Hörmann, H.-J., Soll, H., Dudfield, H. & Banbury, S.: ESSAI – Training Evaluation Study.
Paper presented at the 12th International Symposium on Aviation Psychology. Dayton/OH, 14-17 April, 2003.

ESSAI - TRAINING OF SITUATION AWARENESS AND THREAT MANAGEMENT TECHNIQUES RESULTS OF AN EVALUATION STUDY

Hans-Jürgen Hörmann[1)], Henning Soll[1)], Helen Dudfield[2)] & Simon Banbury[3)]
[1)] German Aerospace Center (DLR), Institute of Aerospace Medicine, Hamburg/Germany
[2)] QinetiQ, Farnborough, Hants/UK
[3)] Cardiff University, School of Aviation, Cardiff/UK

ESSAI – "Enhanced Safety through Situation Awareness Integration in training" – is a project sponsored by the European Commission (DG-TREN). It was carried out in a consortium of European airlines, avionic and aircraft manufacturers as well as aerospace research institutions. The goals of ESSAI were two-fold: First, the identification of factors and strategies for effective Situation Awareness (SA) and Threat Management (TM) by means of flight-crew surveys and analysis of accidents and incident scenarios. Second, the development of training solutions and evaluation of training effects on pilots' and crew performance. As a starting point for the design of a one-day training program, eight distinct core competencies for SA and TM had been described and defined in terms of training objectives. The ESSAI training program itself consists of an awareness phase with computer-based instruction, an interactive tactical decision game designed to practice new behavioral strategies, simulator training with two LOS scenarios, and a facilitated debriefing. In a pre-post design with an experimental (N=16) and a control group (N=16), training effects were analyzed using different SA/TM measures like self- and peer-ratings, behavior observations, and questionnaires. First findings of this training experiment indicate that the ESSAI training tools had a significant impact on SA measures while TM measures remained stable or show only slight increases.

Introduction

In the European Commission Framework V ESSAI (Enhanced Safety through Situation Awareness Integration in training) project, the role of crewmembers' competencies concerning Situation Awareness (SA) and Threat Management (TM) as well as the trainability of these factors was investigated. The goal of this project was to provide training tools for SA and TM and to validate their effectiveness in a high fidelity simulator environment. Part of the solution was found in enhanced training to:

- Minimize (or recover from) loss of Situation Awareness during flight operations that could result in hazardous situations;
- Provide strategies for effective Threat Management during normal and non-normal flight operations.

This was achieved by:

- Reviewing current training in relation to non-normal situations and emergencies at various operators;
- Identifying competencies and procedures relating to threat handling, error management and recovery of an appropriate level of Situation Awareness for safe flight;
- Developing and empirically validating training tools and techniques in support of the above; this will result in training solutions, guidance on their implementation in training programs and training media specification.

The design phase of the ESSAI training solution involved the following steps:

- initial review of theoretical constructs relating to SA and TM
- validation of proposed significant factors in training for SA and TM with line crew using group exercises and interviews
- development of a training needs analysis based on cognitive task analysis
- conceptual shift from the classical SA notion towards the recently proposed model of Situation Control (Amalberti, in prep.)
- definition of a competency framework meeting the needs identified
- generation of an innovative training program focusing on selected competencies.

The objective of the experimental evaluation phase was to test the hypothesis that crews' SA and TM skills were improved post-training with the ESSAI solution in comparison to conventional training through a range of objective and subjective dependent measures.

Experimental Design

Two groups of Airbus qualified operational airline pilots took part in the training experiment, which was based on a mixed model design with one between group (Experimental versus Control Group effects) and one within group factor (pre-post training effects). In the *Experimental Group (EG)*, 8 crews with 16 pilots received the ESSAI training and in the

Control Group (CG), a further 8 crews with 16 pilots received a normal LOFT-session. The ESSAI training comprises an interactive DVD, a Tactical Decision Game and tailored simulator training sessions with facilitated debriefings. All participating pilots were type-rated on Airbus A319/320 with a minimum experience of one year on type and employed by three different European companies with British (n=8), German (n=16), or Italian (n=8) nationality. Nationality is distributed equally over the Experimental and Control Group. Both groups were comparable in terms of age (mean age 37.1 years), flying experience (averages of 6306 flight hours total time and 1938 flight hours on type), as well as prior experience with Crew Resource Management and Human Factors training. As dependent variables of the experiment different methods to measure Situation Awareness, Workload, Threat Management, and Training Satisfaction in general have been applied, for example, questionnaires, self- and peer-ratings or behavior observation techniques. All measures were connected to two simulator sessions: the *Benchmark Scenario (BS)* was designed to capture the baseline of all measures prior to the training; the *Assessment Scenario (AS)* was designed to evaluate expected gains in regard to SA and TM subsequent to the training.

Results for three measures in particular will be reported in this paper: 1) a self-rating scale of attitudes towards different factors affecting SA, called FASA (Banbury et al., in prep.), 2) an index of self-rated threat management strategies – ITMS (Banbury et al., in prep.) and 3) behavior observations during the respective scenarios (Hoermann et al., 2003).

The FASA scale is divided into the following five sub-scales:

- *Attention Management* – Questions pertaining to participants' ability to attend to more than one task and resume a task successfully after being interrupted.
- *Information Management* – Questions pertaining to participants' desire to acquire appropriate information to make rational decisions.
- *Cognitive Efficiency* – Questions pertaining to participants' ability to ignore distractions and maintain Situation Awareness despite external stressors.
- *Automaticity* – Questions pertaining to participants' experience of performing routine tasks in a highly practised, automatic way.
- *Inter-Personal Dynamics* – Questions pertaining to participants' knowledge of non-verbal communication and their views on what team membership entails.

The scale was constructed in a five point Likert format: *Strongly Disagree, Disagree, Neutral, Agree,* and *Strongly Agree.* There were six questions per sub-scale making a total of 30 questions. The scoring of the scale comprised the total score from all responses given, so that the higher the score the more the participant is aware and in control of respective factors that could negatively affect SA maintenance and acquisition.

The Index of Threat Management (ITMS) questionnaire comprises 15 statements related to threat management strategies, like anticipation, prioritisation and communication. The scale was constructed in the same five point Likert format as described above for FASA. The self-ratings were related to a particular threat encountered in the simulator session. The scoring of the scale comprised the total score from all responses given, so that the higher the score the more the participant's perceived effectiveness of their own threat management.

FASA and ITMS questionnaires were presented twice, immediately after the Benchmark Scenario (pre-training) and secondly after the Assessment Scenario (post-training).

The present study adopted and modified previous approaches for the observation and assessment of crew SA and TM. Based on methods like the Line/Los Checklist (Helmreich et al., 1997), NOTECHS (Flin et al., in press), or TARGET (Fowlkes et al., 1994) a behavioural marker system was devised to rate participants' SA and TM behaviour during pre-determined scenario events. Eighteen events were rated in the Benchmark Scenario and 21 in the Assessment Scenario. These events were distributed over all phases of the flight; pre-flight briefing, departure and climb, descent briefing, approach and landing, and overall pax distractions. Both SA and TM were rated on a three-point scale, corresponding to *Notice, Understand*, and *Think Ahead* for ratings of SA, and *Ineffective, Partially Effective* and *Effective* for ratings of TM. Initially, individual ratings were made for the *Captain*, *First Officer* and as a *Crew*. However, as there were only minor deviations crew ratings were used throughout the data analyses. Hereby, some problems of missing data could be solved, but on the other hand the sample size was further reduced from N = 32 pilots to N = 16 crews in these analyses. All ratings were conducted by two observers; an Instructor in charge of the simulator, and an Aviation Psychologist in charge of the experiment. Scores were discussed at the end, and any discrepancies

between the raters were resolved.

For reasons of better comparability it was decided to combine the ratings by aggregation. Both scenarios were organized into five operational phases: 1) pre-flight briefing and taxiing out, 2) departure and initial climb, 3) descent briefing, 4) approach, landing, and taxiing in, 5) distractions from cabin due to passenger problems. Furthermore, respective items were aggregated into totalscores for SA and TM.

Results

Results regarding general training satisfaction:
Generally, the pilots were satisfied with all modules of the provided training. The majority of ESSAI trained pilots (62.5%) strongly agreed that they had learned a lot from the Low-Tech exercise and the simulator training sessions. The simulator-part was evaluated significantly higher by the Experimental Group that had been trained in tailored ESSAI simulator scenarios. Between 62% and 75% of the ESSAI-trained pilots agreed strongly that they have learnt a lot from the sim-sessions, which they found interesting and stimulating. In the non-ESSAI group the corresponding figures varied between 25% and 37% only.

Results regarding self-ratings (FASA and ITMS):
A significant shift in attitudes and self-reflected behaviors could be demonstrated with the FASA scales. As a result of the ESSAI training, pilots revealed more favorable attitudes towards dimensions identified as significant competencies in SA and TM: Information Management, Automaticity, and Interpersonal Dynamics. In order to maintain adequate levels of SA, ESSAI-trained pilots put increasing emphasis on these factors, while the attitudes of the Control Group pilots seemed unchanged. Self-rated effectiveness of TM strategies (ITMS) was also affected by the ESSAI training but less prominent than the SA measures. However, the overall trends for TM strategies are consistent with the hypothesis and statistically significant. The spread between the two groups is increasing in relation to the training events, with ESSAI-trained pilots perceiving a higher effectiveness of their TM strategies than the Control Group pilots. The results for the five FASA-subscales are summarized in table 1 and significances illustrated in figure 1. Only the abbreviations for the significant effects are shown in the table.

In a pre-post comparison CG shows a significant increase on the scale Interpersonal Dynamics, while EG shows increases on Information Management, Automaticity and Interpersonal Dynamics. Two interaction effects are significant and can be attributed to the ESSAI-training: The scores of the EG demonstrated a higher gain than those of the CG for Automaticity and Interpersonal Dynamics. In summary, self-rated behaviours and behavioural intentions as measured with FASA reveal a positive impact of the ESSAI-training tools, while the ITMS scores revealed a moderate increase in both groups.

Results regarding Behavior Observations in flight:
The most crucial test of the effectiveness of the ESSAI training was the comparison of crewmembers' actual behavior in a full-flight simulator before and after the training events. The tasks embedded in the scenarios were twofold: First, in relation to SA, the crew had to notice, understand and project flight related information appropriately in order to maintain Situation Control and "be ahead of the aircraft". Secondly, in relation to TM, overt and hidden threats like disruptive passengers, high terrain, weather, or human error had to be detected and managed effectively. Data analyses provided convincing empirical evidence that the ESSAI training solution substantially enhances SA. From before to after the training the majority of ESSAI-trained crews could progress their SA-level from Understanding to Projection, while the Control crews remained on the level of Understanding. The effectiveness of observed TM strategies does also significantly increase; however, the incremental gain of the ESSAI training compared to a "normal" LOFT-mission is visible but not statistically significant. For both conditions more effective TM strategies were observed at the end of the training. Consequently, ESSAI seems to have a larger and more specific impact on SA than on TM.

A summary of significant effects is shown in table 2 with illustrations of potential interaction effects in figure 2. During BS no significant differences occurred between CG and EG. Therefore, it can be confirmed that the performance of both groups of pilots in terms of SA and TM was comparable prior to the training. After the training the performance of the experimental crews was rated significantly higher than that of the control crews on both briefing items (pre-flight and descent briefing), which is an indicator of their better planning performance during AS. This finding was confirmed by the effects found for the SA Totalscore.

In addition, the EG showed significantly increasing performance from BS to AS on the items descent briefing, approach and landing, overall passenger distractions, SA total and TM total while the control crews' performance did not change to the same degree on these items. The interaction effect Group x

Time is clearly significant for the SA Totalscore, the descent briefing and close to being significant (Alpha = 6%) for the overall handling of distractions coming from passengers or cabin crews.

In summary, the impact of the ESSAI training solution on flight crews' behaviours was demonstrated in the experiment. SA seems especially enhanced on the level of planning and projection through improved briefings. Therefore, potential threats could be avoided in most cases. Also coping with distractions due to passenger problems seemed to be more effective after having received the ESSAI training. Similarly the interaction for the TM Totalscore can be described. Though the EG showed a significant improvement from before to after the training the interaction failed to become significant because also the CG was more effective in terms of TM after the training. Thus, a general improvement of TM techniques could be observed which seemed more attributable to the LOFT-type simulator training as such and less to the ESSAI specific features of the simulator sessions.

Conclusion

The study provides sufficient empirical evidence for the effectiveness of the ESSAI methods to train flight crews' SA and TM techniques. Positive training effects could be demonstrated across all modalities of measurement: knowledge tests, self-ratings, peer-ratings, behavior observations and general questionnaires. Scores for Situation Awareness and Threat Management are increasing as a result of the training throughout all methods of measurement. Observations have shown that especially the briefing quality as well as the approach and landing phase and the management of distractions can significantly be improved by the ESSAI-training sequence of DVD, Low-Tech exercise, simulator sessions and facilitated debrief. Full details of the analyses are described in the respective ESSAI research report (Hoermann et al., 2003).

From before to after the training the majority of EG crews could improve their SA-level from Understanding to Projection, while CG remained on the level of Understanding. This can be demonstrated especially for the briefing phases during the simulator missions. The effectiveness of TM strategies did also significantly increase; however, the incremental gain of the ESSAI training compared to a "normal" LOFT-mission is visible but not statistically significant. For both, EG and CG crews, more effective TM strategies were observed at the end of the training. Consequently, ESSAI seems to have a larger and more specific impact on SA than on TM.

Thus, the extra benefit of the ESSAI training compared to a regular LOFT-type training is more evident for Situation Awareness than for Threat Management. In summary, it can be confirmed here that with the ESSAI training tool, loss of SA is, in fact, minimized. Effective strategies for TM can be acquired through ESSAI but also through other training means (e.g. regular LOFT).

The future of the ESSAI training solution lies in the take-up by operators that should be encouraged by the experimental phase described in this document. Moreover, the power of the ESSAI training solution may only be realised through follow up studies, further development and its transfer to other transport and novel domains.

References

Amalberti, R. (in prep.) The cognitive control of dynamic situations. In P. Barach (Ed.) Clambake Seminar on Human Error. Chicago.

Banbury, S., Hoermann, H.J., Soll, H. & Dudfield, H.J. (in prep.). Development and validation of novel measures of Situation Awareness to assess the effectiveness of commercial airline pilot threat management training.

Flin, R., Martin, L, Goeters, K.M., Hoermann, H.J., Amalberti, R., Valot, C. & Nijhuis, H. (in press). Development of the NOTECHS (Non-technical skills) system for assessing pilots' CRM skills. Submitted to Human Factors and Aerospace Safety.

Fowlkes, J., Lane, N., Salas, E., Franz, T. & Oser, R. (1994) Improving the measurement of team performance: The TARGETs methodology. Military Psychology, 6, 47-61.

Helmreich, R., Butler, R., Taggert, W. & Wilhelm, J. (1997). The NASA/University of Texas/Federal Aviation Administration Line/LOS checklist: A behavioural-based checklist for CRM skills assessment (Version 4.4). Austin, Tx: NASA/University of Texas/Federal Aviation Administration Aerospace Gourp.

Hoermann, H.J., Banbury, S., Blokzijl, C., Dudfield, H., Lamers, J., Lehmann, O., Lodge, M. & Soll, H. (2003). *Experimental Validation*. ESSAI WP5 Workpackage Report. ESSAI/DLR&Q_Q/WPR/WP5/2.0. EC DG-TREN. Contract No.: 2000-GRD1-10450.

TABLE 1 Summary of significant training effects for the FASA scale values
For example "EG: AS > BS" means, the subjects in the Experimental Group score significantly higher after the Assessment Scenario (post-training) compared to the Benchmark Scenario (pre-training).

FASA-Scale	**Effects**
Attention Management	n.s.
Information Management	Main effects: EG: AS > BS; AS > BS
Cognitive Efficacy	n.s.
Automaticity	Main effects: AS: EG > CG; EG: AS > BS Interaction effect: Group x Time
Interpersonal Dynamics	Main effects: CG: AS > BS; EG: AS > BS; AS > BS Interaction effect: Group x Time

TABLE 2 Summary of significant training effects for the behaviour observations
For example "CG: AS < BS" means, the subjects in the Control Group score significantly lower after the Assessment Scenario (post-training) than after the Benchmark Scenario (pre-training).

Observation Cluster	**Effects**
Preflight Briefing	Main effects: AS: EG > CG; CG: AS < BS; EG > CG; AS > BS
Departure and Climb	Main effects: EG > CG
Descent Briefing	Main effects: AS: EG > CG; EG: AS > BS; EG > CG; AS > BS Interaction effect: Group x Time
Approach and Landing	Main effects: CG: AS > BS; EG: AS > BS; AS > BS
Overall Pax Distractions	Main effects: EG: AS > BS; AS > BS
SA - Totalscore	Main effects: AS: EG > CG; EG: AS > BS; EG > CG; AS > BS Interaction effect: Group x Time
TM - Totalscore	Main effects: AS: EG > CG; EG: AS > BS; AS > BS

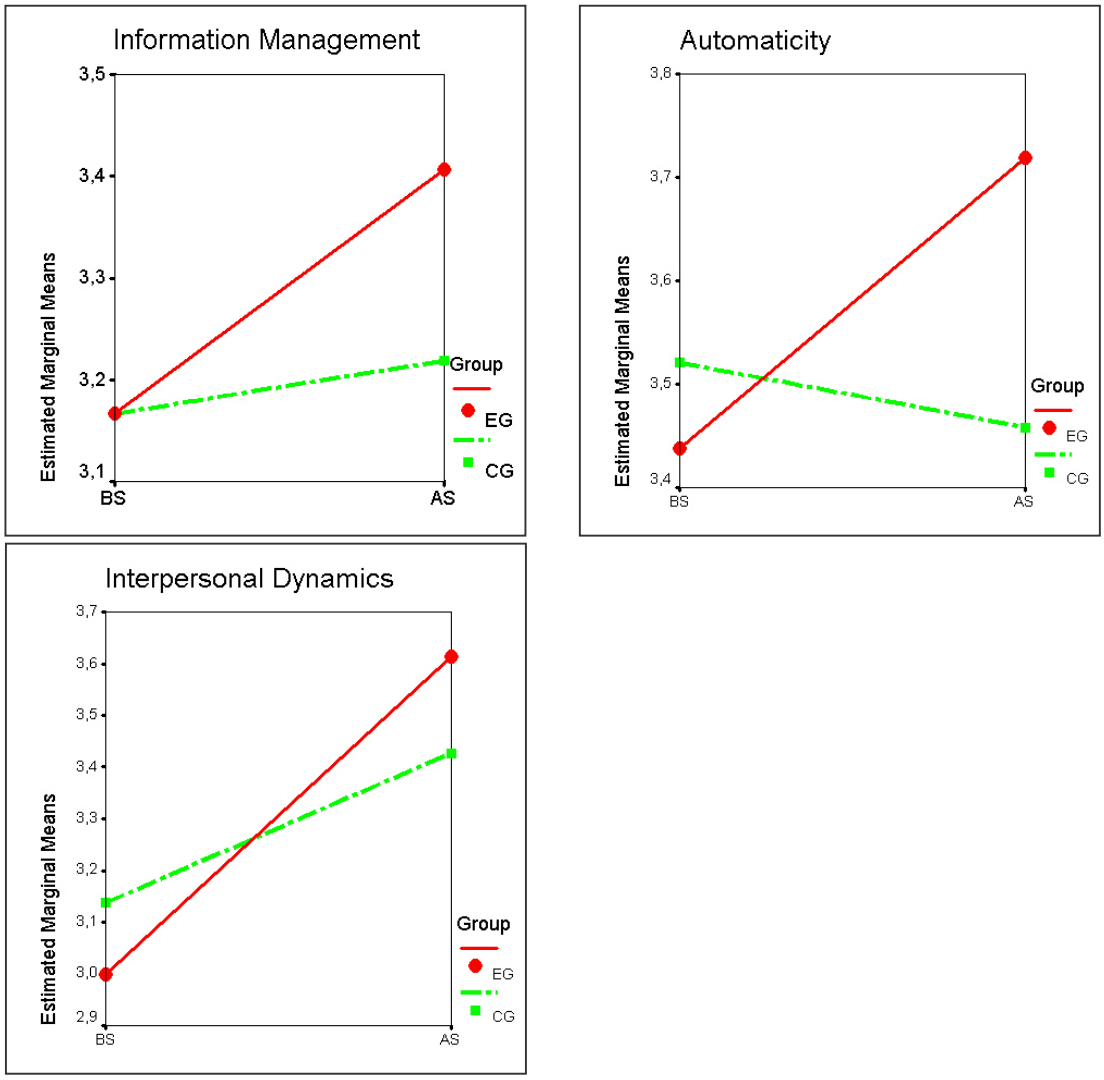

FIGURE 1 Illustration of significant training effects on scores of the FASA subscales

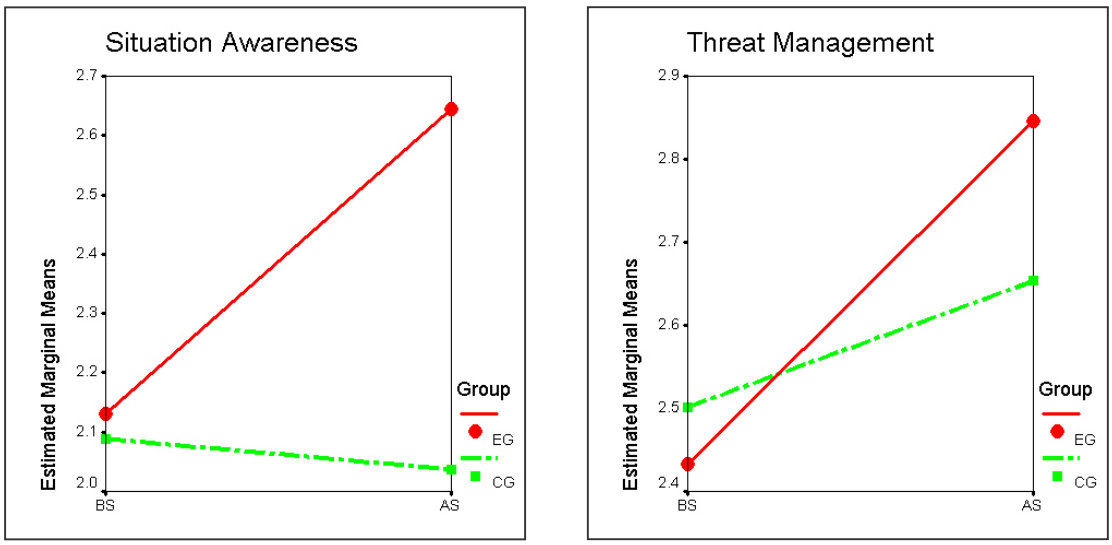

FIGURE 2 Illustration of significant training effects for the behavior observations