Proceeding of 27th Conference of the European Association for Aviation Psychology,
Potsdam, Germany, 24-28 September 2006

The Evaluation of 'Product' and 'Process' of In-flight Decision-making Training

Wen-Chin Li[1], Don Harris[1] & Chih-Chao Hsieh[2]

Department of Human Factors, Cranfield University, United Kingdom[1]
Operational Test and Evaluation center, R.O.C.[2]

Abstract

Forty-One male pilots from ROC Air Force Tactical Training Wings participated in the study. The flying experience of participants was between 354 and 220 hours with an average of 292 hours. Participants were randomly divided into two groups, 21 pilots in the experimental group, and 20 pilots in control group. Two ADM mnemonic methods, SHOR and DESIDE, that had been previously been assessed by instructor pilots as being the most applicable and having the potential to significantly improve the quality of military pilots' decision-making formed the basis of the ADM training programs. Overall, results from both the simulator-based trials (which assessed the *product* of the ADM training programme) and the pencil-and-paper tests (which assessed the *process* that the trainees applied) showed gains being made in both Situation Assessment and Risk Management skills attributable to the decision making training course. The results strongly suggest that such a short training course can be effective in terms of improving pilots' skill in situation assessment and risk management. However, these gains were at the cost of a decreased speed of responding. Nevertheless, it is suggested that a simple, short, cost-effective training program in the appropriate use of ADM mnemonic methods may ultimately produce significant gains in flight safety. Such a course may easily be integrated into current CRM or simulator-based training programs.

Introduction

For military pilots operating in a hostile environment, the normal hazards of aviation are further compounded by the enemy's intent for the destruction of their aircraft. Risk and time pressure are situational variables that further influence the decision making process, as these factors may pressure the pilot into making an immediate response irrespective of whether or not the problem is fully understood (Fischer, Orasanu & Wich, 1995). To manage threats, pilots must first assess the risks associated with them. Risk assessment feeds into decision making in two ways: during the assessment of the precipitating threats and in evaluating potential courses of action (Orasanu, Davison & Fischer, 2001).

There are several approaches to the evaluation of training that differ in terms of their aims, criteria, and methods. The traditional approach to evaluation has as its goal the identification of whether training meets its objectives, and if it doesn't it should suggest how the training programme should be modified. The role of evaluation is to correct and manage training design. Feedback from evaluation may result in revision to either or both of the training objectives and/or the training design (Patrick, 1992 & 2003). The criteria used for this type of evaluation concerns both the processes of training development and the products of training as manifested in performance at both the individual and organizational levels. The purpose of the work reported in this paper was to evaluate the effectiveness of an ADM training programme by using a series of trials in a flight simulator and by the knowledge-based evaluation of trainees using pencil and paper tests.

Method

Participants

Forty-One male cadet pilots from ROC Air Force Tactical Training Wings participated in the study. The flying experience of participants was between 220 and 354 hours with an average of 292 hours.

Scenarios for Pre-training and Post-training Evaluation

Three different types of decision-making scenario were assessed in this study.

Recognition-primed Decision Making Scenario

Pre-training: F-5E right engine fails as a result of Foreign Object Damage just as the nose gear leaves the ground at the speed 165 knots.

Post-training: During an F-5E solo, after taking off at 500 feet, the pilot hears two unusual sounds from the engines and feels the aircraft shake. Engine exhaust gas temperature is increased, and RPM decreased.

Non-diagnostic Procedural Decision Making Scenario

Pre-training: Both the leader and wingman in a formation of F-5Es are unable to land at home-base in a 'bingo' (low fuel) situation during instrument flight in bad weather.

Post-training: When an F-5E pilot is finishing Basic Fighting Manoeuvre training, the Ground Intercept Controller reports that home base weather is worsening. Surplus fuel is down to only 1,400 lb. The pilot asks for weather conditions at alternative airports.

Creative Problem Solving Decision Making Scenario

Pre-training: When flying an F-5F both left and right generators fail at the same time during a tactical manoeuvre.

Post-training: When lowering the landing gear while on the down-wind leg, the landing gear shaft warning light illuminates, indicating the nose landing gear is abnormal.

Research Design

Both the experimental group and control group participated in an initial set of trials in the flight simulator and in the knowledge-based pencil and paper tests to evaluate their baseline decision-making skills. The simulator trials and pencil-and-paper tests used the recognition-primed decision-making, non-diagnostic procedural decision-making, and creative problem solving scenarios. After these initial trials the experimental group attended a four hour 'ADM training programme for military pilots'. Two ADM mnemonic methods, SHOR (Wohl, 1981) and DESIDE (Murray, 1997), that had been previously been assessed by instructor pilots as being the most applicable and having the potential to significantly improve the quality of military pilots' decision-making formed the basis of the ADM training program (Li & Harris, 2005). Both groups then participated in a further set of trials to evaluate their subsequent decision-making performance in both the simulator and in the pencil and paper tests. To eliminate order effects, the three scenarios were presented in a randomized order during both the pre-test and post-test trials. Pilots' performance was evaluated on three dimensions (situation assessment, risk management and response time) by a professional simulator instructor. These dimensions were derived from the earlier study to select the most appropriate ADM training mnemonic methods. Each aspect of performance was rated using a nine-point Likert-type scale (with a high score of 9 and a low score

of 1, or in the case of the dimension of response time a score of 1 indicating the fastest and, 9 is the slowest).

Results

Recognition-primed decisions

With regard to performance in the flight simulator there was an effect approaching statistical significance on situation assessment performance before and after ADM training (table 1). However, there was no corresponding significant difference observed on the pencil-and-paper based tests. The group that had received ADM training also significantly outperformed the group that had not received training on situation assessment performance in the flight simulator but there was again no corresponding significant difference between the groups on the pencil-and-paper tests. The interaction term between the trained/untrained group and before training/after training trials was non-significant on the simulator tests but there was a significant interaction on the pencil-and-paper tests.

There was a significant difference on the dimension of risk management before and after ADM training on the simulator trials, however there was no significant difference on the pencil-and-paper tests. There was a significant difference in performance between the trained and untrained group on simulator trials and a result verging on significance on the pencil-and-paper tests. There was no significant interaction term between the trained/untrained group and trial on the simulator trials. However, there was a result approaching significance on the pencil-and-paper tests.

There was no significant difference on rated speed of response between the first and second trial on the flight simulator and there was also no difference between the trained and untrained group. There was, however, a result verging on significant with respect to the interaction term.

Non-diagnostic procedural decisions

There was a significant difference in pilots' situation assessment performance before and after ADM training on the simulator trials (table 1). However, there was no corresponding significant difference on the pencil-and-paper tests. There was also no significance effect between the trained and untrained groups on the simulator trials, however, there was an effect verging on a significant difference on pencil-and-paper tests. There were significant interactions terms on both the simulator trials and pencil-and-paper tests.

There was a significant difference on flight simulator performance on the dimension of risk management before and after training but there was no significant difference in risk management performance before and after training on the pencil-and-paper tests. There was a result tending towards statistical significance on the simulator trials between the trained and untrained groups, with the group that had received ADM training exhibiting better risk management performance than the other group. There was no significant difference between the trained and untrained group on the pencil-and-paper tests, though. However, there was a significant interaction between the trained/untrained group and trial on the flight simulator trials and a result verging on significance on the pencil-and-paper tests.

There was a significant difference on the dimension of rated speed of response between trials showing that pilots' response times during the second trial in the flight simulator were significantly longer than during the first trial. However, there was no

significant difference between the trained and untrained group in this respect and there was also no significant interaction term.

Table 1

The main effects of before/after training; main effects of trained/untrained groups and interaction effects on both the simulator trials and pencil-and-paper tests on the dimensions of situation assessment, risk management and response time, across all three of the decision-making scenarios. All Analyses of Variance have 1 and 39 degrees of freedom.

Types of decision-making	**Dimensions of evaluation**	**Main effect of before/after training**		**Main effect of trained/untrained**		**Interaction effects**	
		S trials	**P-P trials**	**S trials**	**P-P trials**	**S trials**	**P-P trials**
Recognition-primed decisions	SA	*F=3.520 p=0.068*	F=0.927 p=0.342	**F=6.904 p=0.012**	F=1.337 p=0.225	F=1.735 p=0.195	**F=9.555 p=0.004**
	RM	**F=12.467 p=0.001**	F=0.141 p=0.710	**F=6.736 p=0.013**	*F=2.900 p=0.097*	F=2.248 p=0.142	*F=3.266 p=0.078*
	RT	F=2.778 p=0.104	na	F=0.013 p=0.910	na	*F=3.890 p=0.056*	na
Non-diagnostic procedural decisions	SA	**F=8.216 p=0.007**	**F=8.216 p=0.007**	F=2.484 p=0.123	*F=3.593 p=0.065*	**F=4.237 p=0.046**	**F=19.540 p=0.000**
	RM	**F=6.761 p=0.013**	F=0.067 p=0.797	*F=3.316 p=0.076*	F=1.887 p=0.177	**F=7.743 p=0.008**	*F=3.266 p=0.078*
	RT	*F=3.266 p=0.078*	na	F=1.753 p=0.193	na	F=0.692 p=0.411	na
Creative problem-solving	SA	**F=5.364 p=0.026**	**F=10.320 p=0.003**	*F=3.063 p=0.088*	F=0.187 p=0.668	*F=2.993 p=0.092*	F=2.393 p=0.130
	RM	**F=6.617 p=0.014**	**F=5.885 p=0.020**	F=0.669 p=0.418	F=0.162 p=0.690	**F=4.278 p=0.045**	F=2.393 p=0.130
	RT	*F=3.185 p=0.082*	na	**F=6.164 p=0.017**	na	F=2.132 p=0.152	na

1. **S** means flight simulator trials; **P-P** means pencil and paper trials.
2. Statistically significant results ($p<0.05$) are entered in **bold** type.
3. Results approaching statistical significance ($p<0.10$) are entered in an *italic* typeface.
4. na indicates not applicable

Creative problem-solving

There was a significant difference on the dimension of situation assessment before and after decision-making training on both the simulator trials and pencil-and-paper tests (table 1). There was no significant difference between the experimental and control groups on either the simulator trial or the pencil-and-paper tests, although the former was verging on significance. There was a result approaching significance for the interaction term on the simulator trials, the group that had received ADM training showed a trend toward better performance in the second set of trials after training compared to the untrained group. However, there was no significant interaction term on the pencil-and-paper tests.

There was again a significant difference on the dimension of risk management before and after ADM training on both the simulator trials and pencil-and-paper tests. There was no significant difference in this regard between the trained and untrained group on either the simulator trials or the pencil-and-paper tests. There was a significant interaction term between the trained/untrained group and trial on

performance in the simulator. The group that had received ADM training showed greater gains in performance compared to the untrained group. However, there was no significant interaction term between the trained/untrained group and trial on pencil-and-paper trial.

There was a result verging on a significant difference on the dimension of response time between trials. There was also a significant difference between the trained and untrained group. The group that had received ADM training tended to be rated as having significantly longer response times than the group that had not received training. There was no significant interaction term.

Discussion

Baron and Hershey (1988) suggested that the study of 'outcome' bias on decision evaluation shows the tendency of people to assess the correctness of their decision-making based on the outcome of the decision. However, good decisions can lead to bad outcomes (and vice versa). Decision makers cannot infallibly be graded by their results (Brown, Kahr, & Peterson, 1974) since in a probabilistic environment a good decision cannot guarantee a good outcome. All in-flight decisions are made under uncertainty. Evaluating it as good or not must depend on the stakes and the probability, not on the outcome. It is important to evaluate the both 'products' and 'process' of ADM training efficacy.

Overall, results from both the simulator-based trials (which assessed the *product* of the ADM training programme) and the pencil-and-paper test (which assessed the *process* that the trainees applied) showed gains being made in both Situation Assessment and Risk Management skills which were attributable to the decision making training course. Improvements in situation assessment and risk management performance were, however, sometimes obtained at the expense of a decreased speed of response. In the recognition-primed decision-making scenario (which required a rapid response) after training pilots' responses were rated as being slower than before but the quality of their situation assessment and risk management decisions were improved. This may be a reasonable result as the trainees avoided a rushed, ill-considered response, even in this situation. Similar results were also obtained in the other two scenarios. Unlike the recognition-primed decision-making situation, though, in these cases an immediate decision was not required, hence sacrifices in speed of response were not so critical. However, as before, it can be suggested that 'better' decisions were being made (in terms of situation assessment and risk management) at the expense of longer response times. The results from the pencil-and-paper tests would support this observation. It is possible that further training and practice may subsequently increase the speed of the decision-making process.

Training evaluation involves a series of levels. Simulator trials offer the option of being able to provide pilots with many practice problems, designed to build up their recognition of patterns. They can consistently be adapted to the pilot in terms of difficulty level and in terms of evaluating pilots' skills/knowledge and behaviour. Pilots can perform tactical tasks in the context of a simulated real world environment and missions and pilots can be evaluated in terms of their decision-making performance by an instructor. Pencil and paper trials have limitations for evaluating pilots' behaviour. They can analyze the cognitive processes of pilot, for example what aspects of situation assessment and risk management were conducted by a pilots' mental simulation. However, this research approach can not assess the real-time performance of a pilots' in-flight decision-making performance. It also has a lack of psychological time-pressure. The key cues for pilots' decision-making are time and risk (Orasanu & Fisher, 1997). Based only upon pencil and paper trials it is difficult to

conclude if it was the cognitive processes of pilots for conducting situation assessment and risk management that caused longer response time, as paper-pencil trials were not suitable for assessing this parameter. On the other hand, simulator trials can assess the products of ADM performance of pilots', including situation assessment, risk management, and response time in real-time setting, however, they have limitations for analyzing the cognitive processes of pilots. Using both evaluative approaches, however, provides a far more comprehensive evaluation of the efficacy of the decision-making training program.

Conclusion

The simple, short, cost-effective programme devised to train cadet pilots in the appropriate use of ADM mnemonic methods produced significant gains in decision-making performance. Such a course may easily be integrated into current CRM or simulator-based training programs. However, it still needs to be evaluated if these performance gains continue to be evident at a later date during actual operations. This will be established in the near future, however, the initial results would seem to be most promising.

References

1. Baron, J. and Hershey, J.C. (1988), 'Outcome bias in decision evaluation', *Journal of Personality and Society Psychology,* Vol. 54, pp. 569-579.
2. Brown, R.V., Kahr, A.S., & Peterson, C (1974). *Decision analysis for the manager*. New York: Holt, Rinehart & Winston.
3. Fischer, U., Orasanu, J., & Wich, M. (1995). Expert Pilots' Perceptions of Problem Situations. *Eighth International Symposium on Aviation Psychology* (pp. 777-782). The Ohio State University.
4. Klein, G. A. , & Woods, D. D. (1993). Conclusions: Decision Making in Action. G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Editor), *Decision Making in Action: Models and Methods* (pp. 404-411). Norwood, New Jersey: Ablex.
5. Li, W.C. and Harris, D. (2005), 'HFACS Analysis of ROC Air Force Aviation Accidents: reliability analysis and cross-cultural comparison', *International Journal of Applied Aviation Studies,* Vol. 5, No. 1, pp. 65-81.
6. Murray, S. R. (1997). Deliberate Decision Making by Aircraft Pilots: A simple Reminder to Avoid Decision Making Under Panic. *The International Journal of Aviation Psychology*, 7(1), pp. 83-100.
7. O'Hare, D. (2003). Aeronautical Decision Making: Metaphors, Models, and Methods. P. S. Tsang, & M. A. Vidulich (Editor), *Principles and Practice of Aviation Psychology: Human Factors in Transportation* (pp. 201-237). New Jersey: Lawrence Erlbaum.
8. Orasanu, J., Davison, J., & Fischer, U. (2001). The Role of Risk in Aviation Decision Making: How Pilots Perceive and Manage Flight Risks. *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 58-62). Santa Monica, USA: Human Factors and Ergonomics Society.
9. Orasanu, J. and Fischer, U. (1997), 'Finding Decisions in Natural Environments: The View From the Cockpit', in Zsambok, C.E. and Klein, G. (Editor), *Naturalistic Decision Making*, Lawrence Erlbaum, Mahwah, pp. 343-358.
10. Patrick, J. (1992). Training*: Research and Practice* . London: Academic Press.
11. Patrick, J. (2003). Training. P. S. Tsang, & M. A. Vidulich (Editor), *Principles and Practice of Aviation Psychology: Human Factors in Transportation* (pp. 397-434). New Jersey: Lawrence Erlbaum.
12. Wohl, J. G. (1981). Force management decision requirements for air force tactical command and control. IEEE Transactions on Systems, Mans, and Cybernetics, SMC-11, pp. 618-639.

The evaluation of product and process for in-flight decision-making training

Li, Wen-Chin

2006-09-28T00:00:00Z

Wen-Chin Li, Don Harris and Chih-Chao Hsieh, The evaluation of product and process for in-flight decision-making training, Proceeding of 27th Conference of the European Association for Aviation Psychology, Potsdam, Germany, 24-28 September 2006.

http://dspace.lib.cranfield.ac.uk/handle/1826/8161