320 aircraft and is involved in the development of training programmes, simulation scenarios and procedure design. He has an M.Sc. in Aeronautics and Human Factors.