Incalcaterra, K.A., & Holt, R. W., (1999). Pilot evaluations of ACRM programs. Proceedings of the Tenth International Symposium on Aviation Psychology. The Ohio State University: Columbus, OH.

Pilot Evaluations of ACRM Programs

Kara A. Incalcaterra and Robert W. Holt
George Mason University
Fairfax, VA

ABSTRACT

A survey was administered to pilots who were trained in Advance Crew Resource Management (ACRM) at a regional carrier. The survey was designed to assess several effects of training including, attitudes toward CRM and ACRM, knowledge and practice of ACRM, perceived effects of ACRM, as well as endorsement of ACRM. Data analysis revealed that pilots have positive attitudes toward CRM and ACRM, know the content and timing of ACRM procedures, practice normal ACRM procedures regularly in routine flights, perceive positive consequences of using ACRM procedures, and endorse ACRM. Data also shows that attitudes toward CRM and ACRM positively predict the practice and perceived effects of ACRM, as well as endorsement of ACRM. In general, the results provide indirect support for the ACRM program.

INTRODUCTION

Recently, the United Stated General Accounting Office issued a report on an investigation by the National Transportation Safety Board (NTSB) which looked at 169 accidents between 1983-1985 involving major airline carriers (Dillingham, 1997). Of the 169 accidents, NTSB concluded that about 30 percent were caused in part by pilots' performance. Furthermore, it was reported that improper use of CRM principles (Cooper, White, & Lauber, 1980) accounted for at least one-third of this 30 percent.

Pilot or crew failure to properly use the CRM philosophy and principles in the cockpit may be attributed to the philosophy and principles being unsystematically trained. Pilots are technically trained to perform most flight operations in a sequential, detailed, and defined way. CRM training has traditionally been directed toward changing attitudes and not changing behaviors the way technical skills have been trained. Advanced Crew Resource Management (ACRM) was designed to translate CRM principles into more sequential, systematic, and tangible procedures for crew training and evaluation.

In a quasi-experimental design, ACRM was introduced in one of two comparable turboprop fleets at a regional airline in the Eastern U.S. The other fleet served as the control fleet. At the conclusion of this four-year research project the effectiveness of ACRM training was evaluated. Following Cronbach & Meehl's principle of converging operations (1955), our research team considered multiple training evaluation operations or methods for evaluating the effects of ACRM at the crew and individual pilot levels of analysis. The focus of this paper is devoted to the development, results, and implications of a pilot survey that was designed to assess the effectiveness of ACRM at the individual pilot level of analysis. [1]

Overview of the Pilot Survey

A major focus of the pilot survey was attitudinal reactions to ACRM. Attitudes were correlated with several performance and outcome measures including ACRM knowledge and practice, as suggested by Kraiger, Ford, & Salas (1993), as well as perceived effects and endorsement of ACRM.

Our interest in the investigation of the relationship between attitudes and behavior is rather straightforward. If you feel good about something, you are much more likely to do it. Helmreich, Merritt, & Wilhelm (1999) note that "although attitudes are not perfect predictors of behavior, it is a truism that those whose attitudes show rejection of CRM are unlikely to follow its precepts behaviorally" (p. 23-24). Yet, research has shown that conceptually sound and carefully controlled measurement of the attitude-behavior relationship is necessary to establish the attitude-behavior link (Fishbein and Ajzen , 1975).

In the aviation domain, researchers have specifically considered the relationship between attitudes toward CRM and flight deck performance. Helmreich, Foushee, Benson, & Russini (1986) investigated the relationship between attitudes regarding cockpit management and line flying performance. In this study attitudes toward CRM for 163 pilots were measured using the Cockpit Management Attitudes questionnaire (Helmreich, 1984) and discriminant analysis indicated that these attitudes were a significant predictor of Check Airmen's ratings of pilot performance. This suggests that attitudinal measures can be useful in assessing the behavioral impact of this type of training (Helmreich, et al., 1986).

In all, investigating the relationship between attitudes and behavior is an indirect approach to

[1] For a complete description of each piece of the ACRM evaluation process please see Holt, Boehm-Davis, Ikomi, Hansberger, Beaubien, Incalcaterra, Seamster, Hamman, and Schultz (1998).

assessing the potential effectiveness of resource management training efforts (Helmreich, et al., 1986). As such, this method of ACRM training evaluation complemented the other more direct methods we were concurrently using to assess the overall effectiveness of ACRM training.

Research Hypotheses

To evaluate ACRM training, we considered the following hypotheses:

1. ACRM training should have positive results. Specifically, pilots trained in ACRM (a) should have positive attitudes toward CRM and ACRM, (b) should know the content and timing of ACRM procedures, (c) should practice the normal ACRM procedures regularly in routine flight operations, (d) should perceive positive consequences of using ACRM procedures, and (e) should endorse ACRM.
2. ACRM attitudes should predict ACRM knowledge, practice, perceived effects, and endorsement of ACRM. Specifically, for pilots trained in ACRM their attitudes toward CRM in general and ACRM in particular should positively predict their (a) knowledge about ACRM, (b) practice of ACRM, (c) perceived effects of ACRM, and (d) endorsement of ACRM.

METHOD

Sample

Surveys were distributed to 600 active pilots at a regional airline carrier including both ACRM and non-ACRM trained pilots. There were 184 pilots in the "ACRM-trained" group and 84 pilots in the "non-ACRM" group. The hypotheses presented in this paper involve only the "ACRM-trained" sub-sample of the data collected. For a complete description of all hypotheses related to the ACRM-trained and non-ACRM groups, please see Incalcaterra and Holt (in progress).

Measures

Affective Reactions to CRM. CRM items were taken from the "coordination and communication" and "command control" sub-categories of the CMAQ questionnaire originally developed by Helmreich (1984). Based on feedback from airline pilots and personnel, three of the original CMAQ items were deleted, three items were re-worded, and one item was added. The research team excluded the items from the "stress and fatigue" sub-category of the CMAQ because these items were not consistent with the CRM training philosophy at this airline carrier. For each item, the possible responses were Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree (Likert, 1932).

Overall Impressions of ACRM. Specific attitudes toward ACRM were measured by 5-point bipolar adjective scales (Osgood, 1954) that included Neutral as the midpoint. All of the overall affective impressions items were developed by our research team (academics, pilots, and airline personnel) through a content validation process.

Knowledge, Practice, and Effects Questions. Knowledge was only measured post-training because the ACRM training introduced completely new procedures and nomenclature developed for this project which pilots could not have known about previously. The knowledge questions consisted of 8 multiple-choice items with three alternatives each. These questions were taken directly from information on the ACRM briefing card given to all ACRM trained pilots and concerned the basic content and the appropriate timing of normal ACRM procedures. Unlike abnormal or emergency procedures, the normal ACRM procedures were procedures that should have been frequently performed on a routine basis. For these items, one alternative was correct and the other two incorrect, so that the random baseline of correct answers was one-third.

Our research team developed the items that assessed the practice of ACRM procedures and the effects of ACRM by phase of flight. The eight ACRM practice items all concerned ACRM normal procedures that would occur regularly in line operations (no abnormal or emergency procedures). For each item the pilot estimated whether he or she Always, Frequently, Sometimes, Rarely, or Never performed that procedure. Given that the execution of ACRM normal procedures may depend on situational factors, these procedures are not always performed but should be performed at least *sometimes* during normal flight operations.

The perceived effects of ACRM were measured by 12 items. The items were organized by Departure and Arrival phases of flight, as these phases were considered to be high workload phases for which there were specific ACRM procedural innovations. Each item cited a possible positive effect of ACRM, to which the pilot could respond Strongly Agree, Agree, Neutral, Disagree, or Strongly Disagree (Likert, 1932).

Endorsement of ACRM

Two yes/no endorsement questions developed by our research team were included at the end of the survey. The first question assessed pilot endorsement of ACRM in terms of support for the continuation and expansion of ACRM. The second question assessed pilot endorsement of ACRM in terms of support for the

notion that ACRM has improved their flight performance.[2]

Procedure

600 pilot surveys were hand delivered to each pilot's personal correspondence folder in the pilot lounge at the carrier's hub airport by members of our research team. As most pilots check their personal correspondence folder on a regular basis, this was deemed the most appropriate and economical way to distribute the pilot survey.

About two weeks after the surveys were distributed, pilots were given a reminder note reminding and encouraging them to complete and return their survey to the research team. Data collection lasted for about four months, and ceased when less than one survey per week was arriving by mail. In total, this yielded a 45% response rate (32% "non-ACRM" group and 68% "ACRM-trained" group).

RESULTS

Hypothesis 1: ACRM training should have positive results and ACRM-trained pilots should vote for ACRM

CRM Attitudes. For each of the thirteen items, the distribution of responses was compared to the "Neutral" baseline by a one-sample t-test. The results of these tests are based on a sample size of 179 or more pilots. The results for all items except one were significantly greater than the Neutral baseline ($p< .001$), indicating that pilots trained in ACRM have overwhelmingly positive attitudes toward CRM.

ACRM Attitudes. For each of the seven ACRM attitude items, the distribution of responses was compared to the "Neutral" baseline by a one-sample t-test. The results of these tests are based on a sample size of 175 or more pilots. The results for all items were significantly greater than the Neutral baseline ($p< .001$) except for one item regarding the effect of ACRM on workload ("ACRM adds to workload"). The mean for the workload item was 3.03, which is essentially Neutral on the five point response scale. In all, ACRM-trained pilots are positive toward ACRM with the exception of being neutral about whether ACRM is adding to workload.

ACRM Knowledge. For each of the seven items, the proportion of pilots answering correctly was compared to a random-guessing baseline of .33 using a binomial test. Each non-parametric t-test analysis was based on a sample size of 166 or more pilots. All ACRM knowledge items were answered correctly above the chance baseline of 33 percent ($p< .001$). Except for one item concerning the exact context for using a Statement of Condition ("When is the Statement of Conditions procedure used?") a strong majority of the pilots answered these knowledge questions correctly (76% - 98.8%).

To put this performance in context, for most of these pilots the relevant ACRM training occurred slightly over two years before this survey. Despite this delay between ACRM training and ACRM knowledge assessment, most ACRM-trained pilots had correct knowledge of the content and timing of ACRM procedures. In part, this accurate knowledge could be due to regular practice of the ACRM procedures as indicated by the next set of results.

Practice of ACRM procedures. To test if the average reported frequency was at least *sometimes* or greater, the distribution of responses from the eight practice items was compared to the "Sometimes" baseline (3.0) by a one-sample t-test. The results of these tests are based on 176 or more pilots. The results for all practice items were significantly greater than the "Sometimes" baseline ($p< .001$). The two items with an average frequency between "Sometimes" and "Frequently" were Briefing a Statement of Conditions on the Clearance Brief ("I give a Statement of Conditions in the Clearance Brief") and briefing crew coordination issues on the Approach Brief ("I include crew coordination issues on the Approach Brief"). The other six ACRM items were performed between "Frequently" and "Always". To the extent that these self-reports of enacting the ACRM procedures are accurate, the ACRM procedures are frequently performed in line operations.

Perceived effects of ACRM. For each of these twelve items, the distribution of responses was compared to the "Neutral" baseline by a one-sample t-test. The results of these tests are based on a sample of 176 or more pilots. The results for all items were significantly greater than the Neutral baseline ($p< .01$). Although the means for two items regarding avoiding distractions during departure ("Crews avoid distractions during departure") and conducting briefs during low workload periods of arrival ("Crews conduct briefs during low workload periods of arrival") are between Neutral and Agree, most of the means are either at the Agree level (4.0) or between Agree and Strongly Agree. Overall, these results indicate that the pilots perceive a broad variety of positive effects of the ACRM procedures.

Pilot votes on ACRM. A binomial test of the proportions of pilots endorsing ACRM was positive (alpha .001). About 93% of the pilots (N=177) voted to expand the use of ACRM to other fleets. Similarly, 91% of the pilots (N = 180) endorsed the statement that

[2] Please see Incalcaterra and Holt (in progress) for a copy of the complete pilot survey.

ACRM had improved their flight performance. While neither vote is 100%, the observed percentages are significantly higher than a chance baseline of .50. In addition, power was considerably high for each of these analyses (β is approximately .77 in both cases).

Hypothesis 2: ACRM attitudes should positively predict knowledge, practice, perceived effects, and endorsement of ACRM

Before examining hypothesis two we condensed the multiple items purporting to measure the CRM attitudes, ACRM attitudes, ACRM knowledge, practice and perceived effects constructs into conceptually coherent unit-weighted scale composites. This process was not necessary for the two single-item endorsement measures. Internal consistency reliability analyses indicate that all but one of the scale composites had reliability values at or above the .70 criteria generally accepted for applied research. For the CRM composite scale (13 items) alpha = .6648, ACRM composite scale (7 items) alpha = .8170, Practice composite scale (7 items) alpha = .7522, and Perceived Effects composite scale (12 items) alpha = .9127. The reliability for the Knowledge composite scale (7 items) was near zero indicating that the items designed to tap ACRM knowledge did not hang together well. A simple explanation for this finding may be that pilots' cognitive knowledge structure for ACRM is not a coherent, tightly linked model representative of the knowledge items presented in the survey. Alternatively, it may be the case that the items written to tap ACRM knowledge fall within the broader ACRM knowledge domain, yet represent independent pieces of ACRM knowledge. In any event, factor analysis is arguably a more valid way to establish scale scores based on uni-dimensional factors or latent constructs. Because these are preliminary analyses, factor analytic techniques were not used but will be reported in the final technical report (see Incalcaterra & Holt, in progress).

Correlations. CRM and ACRM attitudes were correlated with knowledge, practice, perceived effects, and endorsement of ACRM. The results are presented in the Table 1. Attitudes toward CRM and ACRM all significantly and positively predict the practice of ACRM, the perceived effects of ACRM, endorsement of the idea that ACRM should be continued and expanded, and endorsement of the idea that ACRM improved flight performance. ACRM knowledge did not correlate significantly with CRM or ACRM attitudes.

Multiple Regression. The noticeable correlation between the CRM attitudes and ACRM attitudes (r=.534, $p<.01$) makes it difficult to tease apart the relative contribution of one predictor over and above the other. Therefore, the possible impact of general CRM versus specific ACRM attitudes on ACRM performance, perceived effects, and endorsement was further analyzed with multiple regression. CRM and ACRM attitudes were block entered into an ordinary least square regression equation for the knowledge, practice, and perceived effects criteria. The predictors were also entered into a logistic regression equation for each dichotomously scored endorsement criterion.

Taken together, general CRM and specific ACRM attitudes are strong predictors of all performance criteria except ACRM knowledge. Multiple R was significant for the practice composite scale (R=.452, $p<.001$) indicating that together CRM and ACRM attitudes predicted about 20% of the total variance in practicing ACRM. Likewise, multiple R was significant for the perceived effects composite scale (R=.680, $p<.001$) indicating that together CRM and ACRM attitudes predicted about 46% of the total variance in the perceived effects of ACRM. However, multiple R was non-significant for the knowledge composite scale (R=.056, *ns*).

Taken together, general CRM and specific ACRM attitudes are also strong predictors of the two endorsement items. Logistic regression indicated that the log likelihood estimation of pilot classification as to whether or not they supported the continuation and expansion of ACRM was significantly improved when the predictors were added into the regression equation (Model chi square = 30.732, $p<.001$). In fact, 95.68% of the pilot classifications based on the logistic regression equation matched their true response to the endorsement item when the two attitudinal predictors were used. Likewise, logistic regression indicated that the log likelihood estimation of pilot classification as to whether or not they supported the idea that ACRM improved their flight performance was significantly improved when the predictors were added into the regression equation (Model chi square = 28.692, $p<.001$). In this case, 92.73% of the pilot classifications based on the logistic regression equation matched their true response to the endorsement item when the two attitudinal predictors were used.

A closer look at the regression analyses is useful for determining the unique contribution of each predictor in explaining variance in each criterion. First, results indicated that CRM attitudes significantly and strongly predicted pilot ratings of practicing ACRM in the cockpit (β=.299, $p<.001$) and ACRM attitudes also significantly, yet not quite as strongly, predicted pilot ratings of practicing ACRM in the cockpit (β=.217, $p<.01$). A significance test for the difference between dependent regression weights indicated that indeed these two regression weights were significantly different from each other ($t(167)=8.39$, $p<.01$).

Furthermore, the squared semi-partial correlations reveal that CRM attitudes uniquely accounted for 7% of the total variance in practicing ACRM, while ACRM attitudes uniquely accounted for 3% of the total variance in practicing ACRM. While neither predictor uniquely explains a large portion of the variance in the dependent variable, the overall multiple R is fairly substantial, calling attention to the moderate degree of collinearity between the two attitudinal predictors.

Next, CRM attitudes strongly predicted pilot perceptions of the perceived effects of ACRM (β=.221, p<.01). Similarly, ACRM attitudes strongly predicted pilot perceptions of the perceived effects of ACRM (β=.533, p<.001). A significance test for the difference between dependent regression weights indicated that again, these two regression weights were significantly different from each other (t(154)=89.31, p<.01). Furthermore, the squared semi-partial correlations reveal that ACRM attitudes uniquely accounted for about 20% of the total variance in perceived effects of ACRM, while CRM attitudes uniquely accounted for only 3% of the total variance (almost 7 times less than that of ACRM attitudes). It is clear in this case, that ACRM attitudes are a substantial predictor of the degree to which pilots perceive ACRM to be effective.

Finally, logistic regression beta weights indicate that for both endorsement items only ACRM attitudes were a significant predictor (β=3.11, p<.001 for the "expand" endorsement item and β=2.26, p<.001 for the "improve" endorsement item). It is no surprise that specific attitudes toward ACRM predicted ACRM endorsement responses. However, it is intriguing that general CRM attitudes did not significantly contribute to this prediction. Accordingly, this is evidence that pilots' attitudes toward CRM are to some degree independent from their feelings about endorsing ACRM.

CONCLUSIONS

In general, the results of the pilot survey support the ACRM program. Hypothesis 1 considered the degree to which ACRM-trained pilots converged on positive ACRM attitudes and performance-related outcomes. The results indicated that when compared to a baseline, ACRM-trained pilots as a group had very positive attitudes toward CRM in general and ACRM in particular. In addition ACRM-trained pilots knew the content and timing of ACRM, performed ACRM procedures frequently, perceived positive effects of ACRM, and overwhelmingly endorsed ACRM when it was put to the vote. These results speak to the success of training in terms of affective outcomes, knowledge outcomes, skill-based outcomes, and perceived effectiveness and endorsement beliefs.

Consistent with Hypothesis 2, bivariate correlations indicated that CRM and ACRM attitudes positively predict pilot practice, perceived effects and endorsement of ACRM. Multiple regression analyses supported this finding and also indicated that general CRM attitudes and specific ACRM attitudes differ in their ability to uniquely predict the various outcome measures.

Specifically, results revealed that CRM attitudes were a significantly stronger predictor of practicing ACRM than were ACRM attitudes. However, after considering the fact that CRM attitudes uniquely account for about 7% of the total variance in practicing ACRM, while ACRM attitudes uniquely account for about 3%, we concluded that while this difference may be statistically significant, it is not practically significant. Next, results revealed that ACRM attitudes were a significantly stronger predictor of the perceived effects of ACRM than were CRM attitudes. In this case, the fact that almost 20% of the variance in the perceived effects criterion was being uniquely accounted for by ACRM attitudes, whereas CRM attitudes only accounted for 3%, lead us to conclude that this was not only a statistically significant difference, but one of practical value as well. Furthermore, results from the logistic regression analyses indicated that for both endorsement items, only the ACRM attitudes predictor contributed significantly to accounting for variance in the criterion.

This evidence leads us to two general conclusions. First, CRM attitudes and ACRM attitudes are somewhat qualitatively distinct, giving rise to the difference in their predictive ability. We believe that ability of ACRM to predict different training outcomes over and above CRM attitudes may be due at least in part to the similarity between ACRM procedures and those for normal technical flight operations. Following from our results, we could hypothesize that pilots see technical procedures as being more congruent with ACRM procedures than with the general CRM philosophy and principles. Further research is needed to address this issue specifically.

Second, this evidence provides support for the hypothesized relationship between attitudes and performance. This provides indirect support for the effectiveness of ACRM training, and also gives us some indication of the degree to which pilots are actually using ACRM procedures in cases of typical flight performance. However, because data were cross sectional, we are unable to demonstrate a direct causal linkage from CRM and ACRM attitudes to performance-related outcomes. Ideally, this could be done by studying the relationship between attitudes and performance over time, showing that changes in attitudes temporally precede behavioral changes (Mill, 1852).

Incalcaterra, K.A., & Holt, R. W., (1999). Pilot evaluations of ACRM programs. Proceedings of the Tenth International Symposium on Aviation Psychology. The Ohio State University: Columbus, OH.

REFERENCES

Boehm-Davis, D.A., Holt, R.W., Ikomi, P.A., Hansberger, J.T., Beaubien, J.M., Incalcaterra, K.A., Seamster, T.L., Hamman, W.R., & Schultz, K. (1998). CRM procedures and crew performance: Executive summary of Advance Crew Resource Management. Technical Report. George Mason University, Fairfax, VA.

Cooper, G.E., White, M.D., & Lauber, J.K. (1980). Resource management on the flightdeck: Proceedings of a NASA/Industry Workshop (Rep. No. NASA CP-2120). Moffett Field, CA: NASA-Ames Research Center.

Cronbach & Meehl (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Dillingham, G.L. (1997). Human factors: FAA's guidance and oversight of pilot crew resource management training can be improved. Report to congressional requesters by the United States General Accounting Office. (GAO/RCED-98-7).

Fishbein, M. & Ajzen, I. (1975). Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley.

Gregorich, S.E., Helmreich, R.L, & Wilhelm, J.A. (1990). The structure of cockpit management attitudes. Journal of Applied Psychology, 75(6), 682-690.

Helmreich, R.L. (1984). Cockpit management attitudes. Human factors, 26, 583-589.

Helmreich, R.L., Foushee, H.C., Benson, R., & Russini, W. (1986). Cockpit management attitudes: Exploring the attitude-behavior linkage. Aviation, space, and environmental medicine, 57, 1198-1200.

Helmreich, R.L., Merritt, A.C., & Wilhelm, J.A. (1999). The evolution of crew resource management training in commercial aviation. The International Journal of Aviation Psychology, 9(1), 19-32.

Incalcaterra, K.A. & Holt, R.W. (in progress). Pilot Evaluations of Advanced Crew Resource Management (ACRM) Programs. Technical Report. George Mason University, Fairfax, VA.

Kraiger, K., Ford, J.K., & Salas, E. (1993). Application of cognitive, skill-based and affective theories of learning to new methods of training evaluation. Journal of Applied Psychology [Monograph], 78(2), 311-328.

Likert, R. (1932). A technique for measurement of attitudes. Archives of Psychology, 140, 44-53.

Mill, J.S. (1852). A system of logic. New York: Harper & Brothers.

Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. (1957). The measurement of meaning. Urbanan, IL: University of Illinois Press.

Seamster, T.L., Boehm-Davis, D.A., Holt, R.W., & Schultz, K. (1998). Developing Advanced Crew Resource Management (ACRM) Training: A Training Manual. Washington, DC: Federal Aviation Administration, AAR-100.

Table 1: Pearson Product-Moment Correlations (N=126)

		1.	2.	3.	4.	5.	6.	7.
1.	CRM Attitudes	1.0	.563**	-.010	.401**	.539**	.191*	.276**
2.	ACRM Attitudes		1.0	-.058	.430**	.637**	.478**	.516**
3.	ACRM Knowledge			1.0	.019	-.058	.086	.012
4.	ACRM Practice				1.0	.437**	.292**	.272**
5.	Perceived Effects					1.0	.339**	.397**
6.	Expand ACRM						1.0	.748**
7.	Improve Flight Performance							1.0