DETERMINING AIRCREW COORDINATION TRAINING EFFECTIVENESS

Janis A. Cannon-Bowers
Carolyn Prince
Eduardo Salas
Naval Training Systems Center
Orlando, Florida

Commander Jerry M. Owens, USN
Naval Air Systems Command
Washington, DC

Ben B. Morgan, Jr.
University of Central Florida

Gregory H. Gonos
Naval Training Systems Center
Orlando, Florida

ABSTRACT

Recognition of the need for aircrew coordination training has led to the development of a number of military and commercial aircrew coordination programs in recent years. The effectiveness of such programs is unclear, however, due to a general lack of training effectiveness data. The purpose of this paper is to present a set of guidelines and recommendations for evaluating aircrew coordination effectiveness in the military. The evaluation plan presented here is drawn from an effort by the Naval Training Systems Center to develop aircrew coordination training for Navy and Marine pilots. Drawing from past work in program evaluation, training effectiveness, and aircrew coordination, this evaluation plan has a number of distinguishing features, including: a) multiple levels of evaluation criteria, b) pre-training assessment, c) recognition of evaluation needs throughout the design cycle, and d) outcome data that can be used for multiple purposes. Such a comprehensive evaluation approach is necessary to ensure that mission safety and effectiveness are increased, and training resources optimized.

OVERVIEW

Statistics indicating that a significant number of aircraft incidents and accidents can be attributed to poor aircrew coordination have led to the development of several military and commercial aircrew coordination training programs in recent years. While the need for such programs has been well established, data regarding their effectiveness are largely unavailable. Furthermore, where evaluation data do exist, they are often incomplete as a means to provide full assessment of program effectiveness. A comprehensive approach to evaluation of aircrew coordination programs is necessary to provide data that will: (a) assess the degree to which mission safety and effectiveness are increased, (b) provide for program improvements, and (c) ensure that the allocation of training resources is optimized.

The purpose of this paper is to address evaluation issues pertinent to aircrew coordination training and to present guidelines and recommendations for a comprehensive approach to evaluating aircrew coordination training in the military. Examples used here to illustrate various concepts will be drawn from a program under development at the Naval Training Systems Center (NAVTRASYSCEN) to train aircrew coordination skills for Navy and Marine platforms. The evaluation plan being developed as part of this program integrates and applies state-of-the-art knowledge in training evaluation and aircrew coordination. Its distinguishing features are 1) attention to evaluation issues throughout the training design cycle, and 2) use of a multilevel, multi-component evaluation approach. Before delineating the details of this approach, a discussion of training evaluation issues and concepts will be presented.

ISSUES IN TRAINING EVALUATION

Purposes of Evaluation

In general, training effectiveness evaluations are conducted to determine whether a training program teaches task-relevant knowledge and skills adequately and efficiently (1). On a more specific level, evaluation results serve a variety purposes in support of training system design and maintenance (2). First, program evaluation results can indicate whether the goals and objectives of a program are appropriate to achieve desired results. In the case of aircrew coordination training, for example, this would entail a demonstration that mission safety (a desired result) is increased due to aircrew coordination training. Second, evaluation data can indicate whether the content and methods employed in the training are appropriate for meeting the program's overall objectives. It is conceivable, for example, that aircrew coordination training may successfully

reduce mishaps, but that a particular program has not developed training content or methods that address aircrew coordination needs adequately.

A third use of evaluation data is to determine how to maximize transfer of training (3). Factors related to the trainees or training situation may be found to have an impact on the extent to which skills are transferred from the training situation to the operational environment. Once identified, these factors can be addressed as a means to maximize transfer. Fourth, training evaluation can be a source of information that suggests improvements to the training program. Feedback regarding specific program deficiencies can then provide for changes in program content or methods (4). On an individual level, evaluation data can provide a critical source of feedback to trainees and instructors. Well developed evaluation instruments will allow an assessment of areas of deficiency in trainees, and an indication of instructor success. Finally, knowledge gained in training evaluation can indicate the efficiency of training. The optimal use of training resources requires that programs reach their goals with the lowest expenditure of time and resources possible (1).

Training Program Evaluability

Designing and conducting a training program evaluation that can serve the various purposes listed above is desirable since it maximizes the likelihood that positive changes to the training program will be initiated. Successfully undertaking such a comprehensive evaluation is complex, however, since different kinds of information are needed for different uses. Further, researchers in the area of program evaluation contend that a program's original design and implementation affect how difficult it will be to determine program effectiveness at a later time. Rutman (1) defines the "evaluability" of a program as the extent to which its effectiveness can be evaluated.

According to Rutman (1), the evaluability of a program is a function of a) the quality of program definition and implementation b) the clarity and specificity of program goals and desired effects, and c) the extent to which the cause and effect relationships between program activities and goals are plausible. In terms of training evaluation, this means that the evaluability of a training program is increased when it is well defined in terms of tasks to be trained, and well implemented with respect to design intentions; when training objectives are clearly stated in terms of specific knowledge, skills, and abilities identified as necessary for mission accomplishment; and when program content and methodology are developed in strict accordance with program objectives, and based on thorough analysis and testing.

To some extent, a well planned and conducted front-end analysis, by definition, will contribute to a training program's evaluability. Beyond this, however, explicit attention to evaluation needs throughout the training design cycle can help to ensure that evaluation of the program's effectiveness will be possible. While many training system designers voice the necessity to plan for evaluation needs, the fact remains that many training programs are designed without attention to evaluation concerns. However, in the present NAVTRASYSCEN effort, steps are being taken throughout the design cycle to enhance the program's evaluability. For example, specific aircrew coordination skills and behaviors are being defined as a basis to derive measurable training objectives (see (5)). This will allow a thorough "summative" evaluation of the program's effectiveness (2). Summative evaluation refers to assessment of the overall outcomes of a program (in this case, how well training objectives are achieved).

Early attention to evaluation needs also will allow for a "formative" evaluation of the program. Formative evaluations are those that are conducted as programs are being developed so that iterative refinement of training methods and content can be initiated before the program is finalized and implemented (2). Comprehensive program evaluation requires that both summative and formative assessments are made. Finally, evaluation needs are being anticipated in the current effort so that, for example, data collection capabilities can be specified in the design of training devices.

Evaluation Hierarchy

Another method to optimize information gained from a training evaluation and to ensure comprehensiveness of results is to adopt a multi-component approach to evaluation. Such an approach is supported by researchers in the general training literature (e.g., 6, 4, 7, 3) as a means to assess training effectiveness. With respect to aircrew coordination training, practitioners and researchers in this area have also advocated such an approach (e.g., 8, 9, 10, 11, 12, 13).

The logic behind adopting a multi-component approach to evaluation is that: a) diverse types of information are needed to fulfill the evaluation purposes outlined above, b) the interpretability of, and confidence in evaluation results are enhanced when multiple sources of data are available, and c) multiple criteria are necessary in absence of a well defined, measurable "ultimate" criterion. In addition,

Kirkpatrick (3) maintains that training success is dependent on different levels of change in trainees, so that comprehensive assessment of training impact must consider multiple levels of change. In fact, Kirkpatrick conceptualizes components of evaluation as an evaluation hierarchy. The levels of evaluation in this hierarchy include: trainee reactions to training, extent of learning by trainees, extent of performance change in trainees, and degree of impact on organizational effectiveness due to training.

Kirkpatrick (3) contends that transfer of training is a function of change at all levels of the hierarchy. Thus, positive trainee reaction to training will presuppose cognitive assimilation of program content (i.e., learning). Learning of targeted concepts, in turn, is necessary for changes in trainee performance to take place. Effective performance change at the individual or crew level then determines the impact of training on overall organizational effectiveness (e.g., reduced accidents, mishaps, etc.). In order to conduct a comprehensive evaluation, therefore, data from all four levels of the hierarchy are necessary. This is especially true in military environments, where it is difficult to define operationally the "ultimate" criteria of combat readiness or effectiveness. Employing multiple criteria in this case can allow evidence of training effectiveness to be harnessed in absence of a single, ultimate criterion.

Information from different levels of the hierarchy also serves different purposes. For example, information about the trainees' level of learning may suggest portions of the program that require refinement. On the other hand, data indicating that the training failed to affect desired organizational changes may suggest that interventions are needed to increase transfer of learned skills and behavior to the operational environment. Overall, Kirkpatrick's hierarchy of evaluation provides a framework in which evaluation activities can be conceptualized and designed. Table 1 presents a summary of the levels of evaluation in the hierarchy, along with sample measure(s) and an indication of past researchers who have successfully developed criteria at each level. Table 1 also shows that in addition to the levels of evaluation included in Kirkpatricks's hierarchy, another evaluation component, pre-training assessment, is being advocated here.

The purpose of pre-training assessment of trainees is twofold. First, a measure of targeted skills prior to training provides a baseline with which post-training measures can be compared (7, 4). Only in this way can it be determined whether trainees are learning intended concepts and skills, and/or improving pre-existing skills. A second purpose of pre-training assessment is as an indication of trainee "readiness" for training. It has been shown that factors such as training-related attitudes, self-confidence, and expectations, and job and career involvement, have an effect on transfer of training (14, 4). These "trainability" factors combine to affect training motivation and outcomes. Hicks and Klimoski (15) found, for example, that trainees who received a realistic preview of training content and could choose whether or not to participate in training were more likely to benefit from it. Further, it is important for trainees to believe that successfully mastering training content will help them to achieve desired goals (16, 14). Trainees with low job involvement, lack of career goals and plans, and lack of understanding of training relevance have been shown to score significantly lower on post-training measures (14).

With respect to training evaluation, pre-training assessment of trainability factors is critical because it can indicate whether unsuccessful training outcomes are attributable to training program deficiencies or to motivational problems. Further, low trainability scores in trainees suggest that interventions may be in order to precede or augment training. For example, remedial training can be instituted for trainees who have negative attitudes toward training.

AIRCREW COORDINATION TRAINING -- EVALUATION COMPONENTS

Aircrew coordination training seeks to modify the behavior of aircrews by teaching them to coordinate the use of resources effectively in the cockpit. To do this, it imparts critical requisite knowledge regarding the importance of aircrew coordination; it seeks to teach specific behaviors (i.e., is performance-oriented); and it answers a particular need at the mission effectiveness level (i.e., is mission-relevant). Given these characteristics, along with the difficulties associated with deriving ultimate criteria a multi-component approach to evaluation of aircrew coordination training is advocated here.

The various components of evaluation being proposed are summarized in Table 2. Table 2 shows each evaluation component as it applies to aircrew coordination training, as well as the kind of evaluation information it will yield. It should be noted that the highest level of the Kirkpatrick's original hierarchy, organizational impact, must be tailored for application to an aircrew coordination context. Organizational impact in this context can be best thought of as mission

Table 1. Kirkpatrick's Evaluation Hierarchy

Level of Evaluation	Sample Item/Measure	Sample Sources
Pre-Training Assessment*	1) Aircrew coordination is critical to mission safety (agree/disagree).	Noe & Schmitt, 1986
	2) List four barriers to effective communication in the cockpit.	Huczynski & Lewis, 1980
		Helmreich & Wilhelm, 1986
Reaction to Training	1) This training was relevant to my performance as a pilot (agree/disagree).	Noe & Schmitt, 1986
		Helmreich & Wilhelm, 1987
Learning	1) List four barriers to effective communication in the cockpit.	Jensen & Adrion, 1988
		Hicks & Klimoski, 1987
Performance Assessment	1) Trainee helped a fellow crew member who was overloaded (during a simulated exercise).	Helmreich & Wilhelm, 1986; 1987
		Helmreich et al., 1989
Organizational Impact	1) To what extent are aircraft incidents reduced after introduction of aircrew coordination training.	Cullen et al., 1976

* Not part of Kirkpatrick's original hierarchy

effectiveness, since the likelihood that a single training program will affect changes at the global organizational level (i.e., across the Navy and Marine Corps) is not feasible. Instead, the term mission effectiveness is used to refer to changes in the extent of overall mission effectiveness within weapon-system communities.

Further, inspection of Table 2 reveals that the evaluation ,7 components are both complementary and interrelated. In isolation, each component provides specific data in support of a subset of evaluation questions. Together they can provide a comprehensive assessment of aircrew coordination training effectiveness, yielding information necessary to judge program effectiveness at the individual and crew levels, to assess the impact of training on mission accomplishment, and to provide feedback that will enhance trainee, instructor, and program effectiveness. The following describes in more detail the recommended components of evaluation being developed for aircrew coordination training.

Pre-Training Assessment

As discussed, use of a pre-training measure of knowledge and skill can provide a baseline for comparison with post training outcomes. Pre-training assessment measures for this purpose should be developed based on the program's training objectives and resemble closely the format of post-training measures. In terms of aircrew coordination, pre-training assessment should determine, for example, the extent to which crews understand the importance of effective aircrew coordination, the skills that comprise effective aircrew coordination, when such skills are most critical and how they are performed in terms of specific behaviors.

Pre-training assessment of trainability factors (described above) for aircrew coordination training should include assessment of crews' attitudes toward aircrew coordination training, their expectations for training and their beliefs about how aircrew coordination training can improve their performance in the cockpit and enhance mission accomplishment. Helmreich &

Table 2. Proposed Components of Evaluation and Resulting Data for Aircrew Training Effectiveness Evaluation

Purpose/Use of Data	Evaluation Component 1	2	3	4
Assess whether program objectives are met	X	X	X	X
Assess validity of program content	X	X	X	X
Assess training methods		X	X	X
Assess trainee readiness for training	X			
Assess impact on individual trainees	X	X	X	X
Assess impact on overall mission effectiveness				
Provide feedback to make program improvements	X	X	X	X

Component 1 = Pre-Training Assessment
Component 2 = Trainee Reactions to Training
Component 3 = Learning Assessment
Component 4 = Performance Assessment
Component 5 = Mission Effectiveness Assessment

Wilhelm (17, 11) have developed the Cockpit Management Attitudes Questionnaire (CMAQ) as a means to assess crew attitudes about aircrew coordination. This instrument will serve as a basis to develop measures of trainability for aircrew coordination training in the NAVTRASYSCEN effort.

Reaction

The first level of Kirkpatricks's hierarchy of evaluation, trainee reactions to the training, involves a post-training assessment of the degree to which trainees believe the training was relevant, worthwhile, interesting, well conducted, and the like. This type of information is useful for a number of reasons, including: 1) positive reactions to training help to ensure top-level support for the training program; 2) favorable reactions can enhance trainee motivation to learn and succeed in training; 3) poor trainee reactions may pinpoint areas of training content and/or methods that need to be upgraded.

With respect to aircrew coordination training, Orlady and Foushee (8) recommend use of trainee reaction measures as a means to ascertain trainees' perceived relevance of training concepts and methods. According to Helmreich and Wilhelm (11), enthusiastic endorsement of a program does not ensure desired behavioral change, but can help indicate when a program is without value, or has failed to reach its goals. Further, selfreported evaluations of the training program can indicate a program's level of credibility, and provide important guidance in determining which aspects of the program need to be modified or strengthened. For example, trainee reactions to aircrew coordination training that employed lectures, video tapes and simulated exercises may indicate that trainees found little value in the video tapes, but believed the lectures and simulated exercises to be worthwhile. Examination of the video tapes by training designers would then be in order.

Trainee reaction data can also provide insights as to why desired changes in the other levels of evaluation (i.e., learning, performance, and mission effectiveness) did not occur. Specifically, a program that failed to teach trainees targeted aircrew coordination concepts may have done so because training content was insufficient, or because trainees did not believe the program to be relevant, and thus were not motivated to learn the material. Finding negative trainee suggest that the methods of presenting material be upgraded, or program credibility enhanced, rather than changing program content.

Instruments designed to assess trainee reactions are relatively simple to design and administer. Such instruments are most often paper-and-pencil measures that can be completed by trainees at the conclusion of training. According to Goldstein, (4) instruments developed to measure trainee reactions to training should be tailored to elicit responses that are quantifiable, should provide for

anonymity of trainees (so that responses will be candid), and should be pre-tested to ensure completeness and psychometric soundness. The NAVTRASYSCEN aircrew coordination training program will employ a simple paper and pencil measure of trainee reactions. Trainees will be asked to report their reactions on Likert-type scales with respect to such things as how worthwhile, valuable, relevant, interesting and challenging they believed the training to be.

Learning

The second level of evaluation in Kirkpatrick's hierarchy involves measuring the extent to which trainees learn the concepts, principles and information necessary for successful performance. At the individual/crew level, such information indicates how successfully the training program imparted targeted knowledge. At the program level, data regarding the extent of learning exhibited by trainees can provide feedback that determines whether the program is reaching its objectives, and pinpoint areas of the program in need of enhancement. In addition, as with reaction data, an evaluation of learning may be necessary to help interpret data from other levels of evaluation. Specifically, failure of training to affect changes in performance or mission effectiveness may be due to the fact that trainees did not learn the critical knowledge and concepts that underlie successful performance.

Aircrew coordination programs often employ measures of learning to evaluate trainees. A recent review by Jensen (13) revealed that a variety of techniques have been used to assess learning in past aircrew coordination training, including paper and pencil questionnaires, interactive lectures, and small group discussions. According to Wexley & Latham (7), instruments designed to measure learning must be based directly on the program's objectives, i.e., they must determine whether the training successfully taught the concepts, principles and knowledge specified as requisite for effective performance. Other concerns, such as psychometric soundness, difficulty, and length, are also important factors in designing a learning test for aircrew coordination training, regardless of the scale's format. The present approach will incorporate measures of learning that display a one-to-one correspondence with the knowledge objectives identified in the training needs analysis, and is based on a substantial body of research in the team training and performance area (e.g., 18).

Related to learning assessment, the majority of present aircrew coordination programs also seek to assess the extent to which training leads to positive attitude change toward aircrew coordination. According to Helmreich and associates, a pilot's attitude toward effective aircrew coordination (i.e., the extent to which he/she believes that aircrew coordination is important for mission accomplishment, and that behaviors involved in aircrew coordination are necessary, etc.) has a direct impact on how well he/she performs in the cockpit. These researchers maintain that aircrew coordination training designed to produce positive attitudes will lead to effective aircrew performance.

This use of attitude data differs from the discussion of pretraining assessment of attitudes above. As a pre-training trainability factor, it was maintained that crew attitudes would have an impact on the crew's motivation to learn targeted knowledge and skills in training. Helmreich et al. (17) maintain that attitudes are also a viable (post-training) index of training success. In fact, according to Helmreich et al., the goal of aircrew coordination training should be to produce more positive attitudes toward aircrew coordination in trainees. In support of this position, Helmreich and Wilhelm (11) found that aircrew coordination training produced *significant attitude change in trainees.* Further, Helmreich et al. (19) found that trainee attitudes were significantly related to aircrew coordination performance as rated by specially trained evaluators.

The discussion of attitude change as a dependent measure in training effectiveness has been introduced at this point for several reasons. First, attitude change is a cognitive event; as such, it can be conceptualized at the same level of evaluation as learning. Second, attitude change and learning are processes that mediate performance change, and they must be assessed to provide necessary information regarding training effectiveness. However, These measures cannot substitute for assessment of overt aircrew performance change. This is in line with the overriding goal of aircrew coordination training to identify and teach specific aircrew coordination skills and behaviors. Evaluation of post-training aircrew coordination performance, therefore, is paramount to the program under development at NAVTRASYSCEN.

Performance

Evaluating post-training performance (the third level of Kirkpatrick's hierarchy) provides information regarding 1) the extent to which trainees learned how to perform necessary skills and behaviors, and when these are appropriate, 2) the extent to which training transferred to the operational environment, 3) the relative success of various parts of the training program in achieving desired results. Such information is pertinent at both

the individual and program levels, as an indication of individual trainee readiness, and of overall program effectiveness, respectively.

While measures of performance change are common as indices of training effectiveness, they are often difficult to design and collect. To begin with, objective measures of performance (e.g., instrument readings, time to complete a task, etc.) are not typically sensitive enough to detect changes in behavior (20). Subjective measures (i.e. ratings by instructors) can be more appropriate, but are subject to error and bias, and can be difficult to develop. In aircrew coordination training, for example, optimal ratings of crew coordination skills should take place in the operational aircraft, which is difficult, if not impossible, in most cases. For this reason, Benson et al. (6) and others (e.g., 8) recommend use of flight simulators to provide necessary performance information. Helmreich and associates (e.g., 17, 11) designed an aircrew performance evaluation scheme that successfully trained expert observers in aircrew coordination, and then had them observe and rate crews during full mission simulation and during operational flight. Video taping crew performance and deriving performance ratings from the tapes can make this an easier task, and enhance accuracy since tapes can be replayed, stopped, etc.

In order to be useful, rating schemes must be developed and implemented carefully. The approach advocated here is to rely on needs analysis data indicating the specific behaviors that are characteristic of effective aircrew coordination. These behaviors should form the basis of an observational protocol to be used by training instructors in evaluating crews. The criteria and associated scoring procedures used in the rating scheme should then be determined empirically; studies with flight instructors and student pilots are being conducted at present by NAVTRASYSCEN for this purpose.

Other potential performance ratings include trainee evaluations of their own or other crew members' performance. Awe and Murphy (21), for example, successfully employed peer ratings of aircrew coordination and crew decision making. Such measures may have an advantage over instructor ratings, since peers may be less threatening than instructors and may be in a better position to observe critical behaviors than outside observers. Self-evaluations have met with somewhat less success than peer ratings for aircrew coordination. Helmreich and Wilhelm (17) found that self-evaluations of performance were inflated (i.e., higher than they should have been). However, as a means to provide feedback to trainees, self-evaluations may be helpful. Crews must be given information that helps them form a realistic picture of their aircrew coordination skills, so that they are aware of areas that need improvement.

Mission Effectiveness

The final level of Kirkpatrick's evaluation hierarchy is overall organizational impact. This refers to the extent to which a training program has an impact on overall organizational performance. Such measurement has traditionally been most difficult to accomplish since a host of extraneous factors influence mission performance and serve to confound assessments related to performance in a specific domain.

In the case of aircrew coordination, mission effectiveness can be defined as a significant reduction in aircraft accidents and incidents. With respect to accidents, these occur so infrequently that it may be impossible to establish a statistical relationship between training and accident reduction. Incidents, on the other hand, occur often enough to provide a meaningful index of aircrew coordination performance. With careful analysis and coding, incidents that are caused by poor aircrew coordination can be tracked before and after the introduction of training. Of course, the feasibility of using such data will depend on its accuracy, availability and completeness.

SUMMARY

The multi-component evaluation plan for aircrew coordination training advocated here has the potential to provide comprehensive, reliable information that can enhance training system design, indicate training effectiveness at various levels, and provide feedback to trainees, instructors, training administrators and training designers. In addition, the probability that the program's effectiveness will be amenable to evaluation is increased via recognition of evaluation needs throughout the training design and development cycle. Overall, the approach delineated here will allow training designers to gather relevant, valid and useful information about aircrew coordination training for incorporation into programs introduced at flight replacement squadrons.

REFERENCES

[1] Rutman, L. (1980). Planning Useful Evaluations. Beverly Hills, CA: Sage Publications.

[2] Rossi, P. H., & Freeman, H. E. (1982). Evaluation: A Systematic Approach, (2nd Edition). Beverly Hills, CA: Sage Publications.

[3] Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig (Ed.), Training and Development Handbook: A Guide to Human Resource Development. New York, NY: McGraw-Hill.

[4] Goldstein, I. L. (1986). Training in Organizations: Needs Assessment, Development, and Evaluation. Monterey, CA: Brooks/Cole Publishing Company.

[5] Prince, C., Cannon-Bowers, J. A., Salas, E., Owens, J., & Morgan, B. B., Jr. (in preparation). Aircrew coordination training: Requirements and challenges. Paper submitted to the 11th Interservice/Industry Training Systems Conference.

[6] Benson, P. G., Riddlegarger, S., Hornsby, J. S., & Salas, E. (1988). Assessment of Team Performance Measurement Techniques for Training Systems, Technical Report Number DAALO3-86-d-0001. Orlando, FL: Naval Training Systems Center.

[7] Wexley, K. N., & Latham, G. P. (1981). Developing and Training Human Resources in Organizations. Glenwood, IL: Scott, Foresman and Company.

[8] Orlady, H. W., & Foushee, H. C. (1987). Cockpit Resource Management Training, Technical Report Number NASA CP-2455 Moffett Field, CA: NASA Ames Research Center.

[9] Chidester, T. R. (1987). Use of full-mission simulation to identify variables influencing crew effectiveness. In Proceedings of the Commission of European Communities Workshop on Simulation in Traffic Systems. Bremen, West Germany.

[10] Lauber, J. K., & Foushee, H. C. (1981). Guidelines for Line Oriented Flight Training, Volumes I & II, Technical Report Number NASA CP-2184. Moffett Field, CA: NASA Ames Research Center.

[11] Helmreich, R. L., & Wilhelm, J. A. (1987). Evaluating Cockpit Resource Management Training. In Proceedings of the Fourth International Symposium on Aviation Psychology. Columbus, Ohio: The Ohio State University.

[12] Helmreich, R. L., & Wilhelm, J. A. (1988). Reinforcing and Measuring Flight Crew Resource Management: Training Captain/Check Airman/ Instructor/Reference Manual, NASA/UT Technical Manual Number 87-1. Moffett Field, CA: NASA Ames Research Center.

[13] Jensen, R. S. (1987). Cockpit Resource Management Training, Final Report, Contract Number DTFA01-80-C-10080. Arlington, VA: Systems Control Technology, Inc.

[14] Noe, R. A., & Schmitt, N. (1986). The influence of trainee attitudes on training effectiveness: Test of a model. Personnel Psychology, 39, 497-524.

[15] Hicks, W. D., & Klimoski, R. J. (1987). Entry into training programs and its effects on training outcomes: A field experiment. Academy of Management Journal 3, 542-552.

[16] Vroom, V. H. (1964). Work and Motivation. New York, NY: Wiley.

[17] Helmreich, R. L., & Wilhelm, J. A. (1986). Evaluating CRM Training: I. Measures and Methodology, NASA/UT Technical Report Number 86-8. Moffett Field, CA: NASA Ames Research Center.

[18] Glickman, A. S., Zimmer, S., Montero, R. C., Guerette, P. J., Campbell, W. J., Morgan, B. B., & Salas, E. (1987). The Evolution of Teamwork Skills: An Empirical Assessment with Implications for Training. Technical Report Number TR87-016, Orlando, FL: Naval Training Systems Center.

[19] Helmreich, R. L., Chidester, T. R., Foushee, H. C., Gregorich, S., & Wilhelm, J. A. (1989). How Effective is Cockpit Resource Management Training? Issues in Evaluating the Impact of Programs to Enhance Crew Coordination, NASA/UT Technical Report Number 89-2, Draft 6.0. Moffett Field, CA: NASA Ames Research Center.

[20] Rockway, M. R., & Nullmeyer, R. T. (1987). The relative utility of various types of performance measures for aircrew training and evaluation. In Proceedings of the Fourth International Symposium on Aviation Psychology. Columbus, Ohio: The Ohio State University.

[21] Awe, C. A., & Murphy, M. R. (1986). Aircrew coordination and decision-making: Peer ratings of video tapes made during a full mission simulation. In Proceedings of the 21st Annual Conference on Manual Control. Washington, D.C.: NASA Scientific and Technical Information Office.

[22] Huczynski, A. A., & Lewis, J. W. (1980). An empirical study into the learning transfer process in management training. Journal of Management Studies 17, 227-240.

[23] Jensen, R. S., & Adrion, J. (1988). Aeronautical Decision Making for Commercial Pilots. Report Number DOT/FAA/PM-86/42. Washington, D. C.: U.S. Department of Transportation, Federal Aviation Administration.

[24] Cullen, J. G., Sawzin, S. A., Sisson, G. R., & Swanson, R. A. (1976). Training: What's it worth. Training and Development Journal, 1976, 30, 12-20.

ABOUT THE AUTHORS

DR. JANIS A. CANNON-BOWERS is a Research Psychologist with the Human Factors Division at the Naval Training Systems Center (NTSC). She has been involved in R&D in team training and performance, aircrew coordination training, trainability factors, training for tactical decision making and expert tutoring systems.

DR. CAROLYN PRINCE is a research psychologist with the Human Factors Division of the Naval Training Systems Center, where she has worked on a variety of aviation training issues. Prior to government employment, she worked as a consultant in management selection and development. DR. EDUARDO SALAS is a research psychologist with the Human Factors Division of the Naval Training Systems Center. He is principle investigator for NTSC's R&D program on team training and performance, and project manager of an effort to develop aircrew coordination training for Navy and Marine platforms. He has done research on training evaluation, human performance assessment, job task analysis, skill acquisition and personnel psychology.

COMMANDER JERRY M. OWENS is a Naval Aerospace Experimental Psychologist and holds a Ph.D. in Experimental Psychology. He has served in various assignments related to aviation RDT&E management and execution, including Head, Psychological Sciences Department, Naval Aerospace Medical Research Laboratory from 1981-1984. He has performed research in the areas of aviation personnel screening and performance prediction, performance assessment in high demand environments, sustained performance and voice interactive systems technology. On the staff of the Program Manager for Aviation Training Systems (PMA205), Commander Owens currently serves as Assistant Deputy for Programs and Evaluation and provides training human factors support to the V-22.

DR. BEN B. MORGAN, JR. is a Professor of Psychology at the University of Central Florida. In the past, he has served as Director of the Center for Applied Psychological Studies and as Chair of the Department at Old Dominion University. His major research interests include human performance assessment, the performance effects of stress, the recovery of performance from stress effects, the acquisition of complex skills, and team training and performance. GREGORY H. GONOS works in the Special Projects and Analysis Division of the Naval Training Systems Center. He is currently completing a M.A. degree in Industrial/Organizational Psychology at the University of Central Florida. He has been involved in research into team performance and training, aircrew coordination training and trainability factors.