SPECIFICATION OF OBSERVABLE BEHAVIORS WITHIN LOE/LOFT EVENT SETS

Thomas L. Seamster, Ph.D.
Cognitive & Human Factors
Santa Fe, New Mexico

William R. Hamman, M.D., Ph.D.
United Airlines
Denver, Colorado

Eleana S. Edens, Ph.D.
Federal Aviation Administration
Washington, D.C.

INTRODUCTION

This paper presents the methods used to identify observable behaviors for Line Operational Evaluation (LOE) and Line Oriented Flight Training (LOFT) event sets along with a set of recommendations for specifying and representing observable behaviors. As the assessment of Crew Resource Management (CRM) is placed on an equal par with technical assessment, especially under Advanced Qualification Program (AQP), the specification of what is being trained and assessed takes on greater importance. The approach described in this paper is instructor-based and emphasizes the importance of the instructor/evaluator in the identification of the observable behaviors related to CRM and in the representation of those behaviors on a check or evaluation instrument. In short, the specification of observable crew behaviors within the LOE or LOFT design and assessment process are by and ultimately for the instructor/evaluators.

This paper presents the identification and specification of observable behaviors within the context of the LOE or LOFT scenario design process. That design process starts with the identification of line incidents and training issues relevant to the carrier. Instructor/evaluators are one of several sources for identifying training issues that can be translated into scenario event sets. Instructors often specify the training issues in terms of what they observe in crew performance, and that can be used to make a preliminary specification of observable behaviors. Although instructor/evaluator wording of training issues is frequently translated into more generalizable observable behaviors, the instructors are an extremely valuable and relevant resource because they are the ones that will ultimately be observing for the specified behaviors during the conduct of the LOE in the flight simulator. When properly specified and formatted, observable behaviors can serve as a valuable tool for instructor/evaluators in helping standardize the assessment of crew performance within the simulator environment and, with some modifications, within the line check environment as well.

SCENARIO DESIGN PROCESS

Because of the increased emphasis on CRM training, a structured design process is needed for LOE/LOFT that specifies and integrates required CRM and technical skills into training or evaluation scenarios. The need for a LOE design process has been made evident by AQP's mandate that CRM and technical skills be trained and assessed. The framework described here for developing LOE scenarios (ATA LOFT Design Focus Group, 1994) is based on the concept of an event set, a group of related events that are part of the scenario and are inserted into the LOE session for specific CRM and technical training objectives. Included in the framework is a method for identifying specific CRM behaviors appropriate for the event sets. Utilizing this LOE design process results in LOE sessions that are more manageable and easier to assess because they allow the instructor/evaluator to concentrate on a few CRM categories within any given event set.

The overall purpose of the LOE design process is to define a group of event sets that allows for the examination of the crew's CRM and technical skills through their ability to respond to situations of the type encountered in the course of actual line flights. The LOE/LOFT design process is made up of the five steps listed in Table 1. Steps 1, 3, 4, and 5 include activities related to the identification, validation, and representation of observable behaviors.

Table 1. Steps in the LOE/LOFT Scenario Development Process

1) Identification of primary CRM/technical training objectives

   Identify the related observable behaviors for the CRM categories.

2) Identification of possible incidents that will produce the training objectives

3) Specification and development of LOE scenario event sets

   Specify LOE scenario objectives, related TPOs, primary and secondary CRM categories, and observable crew behaviors for each scenario event set.

   Administer LOE validation instrument to ensure event sets and observable behaviors are specified and organized consistent with the CRM and technical training objectives or TPOs.

4) Evaluation and modification of the LOE scenario

   Administer the LOE validation instrument form to crews and instructors that fly the scenario.

5) Instructor training implementation and evaluation of LOE scenarios

   Develop the training materials for recurrent training instructors including the LOE Worksheet and train the instructors/evaluators.

SCENARIO EVENT SETS

The scenario event set is the principal unit of LOE scenario design and of the subsequent CRM assessment. The event set is a modification of the AQP concept of "event" and is an integral part of aircrew training and evaluation. The event set includes an event trigger, supporting conditions, and optional distracters. The event trigger is the condition which fully activates the event set and provides the instructor/evaluator with a specific time segment to focus the assessment process. Supporting conditions are other events taking place within the event set designed to further CRM and technical training objectives and to increase event set realism. The optional distracters are conditions inserted within the event set time frame that are designed to divert the crew's attention from the event trigger or other events taking place within the timeframe of the event set.

In LOE scenario design, the CRM and technical training objectives should be integrated into the event sets. This event set framework allows the design team to present the appropriate degree of realism in the LOE. Instead of focusing on a specific technical issue, the event set presents complex conditions of the line environment such as terrain, ATC, and weather issues, organized in ways to stimulate decisions, situation awareness, workload, and team issues. With the LOE scenario defined by a group of event sets, scenario validation and CRM assessment is performed at the event set level rather than at the overall LOE scenario or session level.

The significance of the scenario event set in the identification and specification of observable crew behaviors is that it can provide bounded periods of time and sets of conditions to help specify the relevant behaviors. Other units of analysis, besides the event set, are possible. NASA/UT has proposed using phases of flight (i.e., Pre Departure, Takeoff & Climb, Cruise, and Descent/Approach & Landing) in order to guide assessments of CRM behaviors. Kaempf & Klinger (1992) proposed the use of major tasks (i.e., Takeoff, CAT II ILS Approach, Missed Approach, and Enroute) to help organize the assessment instrument. In comparing these three units of analysis, the event set can provide the most control in the assessment of scripted LOEs, while major tasks or Terminal Proficiency Objectives (TPOs) may provide greater flexibility in the less structured line check.

Whether using the event set, TPO, or other element as a unit of analysis, the identification and evaluation of observable behaviors can be better focused if a clearly definable unit is specified and used to delimit the observable crew behaviors. TPOs or event sets help designers to narrow the list of all possible CRM behaviors to those most relevant under the TPO or event set conditions. Event sets have the advantage that if they are is carried through to the assessment process, there is evidence (Seamster, Edens, & Holt, in press) that their use can increase rater reliability.

CATEGORIES OF CRM BEHAVIORS

In the specification of observable behaviors, it is desirable to employ a classification system that can divide CRM behaviors into groupings that facilitate the assessment process. An analysis of how more experience instructor/evaluators group CRM concepts during assessment (Seamster & Edens, 1993) showed that experienced LOFT instructors organized CRM elements to guide the assessment process. In that categorization, CRM behaviors were divided into two groupings, one included the interpersonal CRM elements and the second grouping consisted of the CRM mental activities. The advantage of such an organization is that it helps the instructor/evaluator in the observation and assessment of the CRM performance. The interpersonal or team behaviors are more directly observable. On the other hand, the mental activities are less directly observable, and the LOE facilitator often has to make inferences based on interpersonal or technical behaviors.

Interpersonal factors (see Table 2) are relatively easy to specify and monitor because a majority of their behaviors can be observed through crew communication and the way that the crew coordinates technical tasks of flying the aircraft. The mental activities are less directly observable and often have to be inferred based on the observation of interpersonal or technical behaviors. Therefore, when asked to observe for decision making, instructor/evaluators should not be asked to observe whether the crew made a decision, rather, they should be asked to observe for specific crew communications. This organization of CRM elements helps to identify the two groups of observable behaviors shown in Table 2.

Table 2. Categorization of CRM

INTERPERSONAL/TEAM FACTORS (Directly Observable)
Communications
Group Climate
Crew Coordination
Leadership/Followership
INDIVIDUAL/MENTAL FACTORS (Indirectly Observed or Inferred)
Decision Making
Situation Awareness
Workload Management
Planning

IDENTIFICATION OF OBSERVABLE BEHAVIORS

There are a number of different approached that can be used in the identification of observable behaviors. At one air carrier (Seamster, Edens, McDougall, & Hamman, 1994), instructor remarks collected from First Look (FL) and Proficiency Checks (PC's) within two fleets were analyzed. Because these instructor remarks were recorded only when there were problems, that analysis provided a review of the types of problems that crews have in recurrent training, both technical and CRM. The analysis concentrated on remarks coded as related to CRM. Results of the analysis of instructor remarks highlighted the most frequent CRM categories that caused problems in simulator-based assessments and helped to identify the types of crew behaviors that should be observed within the different CRM categories.

For air carriers that do not collect instructor comments, there are a number of other methods that can be used to help identify training issues or line incidents. For example, one carrier had instructor/evaluators rank CRM items based on where crews showed the most problems in the recurrent training environment. The results of those rankings identified the following five CRM items for which crews had the most problems: Inquiries and Assertiveness, Briefing, Decision Process Participation, Workload/Distractions Avoided, and Bottom Lines Established. Instructor/evaluators gave specific examples for these CRM items. For example, the following problems were identified for Briefing:

- It is rare that any crew gave good briefings. Therefore, there were many unanswered questions.
- No briefing to F/A or F/O. No set guidelines. Captains afraid to be assertive.
- Briefings I have see are typically too general and often fail to specifically address how a problem should be handled. Also, I think the briefings to F/As are usually cursory and often totally inadequate.

A third source of information to help with the identification of observable behaviors is the fleet-specific task analyses conducted under AQP. As Kaempf and Klinger (1992) indicate, care must be taken when using general task analyses, since they do not account for all the performance elements or behaviors especially those related to CRM. Those researchers recommend using a supplemental form of Subject Matter Expert (SME) interview. In their study, they used the Critical Decision Method in the context of specific scenarios. Thus, a task analysis may be used as a starting point in the identification of key behaviors, but some additional method will have to be used in order to identify all the key technical and CRM behaviors within the context of specified conditions.

In most cases, the air carrier will have existing data such as instructor remarks or task analysis results that can be used to help in the identification of observable crew behaviors. It is important to note that the identification process can be made more efficient if it is guided or constrained by specific scenarios or scenario event sets that specify the exact conditions under which the behaviors are to occur.

SPECIFICATION OF OBSERVABLE BEHAVIORS

The specification of observable behaviors is an important step that can affect the outcome of CRM assessment. Several examples of current behavior specifications demonstrate a number of problems that can be avoided if a standard set of guidelines are used in behavior specification. Existing LOFT evaluation forms have listed the following behaviors:

- Team concept and environment for open communications established and/or maintained
- Captain coordinated flightdeck activities to establish proper balance between command authority and crew member participation, and acts decisively when the situation requires
- Made decision

Two of the above specifications have more than one behavior making them more difficult to rate or check than a single behavior. In general, specify just one behavior and only add a second if that second behavior is necessary to properly qualify the main behavior being observed. Concise and simple wording is recommended, and the verb should refer to a clearly observable behavior. In the case of the above, "Made decision," a mental activity that is not observable is being specified. When specifying mental activities, use a resulting interpersonal or technical behavior that is clearly observable. In this case, the observable behavior may be, "Communicated decision to crew members." Ultimately, the specification has to be done in such a way that the behaviors can be understood and used by instructors who often experience high workload in the assessment process. Therefore, instructor/evaluators should be used to review and validate the observable behaviors.

VALIDATION AND IMPLEMENTATION OF OBSERVABLE BEHAVIORS

Once the observable behaviors have been specified, they should be incorporated into an assessment form and validated by a group of SMEs or instructor/evaluators. Kaempf and Klinger (1992) worked with eight instructors and collected performance marker data, ratings, and instructor comments about the forms. In a separate study (Seamster, Edens, McDougall, & Hamman, 1994) eight members of a specific air carrier fleet participated in validating the links between observable crew behaviors and scenario event sets. These participants were either instructors/evaluators or check airmen.

Seamster et al. (1994) collected validation data using an observable behavior form developed to help air carrier personnel rate observable crew behaviors in the context of scenario event sets, crew tasks, and CRM assessment categories. The form was designed to be used by instructors/pilots who had either flown the LOE or had observed video tapes of LOE sessions and had some experience with the specific scenario event sets. The form presented sufficient background information to explain the scenario event set approach to LOE assessment and was limited to a set of observable behaviors that could be rated in about one hour. The ratings were on a five-point scale where "1" signified that there was a very low probability that the observable crew behavior was important in the assessment of tasks being performed, and the number "5" signified that there was a very high probability that it was important. Additional approaches to observable behavior validation can be developed along the lines of the reliability study presented in Seamster, Edens, and Holt (in press).

The results from Seamster et al. (1994) showed that a majority of the observable behaviors (60%) were rated as having a High to Very High probability of being important to the assessment process for the specified event set.

Once observable behaviors have been validated for a specific event set, they can be used in a LOFT or LOE worksheet to prompt instructor/evaluators to assess specific crew performance items that are both central to event set success and also have been problems in past training. An example of a LOE worksheet is shown in Figure 1 where the checking of specific observable behaviors can be used as debriefing items and as a focus to help standardize the rating process. Both instructors and check airmen have shown a strong interest in working with such instruments in the simulator and line environment.

RECOMMENDATIONS

Properly identified and specified observable behaviors play a central role in the design and assessment of LOE and LOFT scenarios. The recommendations outlined in this paper also apply to the design and implementation of line check forms. Properly represented observable behaviors are the basis for evaluation forms that can be used as tools to guide the instructor/evaluator's assessment and debriefing process. When developing crew assessment instruments, the general guideline is that better instruments are often designed by and always developed for the instructor/evaluator. Specific recommendations include:

- Be aware that existing task analyses and related flight and training manuals do not include all the key performance elements, especially CRM elements.
- Use instructor/evaluators to identify training issues, to determine possible relevant behaviors, and, to validate the observable behaviors.
- Specify observable behavior as simple statements of actions that can be clearly detected by instructor/ evaluators who may be under a relatively heavy workload
- Validate observable behaviors working with both those who have observed training/evaluation sessions and those who have flown the scenarios
- Do not ask instructor/evaluators to observe mental activities, specify the relevant interpersonal or technical activities.
- Organize observable behaviors by event sets or TPOs into worksheets that can be used as tools to produce more reliable assessments and more systematic debriefings.

ACKNOWLEDGMENTS

Part of this research was supported by the Aviation Research and Development Service of the Federal Aviation Administration. We thank them for their help and useful review comments. The views expressed in this paper do not necessarily represent the views, policies, or plans of the Federal Aviation Administration, nor those of any of its organizational elements.

REFERENCES

ATA LOFT Design Focus Group (1994). Line Operational Simulations: LOFT Scenario Design, Conduct and Validation. LOFT Design Focus Group, ATA, AQP Subcommittee, November 2, 1994.

Kaempf, G. L., & Klinger, D. W. (1992). *Integrated measurement of crew resource management and technical flying skills*. Final Technical Report prepared for the Battelle Memorial Institute, Fairborn, OH: Klein Associates, Inc.

Seamster, T. L., & Edens, E. S. (1993). Cognitive modeling of CRM assessment expertise: Identification of the primary assessors. *Proceedings of the Human Factors Society 37th Annual Meeting* (pp. 122-126). Santa Monica, CA: Human Factors Society.

Seamster, T. L., Edens, E. S., & Holt, R. W. (in press). Scenario Event Sets and the Reliability of CRM Assessment. *Proceedings of the Eighth International Symposium on Aviation Psychology*. Columbus, OH: Ohio State University.

Seamster, T. L., Edens, E. S., McDougall, W. A., & Hamman, W. R. (1994). *Observable crew behaviors in the development and assessment of Line Operational Evaluations (LOE'S)*. Washington, D.C.: The Federal Aviation Administration.

1 = Unacceptable
2 = Minimally Acceptable
3 = Standard
4 = Above Standard

CHECK (Full Obs./Part Obs./Not Obs.) **RATE Each From 1 to 4**

SEGMENT 1 EVENT SET 1 (Pre Departure through Taxi)

CRM Behaviors (Check Fully Observed/ Partially Observed/ Not Observed)	Fully Obs.	Part Obs.	Not Obs.	**Technical/CRM** (Rate 1-4)	Rate below
Crew discussed A/C performance in briefing				**TECH:** Interpretation of Airport Analysis	
Crew discussed summer operations SOP				**TECH:** Two-Engine Taxi	
Captain stated that crew ask questions				**CRM:** Crew setting tone of flight deck	
Captain specified how SOP deviations would be handled				**Overall Event Set 1.1 (Rate PIC & SIC from 1-4)**	PIC / SIC

Comments - PIC	**Comments - SIC**	**Comments - Crew**

SEGMENT 1 EVENT SET 2 (Immediately prior to Takeoff)

CRM Behaviors (Check Fully Observed/ Partially Observed/ Not Observed)	Fully Obs.	Part Obs.	Not Obs.	**Technical/CRM** (Rate 1-4)	Rate below
Crew stated task priority				**TECH:** Solution of issue of tailwind	
Crew specified which task(s) needed to be performed immediately				**CRM:** Discussion of the problem	
Crew stated the tailwind condition					
Crew stated solution to tailwind condition				**Overall Event Set 1.2 (Rate PIC & SIC from 1-4)**	PIC / SIC

Comments - PIC	**Comments - SIC**	**Comments - Crew**

SEGMENT 1 EVENT SET 3 (Takeoff to Initial Climb)

CRM Behaviors (Check Fully Observed/ Partially Observed/ Not Observed)	Fully Obs.	Part Obs.	Not Obs.	**Technical/CRM** (Rate 1-4)	Rate below
Crew stated bottom lines for abort				**TECH:** Takeoff	
Crew called out problem immediately				**TECH:** Handling system prob. at T/O	
PIC discussed decision in timely manner				**CRM:** Participation in decision	
Crew carried out all decision actions				**Overall Event Set 1.3 (Rate PIC & SIC from 1-4)**	PIC / SIC

Comments - PIC	**Comments - SIC**	**Comments - Crew**

Figure 1: Sample LOE Worksheet