EVALUATING THE EFFECTS OF MAINTENANCE RESOURCE MANAGEMENT (MRM) IN AIR SAFETY[1]

James C. Taylor, Ph.D.
School of Engineering
Santa Clara University
Santa Clara, CA 95053-0590

Jean Watson
Federal Aviation Administration

October 24, 2000

SUMMARY

This research project was designed to help understand, evaluate and validate the impact of Maintenance Resource Management (MRM) training programs, and other MRM interventions, on participant attitudes, opinions, behaviors, and ultimately on enhanced safety performance. It includes research into evaluation methodology as well as examination of the range of change models used by the aviation companies studied.

Longitudinal comparisons. The present report begins with the measures of attitudes and opinions about "communication and coordination," "shared decision making," "assertiveness," "stress management, "goal sharing" and "safety practice" before training is undertaken. After examining changes over time in participant attitudes and opinions, assessment measurements were made of self-reported intentions to change behaviors and attitudes, and self-reports of subsequent changes actually made. Enthusiasm for the training, and participant reaction to the training, are reviewed. The design of this project's evaluation method incorporates a longitudinal, time sensitive data collection approach by comparing the baseline and pre-training measures with follow-up attitude and opinion surveys (immediately after training, and two, six and twelve months following training). Finally these attitudes and opinions are tested for correlations with monthly or quarterly safety performance measures.

Benchmark comparisons. Data collected since 1991, using the same standardized questions, are used to create profiles of benchmark MRM attitudes and opinions – against which specific MRM sites can be compared with standard scores for any of the waves (i.e., pre-training, post-training, 2-, 6-, 12-month) measured.

This report focuses on two of those companies ("A," and "D"). In the first case ("A") analysis is directed to the attitudes and reported behaviors of Aviation Maintenance Technicians (AMTs) and the safety performance of their maintenance work units following a one-shot safety awareness training program. Examination of Company A data reveals a mainly positive effect of the initial training, as well as some signs that additional intervention such as recurrent training is required in order to sustain enthusiasm and thus continue MRM success. Company D analysis examines the effects of distributing its MRM training over several months for each participant while concentrating the training in one city/station at a time. These results are set against the backdrop of the four other companies' MRM experience from 1998-1999, as well as the "benchmark" results from all MRM programs studied since 1991.

OBJECTIVES

The present report describes the 1998-1999 results of examining several programs implemented by commercial aviation maintenance operations headquartered in the USA.

These programs represent varied combinations of several "change models," for MRM interventions. Those change models include "individual *vs.* system change," "behavior *vs.* awareness," "one-time *vs.* ongoing change," and "MRM for AMTs-only *vs.* MRM for all Tech Ops personnel."

The larger implication of these change models is discussed below in that section addressing the effect of national, occupational, and organizational "cultures" on the creation of a "safety culture."

The evaluation methodology applied to the several MRM programs has involved the creation of reliable and valid psychological scales (both attitude scales and opinion scales), examination of "before-after" changes in the attitude scales, valid survey measurement of expected future behaviors, and self-reports of behaviors changed. These survey data are also compared with subsequent reports of behavior changes. The methodology further includes correlating post-training attitudes and opinions with subsequent safety performance, the calculation of financial returns for MRM programs, and the collection and documentation of field observations and interviews.

The multiple purposes of the evaluation methodology include 1) the scientific and objective measurement of MRM program success, 2) the discovery and dissemination of "best practices," and 3) the development of simple and accessible evaluation tools for use by the partner companies themselves.

BACKGROUND

Research activities in the field of "macro" human factors in aviation maintenance since the early 1990s indicate that many airlines have opted to improve awareness of communication, safe practices, and professionalism. A few of these programs have also included skill-based training in decision-making, employee participation, assertiveness, and effective safety systems. "Maintenance Resource Management," (MRM) is the generic term used here for these programs. Most of these airline MRM efforts have focused on training employees – some emphasizing training mechanics (or aviation maintenance technicians – AMTs) and others training both maintenance management and AMTs. As reported earlier (Taylor & Robertson, 1995; Taylor, 1998) the training in both cases has succeeded in achieving initial and significant attitude changes among the participants. On the other hand there is little evidence that system change follows, or is sustained following, the initial, individual, attitude improvement. The data continue to show that individual, awareness-only efforts, and initial enthusiasm for MRM fail to create an effective safety culture conjoined with sustained safety performance.

This research project is a university-industry-government-labor partnership that involves numerous airline operators, and aviation maintenance facilities based in the U.S. It is planned, integrated, and coordinated with the FAA/NASA Aviation Safety Program (ASP). The research program has earned the reputation of representing the "real world" of aviation maintenance and addressing maintenance human factors issues accordingly. It has raised the awareness of the importance of "macro human factors" to the aviation industry, and has served a number of organizations that have introduced MRM programs, some of which have been specifically designed to reduce maintenance errors

Since February 1998 the research program has entered into to its databases over 19,000 surveys completed by MRM program participants. 1998-1999 survey data have been combined with the 19,000 surveys contained in the program's databases from prior years (Taylor, 1998). This combined database has been used to create a template of "typical responses" (i.e., a normative profile) to which any airline or repair station can compare its own workforce at any stage in it's MRM program.

Historically, this program has provided to each airline partner a timely feedback of its own survey data. The measurement tools, produced as one of the program's long range objectives, have also been made available to the airlines as the development of such tools is completed. During 1998-1999 this tool development and feedback process has continued.

Two new tools developed for use by MRM developers and administrators are reported here. One tool is the "MRM attitude and opinion profile." It is the calculation of percentile scores for any maintenance work unit or site. These profiles in the form of standard scores ("Z") can be used to compare the percentile rank of MRM attitudes and opinions in any given company at any stage in its MRM program with attitudes from a large database of like employees – called the "Benchmark dataset" -- during a similar period in their MRM involvement. This panel of Benchmark comparison profiles for attitudes and opinions is one tool to help audit the relative effectiveness of a maintenance human factors program. The second tool, described below, helps assess the financial return on investment of MRM programs. Called MRM/ROI, this tool helps justify MRM through cost savings from improved safety – and additionally in proportion to the program's measurable effect on a safety outcome.

MRM CHANGE MODELS

In the years since 1990, there has been ample opportunity to observe the evolution of MRM as a tool for change directed toward safety improvement. Several papers based on the research program have been published on this topic (Taylor & Robertson, 1994; Taylor, 1995, 1996; Taylor & Patankar, 2000). Although there are myriad ways to bring about change in an organization there are at least four different models we have observed in the aircraft maintenance organizations we have studied. The models are "ideal types" and thus emphasize a pure form which is usually not found as such in the "real world." Such ideal types however help distinguish the major differences among the models observed – even if not in such purity. The models are also not mutually exclusive and can be combined for greater or lessor effectiveness. Thus, companies can provide illustration of more than one model and in fact the companies studied during 1998-1999 do so. The four ideal models are 1) individual vs. system change, 2) awareness vs. behavior change, 3) episodic vs. ongoing programs, and 4) AMTs only vs. all maintenance personnel.

1) "Individual MRM vs. system change for MRM"

This contrasting model emphasizes either changes within the individual mechanic or changes in the larger maintenance system. As such the larger "system" is comprised of individuals, groups of people together, their work processes, their technologies, their output, as well as management practices and policies. The individual approach changes people as individuals, and as a consequence their changes are expected to "add up" to larger changes. When the emphasis is on "system" instead of individual, the person is seen as part of a larger whole into which he/she is embedded and connected. Changing any part of a system will affect all the other parts and so the systemic change effort takes the individual and his/her context into account. The now completed MRM programs of companies A, E and F best illustrate the individual approach. The still ongoing programs of companies H and D take a somewhat larger and more systemic approach. Company H changes the structure and process of mechanics' decision making, while Company D affects the larger shift turnover process as well as individual mechanic's care and coping.

2) "MRM awareness vs. MRM behaviors"

The contrast here is between "talking the talk," and "walking the walk." Some MRM programs are designed to impact how people think and how they see things. Other programs are designed to change how people behave – especially how they interact with other people. The positive effect of skills training is documented in some early survey reports (Taylor & Robertson, 1995; Taylor, et al., 1997), but most MRM programs documented between 1996 and 1998 have been specifically designed for awareness-building (Taylor & Patankar, 2000). In the present case, Company E's MRM training program was oriented toward imparting ideas and concepts for safety – in other words "awareness" of ideas. Company H on the other hand focused exclusively on the use of a well-defined, joint decision making process (Patankar & Taylor, 1999). Companies D and F emphasized some passive awareness and some active response – such as assertiveness .(Company F), or better written communication (Company D).

3) "Episodic MRM vs. ongoing MRM"

The episodic model emphasizes a one-time event – such as the MRM training; that once completed is set aside and not referred to officially again. In training parlance episodic changes can be "spray and pray" (train everyone and hope it make some difference), or "blame and train" (the "traffic school" approach to correcting rule infractions). In either event, once the intervention is complete, the program is over. In past studies we have reported the apparent continued effects of episodic programs (Taylor & Patankar, 2000), but have also seen positive effects diminish within a few months of the end of a program (Taylor, 1994; 1998). Companies A and E illustrate the episodic approach. Despite the individual efforts of some of Company A's MRM facilitators to keep the ideas alive with the many mechanics who completed the course, there is no official or sanctioned process to follow-up Company A's MRM training. "Ongoing MRM" attention to the program continues with -- and for -- all participants long past their initial official activities. Ongoing or follow-up activities are officially recognized and supported by a company's senior managers and union officials – they are a part of the program, not a "new program" for safety. Recurrent MRM training would be an example of "ongoing MRM," but it is not the only case. Companies D and H exemplify different "ongoing MRM programs." In the case of Company D the training is publicized to happen at least twice (Taylor & Patankar, 2000), while Company H assures that changed behaviors are reinforced through a standing agenda item to address them at every daily shift change meeting (Patankar & Taylor, 1999).

4) "MRM for AMTs-only vs. MRM for all Tech Ops personnel"

A maintenance human factors and safety program exclusively for mechanics and inspectors is easier to design and deliver than its contrasting counterpart that is administered to a mixture of occupational and professional specialties and different hierarchical levels. Referring to Appendix "B", Company A's program is seen to be a prime example of the former type while Company D's program illustrates the latter approach. In past studies, the exclusive focus on mechanics alone was found to be a liability to further diffusion of MRM programs because of a management ignorance of and lack of support for this kind of training (Taylor, 1998).

An overview of these four models for the six companies is shown in the summary table below.

Summary Table				
Site	Individual vs. Systemic	Awareness vs. Behavior	Episodic vs. Ongoing	AMT only vs. All Personnel
Company A	Individual	Awareness	Episodic	AMTs only
Company D	Combination	Combination	Ongoing	All personnel
Company E	Individual	Awareness	Episodic	AMTs only
Company F	Individual	Combination	Episodic	AMTs only
Company G	Systemic	Combination	Not determined	All personnel
Company H	Systemic	Behavior	Ongoing	All personnel

METHODS & TOOLS

SAMPLES AND UNITS OF ANALYSIS

Individual respondents as the focus of analyses. Table 1 provides the basic characteristics of the six sites studied during 1998-1999 and are reported here.

Companies "A" and "D" have been analyzed in detail in this report. The former, Company A, has completed a one-time MRM training program for most of its mechanics and leads. Company D is still in process of a two-phase MRM program involving all of its maintenance employees.

Table 1. Site Samples

Site	Industry Type	MRM program	Employee focus	Training Length	Status	Present Sample size
Co. A	**Airline**	**One-time Training**	**AMTs**	**2 Day**	**Completed**	**6,265**
Co. D	**Airline**	**Phased Training**	**All maintenance employees**	**2 Day**	**In process**	**2,600**
Co. E	**Airline**	**One-time Training**	**Component shop mechanics**	**4 hour**	**Completed**	**1,220**

Co. F	**Manufacturer's contract maintenance**	**One-time Training**	**Line maintenance AMTs & Foremen**	**2 day**	**Completed**	**135**
Co. G.	**Airline**	**Training, plus process intervention**	**All Maintenance employees**	**1 day**	**Beginning 2000**	**124**
Co. H	**Corporate aviation department**	**Process & structure intervention**	**Line maintenance AMTs & Foremen**	**N/A**	**In process**	**7**

The six sample programs vary in the composition of their respondents. Appendix B contains these demographic statistics for Companies A and D, as well as from all site samples reported here. The Company A sample contains a high proportion of mechanics, and equal proportions of line and base workers. Its sample has a lower average age, and employees have fewer years in their jobs. Company D's sample includes a higher proportion of inspectors and foremen than Company A. Company D's sample also includes a higher ratio of shop employees, and a higher proportion on day shift than does A. Two of the other samples (Companies D and G) consist of representative proportions of maintenance management, foremen, leads, support personnel and AMTs. One sample (Company E) contains only component shop personnel. Companies A, F, and H are composed mainly of AMTs.

In exploring the effects of the training on all individuals, the data from all training participants totaled for each MRM program sample site will be used. Remaining analyses will examine the attitudes of respondents in combination with the others in their same work units.

Maintenance work units as the focus of analyses. The maintenance performance data (classified into categories of "Occupational Safety," "Ground Damage," and "Paperwork Errors") are measured by work units, not by individual respondents. The correlation analyses described in this report illustrate the effect of changes in respondent attitudes associated with the maintenance performance of their work-units. For managers these are the units they lead. For staff professionals, maintenance foremen, leads and AMTs, these units are the stations and locations to which they belong. In order to accomplish the examination of attitudes correlated with performance, the individual respondent's attitude data were combined into averages for their appropriate work units.

Field visits were made to two sites in Company A and one site in Company D during 1998-1999. These visits involved several investigators observing and informally interviewing maintenance employees on all shifts over several days. They occurred within a year following the conclusion of training. The data collected were intended to corroborate and validate the data collected in the post-training survey and the subsequent follow-up surveys. They were also intended to provide additional information about team or workgroup structure and process in the months following training.

MEASURES

THE ATTITUDE MEASURE: The "Maintenance Resource Management Technical Operations Questionnaire" (MRM/TOQ)

Since 1991 the attitudes, opinion and self-reported intentions and behaviors associated with MRM interventions have been measured and analyzed. The raw data for that analysis have been provided from time-series surveys conducted by the airlines themselves with the assistance of the investigators. These surveys use a number of standardized questionnaire items agreed-to in advance of the training. These questions are used alone or combined into Likert-type scales [*cf.*, Festinger & Katz, 1953, pp. 530-531; Selltiz, et. al. 1976, pp. 418-421] to assess the degree of improvement achieved by the airlines' various MRM programs. Together these items are called the Maintenance Resource Management/ Technical Operations Questionnaire (MRM/TOQ). Ideally each airline partner in this research program would survey its MRM participants before an intervention begins for a given sample population ("baseline"). It would then measure again immediately prior to a planned MRM intervention ("pre-"training), immediately following the intervention ("post-"training), as well as time periods following the intervention ("two-month," "six-month," and "12-month" follow-up surveys). In reality each partner company in this 1998-1999 sample set differs from that ideal measurement model in some way, but the resulting data are adequate for assessment in every instance. These time-series data points have been previously used only to compare a single airline's results over time. They have now been recombined to yield normative profiles

Thus the common evaluation methodology begins with a survey -- the MRM/TOQ. Several versions of the MRM/TOQ will be discussed below. Each is designed to measure attitudes, opinions and other information during a different time period in each MRM program. All versions of the MRM/TOQ include demographic or background questions, and closed-ended multiple choice attitude and perception questions. Some standard versions of the MRM/TOQ also include open-ended questions to be answered in the respondent's own words.

Five Uses Of The MRM/ TOQ

There are four versions of the MRM/TOQ which were used in five different ways in the present project. The four versions are included in Appendix A.

1. "Baseline MRM/TOQ Questionnaire." This baseline survey measures attitudes and opinions before an MRM intervention is begun. The baseline questionnaire is typically mailed by company to a small sample (typically 10%) randomly drawn from all or a subset of their maintenance personnel. Recipients are instructed to complete the surveys and return them in envelopes provided either to the MRM administrator or to Santa Clara University (SCU).

The form of the baseline questionnaire is identical to the "pre-training" questionnaire described below, and like all versions of the MRM/TOQ it includes eight employee background items. Because all the questionnaires used in this present study are based on the MRM/TOQ core questions, the results can be compared across time for one company as well as between participating companies. Earlier experience with the baseline survey shows the return ranges from 50% (Taylor, Bettencourt & Robertson, 1993; Taylor, 1998) to 76% (Choi, 1995). In the present analysis, Company E's baseline survey return rate was 38%, but some in the sample received baseline surveys in the mail after they had already attended their MRM training, and they thus ignored the baseline form when they received it. For Company G, where efforts were made to maximize returns, the final rate was nearly75%[2].

2. "Pre-training MRM/TOQ Questionnaire." Pre-training questionnaires are completed immediately before MRM training sessions by people currently attending the session. The pre-training surveys are sent to SCU. These pre-training attitudes are compared with "post-training" attitudes immediately after the training, as well as with attitudes measured months later. Instructors are expected to introduce the survey as "voluntary." Normal return rate is very high (90-95%), but varies from a total census – and is usually due to some participants coming late to their session and thus missing the survey administration.

3. "Post-training Questionnaire." The post-training survey is completed by participants at each workshop's conclusion. Data from the post-training MRM/TOQ are sent to SCU at the same time as the pre-training survey. Typically all surveys, including the post-training questionnaire, include the same 26 attitude and opinion items and the eight background questions as the baseline and pre-training questionnaires. In addition, the post-training questionnaire (and all follow-up surveys described below) contains three more multiple response items. These are used to measure enthusiasm for the training. The post-training survey also includes several open ended (or "write-in") questions some of which ask respondents to evaluate the training content; and others of which measure intention to change and self-reports of changes made as a consequence of the training. These latter behavioral descriptors will be further described below.

Post-training surveys are distributed and completed in class and the normal return rate is also between 90-95% (Taylor, 1998). This deviation from 100% can come from several sources. The training facilitators emphasize that the survey is strictly voluntary and confidential and, because it is voluntary, some fraction of all trainees will choose not to complete it. Such shortfall can also result from participants who need to leave the training session early. A larger discrepancy between the numbers of pre and post-training surveys returned for analysis is noted for Company E in the present study. This was the result of some of Company E's facilitators who either did not distribute the MRM/TOQ forms, or who failed to remind participants to complete them.

4. "Follow-up MRM/TOQ Questionnaires". A questionnaire form similar to the post-training instrument is sent to participants in the months following their initial training. Differences from the post-training include wording questions in the past tense instead of the present or future (i.e., "this training has been useful to others," rather than "this training will be useful to others"). Another difference between the follow-up and post-training surveys is that the former includes a write-in question asking respondents to describe how they will use the training on the job. Although the time period can vary, these follow-up surveys are designed to be collected 2, 6, and 12 months afterwards. In the present study these follow-up surveys were all identical in form, they measured the respondents' thoughts, assessments, and attitudes over increasingly lengthy periods from the training.

An example of the MRM/TOQ 2/6/12-month follow-up questionnaire is included in Appendix A.

5. Use Of The "follow-up" MRM/TOQ As A Stand-alone Survey Instrument. Company E had begun its MRM training before deciding to use the MRM/TOQ surveys. In this case the AMTs' MRM program had been completed in several line stations and a heavy maintenance hangar. In that period some 2,500 AMT participants had attended the training. The Company E Human Factors manager decided not only to begin using the pre-training and post-training in January 1998, but also agreed to send out MRM/TOQ questionnaires to previous participants. It was decided that these past participants would be asked to complete a 12-month questionnaire. The questionnaires used were identical to the "follow-up" instrument in Appendix A. Those questionnaires received back act as a stand-alone, post-hoc survey of MRM attitudes, opinions, self-described behaviors, and future intentions to use the training. To distinguish these stand-alone 12-month surveys from the "follow-up" surveys -- which are mailed to participants who had earlier been asked to complete pre-training and post-training surveys in the training session -- they will be referred to as the "12-month Survey." This convention is consistent with that used in earlier MRM research (*cf.*, Taylor, 1998).

THE MRM/TOQ ATTITUDE AND OPINION SCALES

Many of the closed-ended, individual questions in the MRM/TOQ are combined into scales or indices. These scales are measures of attitudes and opinions about "communication & coordination," "shared decision making," "assertiveness," "stress management, "goal sharing" and "safety practice". All six of these scales have been extensively tested during 1998-1999 and the results of that work are reported elsewhere (*cf.*, Taylor, 2000b).

The MRM/TOQ has developed and evolved over the last decade. Most core items in the MRM/TOQ are derived from a 1990 version of the Cockpit Management Attitudes Questionnaire (CMAQ) -- a training, evaluation and research tool developed by investigators at the University of Texas (Helmreich, Foushee, Benson, & Russini, 1986; Taggart, 1990). The CMAQ questionnaire contained a number of items measuring attitudes that are either conceptually or empirically related to communication and teamwork training provided to flight and maintenance personnel.

Methodology for Combining Survey Items Into Scales

Four previous studies have used Factor Analysis to explore and confirm a consistent internal structure for the core questionnaire items of the CMAQ and the MRM/TOQ (Gregorich, Helmreich, & Wilhelm, 1990; Sherman, 1992; Choi, 1995; Taylor, in press).2000b). In those four studies, using samples of flight crews, air traffic controllers, AMTs, and maintenance managers respectively, the authors have shown that the relationships among the 18 core items clustered into four "factors," or constellations of attitudes. The four factors are shown in Table 2.

Table 2. Common Attitude Factors Across Four Studies
1) Value Shared Decision Making,
2) Value of communication and coordination,
3) Value of stress recognition & management,

4) Value of assertiveness (or willingness to voice disagreement).

All four factor analytic studies combine the basic items into composite index scales to obtain more stable indicators of underlying concepts. Such indices permit a more detailed assessment of the separate but related attitudes than a single total score for the entire questionnaire, but they also provide more accurate and reliable results than are available from each of the individual questionnaire items alone.

Measuring both attitudes and opinions. For the purposes of this report, attitudes will be defined as "values or feelings respondents hold toward certain abstract concepts" (e.g., employee participation, or coordination) "and behaviors" (e.g., stress management or assertiveness). On the other hand, opinion refers to "personal judgement," or "evaluation of some person, act, or thing."

Since 1991 the MRM/TOQ has included six items measuring the degree to which a respondent's work group practices goal sharing and setting. Taylor (2000b) included these items in his confirmatory factor analysis and reports they form a fifth factor, "Goal setting and sharing." In 1998 the MRM/TOQ was expanded to include five questions measuring knowledge and opinion of safety practices in the maintenance department. Those five questions were adapted from items in later versions of the CMAQ developed and tested by researchers at the University of Texas. In the MRM/TOQ the five are combined into a single opinion scale called "Safety Practice."

The composition of the four attitude and two opinion scales in the MRM/TOQ are shown in Table 3.

Table 3. Scales Used to Measure Human Outcomes of MRM Training

MRM/TOQ SCALE	CONSTITUENT ITEMS (See Appendix A)
Attitude Scales	
Value Shared Decision Making	6, 8, 11, 13 (reflected)
Value Communication & Coordination	12, 14, 15, 16
Value Stress Management	9, 17, 18
Value Assertiveness	1, 2 (reflected)
Opinion Scales	
Opinions of Goal Setting & Sharing	20-26
Opinions of Company's Safety Practice	3, 4, 7, 10, 19

The individual item numbers in Table 3 correspond to the numbering of the sample questionnaire found in Appendix A. A "reflected" scale means that the scoring of individual responses for constituent items comprising that scale are reversed (i.e., 1=5, 2=4, 4=2, 5=1) before the scale averages are calculated

As noted above, each MRM/TOQ survey also includes several questions to collect demographic data (i.e., information about respondent background). These questions include years with the company, years in present job, age, gender, education, job title, department, and shift.

Finally, all MRM/TOQ surveys administered after the training include three more multiple response items that are used to measure enthusiasm and personal expectations for the training (*cf.*, Appendix A, pp. 4 and 6, items III-1 through III-3).

OPEN-ENDED QUESTIONS IN THE MRM/TOQ

Each version of the MRM/TOQ after the pre-training survey include open ended (or "write-in") questions soliciting reactions to the training and/or improvements to the training (Appendix A, pp. 4 and 6, items III-4 through III-6.

What aspects of the training were particularly good?

What do you think could be done to improve the training?

Another open-ended item collects respondents' intentions to change.

How will you use this training on your job?

And in the follow-up (2,6, and 12-month) surveys, an open-ended question asks for respondents' self-reports of changes made as a consequence of the training (Appendix A, p.6, item III-4).

What changes have you made as a result of attending the MRM training?

INTERVIEWS AND FIELD OBSERVATION TO SUPPLEMENT AND SUPPORT MRM/TOQ

An additional measurement took the form of field visits made during 1998-1999 to two sites in Company A and one site in Company D. These visits involved several investigators observing and informally interviewing maintenance employees on all shifts over several days. They occurred within a year following the conclusion of training at the given site. The data collected were intended to corroborate and validate the data collected in the post-training survey and subsequent follow-up surveys. They were also intended to provide additional information about team or workgroup structure and process in the months following training.

Comparing Answers from Open-ended Survey Questions and Interview Results

Assessing self-reports of changes made. In the question of efficiency in data collection, self-reports of changes actually made are direct and easier to obtain than actual observations. When those self-reports are written in response to the last survey question described above the efficiency improves further. Such efficiency must come at some price in the goodness (i.e., reliability and validity) of the results obtained. Thus the degree of goodness of self-reported statements of changes made in response to MRM programs should be assessed.

Assessment in this case takes the form of comparing the self-reported, written answers with subsequent interviews and observation of changes attributed to MRM training interventions. Two large line stations and one large, heavy maintenance station were part of this study. All three sites had completed MRM training in a relatively brief period -- ranging from 3 months for a 450 mechanic line station, to 10 months for a 1,000 employee maintenance base. All stations had collected MRM/TOQ survey data in the months following the training -- and these data included written answers about respondents' changes made as a result of the training. These responses are collapsed into several categories -- "greater awareness," "better communication," "greater care at work," and "better stress management." They are further analyzed and described in Table 4, p. 15.

Field visits were structured such that observer/interviewers could remain on each of the three shifts at least two days during a three or four day period. In each site, the sample of interview respondents was obtained via introductions from the shift foremen, leads, and union representatives. The interviews were usually held at the individual's work place or in the employee break areas. In this setting, investigators had ample time to observe normal operations as well as to discuss MRM with employees. These interviews allowed respondents to discuss their recollections of their MRM training and to report their observations and impressions of changes made as a result of the training.

These field visits took place in the two Company A sites about one year after their training had been completed. In the first of those visited (a line station of 750 employees) 40 of the prior participants (about 5% of the total) were interviewed. In the second site (the maintenance base of over 1,000) 220 prior training participants (over 20%) were interviewed. In the third site (a line station in Company D, with some 450 mechanics) field visits occurred about three months following each of two phases of the program. Following the first phase, 140 (about 30%) of all training participants were interviewed and during the second field visit after their completion of the second phase of MRM training about 90 (20%) were interviewed.

For the purposes of this analysis the data between the written answers to the questionnaires and the interview results were reduced from some two dozen coded responses and summarized into five general categories of MRM outcomes. The five included "greater awareness," "more communication," "greater care in task accomplishment," "better stress management," and "there was no effect." Four separate measures of MRM/TOQ had been previously administered among the three field sites. Each site had at least one survey questionnaire in which employees were asked to write down what they had done on the job as a consequence of MRM training. One site had two waves of survey data – the first two months following the training and another six months after. Each site also had two interview measures. Line station D-2 had two visits about six months apart. The other two sites had only one visit each, but these were structured such that individuals could report and illustrate their own changes, as well as those observed in others. Table 4 shows the measures and sites along the left axis, and the proportions of responses for each of the five categories of MRM outcomes populate the body of the table.

The results in Table 4 were tested for agreement among the 10 measures (4 surveys and 6 interview measures) over the 5 related MRM outcomes. This test (Kendall's Concordance, or *W*) determines the overall relationship among the ratings from the 10 measures. The Kendall *W* ranges between "0" (no agreement) and "1" (complete agreement). The statistic obtained for the results in Table X (*W* =.717; N=10, df=4) is large and significant above the .0001 level of confidence. This result shows that there is a high degree of agreement or concordance among the 10 measures, and thus between the questionnaire answers and the interview outcomes as well.

Those field visits helped confirm that the changed behaviors people reported were actually practiced. Observations of team meetings, shift turnovers, personal awareness and interpersonal safety reminders (this last target including posters, but not limited to them) were noted and trends observed were compared with the survey and interview results.

Table 4. Comparing Answers to Open-Ended Survey and Interview Questions
Percentage of Five MRM Outcomes For 10 Measures over 3 Sites*

	1. Greater awareness	2. Better communication	3. Greater care in work	4. Better stress mgt.	5. No impact	(other)
Site A-1**, 2Mo Follow-up Survey	39	14	4	6	30	7
Site A-1, 6Mo Follow-up Survey	30	12	12	4	30	14
Site A-1, Interview: Self-Change	33	0	13	9	45	0
Site A-1, Interview: Others' Chg	5	33	5	0	42	14
Site D-2**, 2Mo Follow-up Survey	52	8	21	0	13	5
Site D-2, 1998 Interview	15	13	4	0	67	2
Site D-2, 1999 Interview	25	5	2	0	68	0
Site A-3, 6Mo Follow-up Survey	36	24	3	4	26	7
Site A-3, Interview: Self-Change	22	27	3	3	33	12
Site A-3, Interview: Others' Chg	14	27	7	2	55	4

*** *W* =.717, N=10, df=4, p<.000**
**** Sites Ai and D2 are both Line Maintenance Stations**

MEASURES OF MAINTENANCE SAFETY PERFORMANCE

Company A has provided monthly statistics for Occupational Injury and Aircraft Damage for the years of 1995 through 1999. The total of all 1,250 AMT respondents in the 6-month Company A sample includes the members of over 60 organizational units drawn from many parts of maintenance (base maintenance, line maintenance, shops, quality, and stores). The total numbers of units in each analysis vary depending on the specific performance indicator because not all the work units are measured on the same performance, nor are the same units always reported every month. Table 5 shows the range of numbers of work units available for each of the measures available monthly over the 1995-1999 period.

Table 5. Sample of Work Units Available for Safety Performance Measures Company A Study	
Safety Performance Measure	**Number of Work Units**
Lost Time Injuries (LTI)	· Line Maintenance Stations: 30-31 · Base Maintenance Units: 28-33 · All Sites in Sample: 58-64
Ground Damage-Aircraft (GD)	· Line Maintenance Stations: 30 · Base Maintenance Units: 27-29 · All Sites in Sample: 57-61

Company A Lost Time Injuries (LTI)

The injury rates are expressed in terms of the number of injury incidents which result in days lost to treatment and recovery (termed Lost Time Injuries, or "LTI"). The AMT MRM training undertaken by Company A is expressly intended to reduce LTI, and time is spent at the conclusion of each two-day session in reviewing an injury case and discussing ways to avoid such cases in future. LTI data are available by maintenance cost center by month. There are some 90 Company A cost centers reporting LTI from which personnel attended the AMT MRM training through the period studied here. Actual personnel count per month for each line station, or heavy maintenance work unit were also provided by Company A, and those data were used to control the error statistics by size (i.e., head count) per station. Those corrected LTI data for all months of 1995 through 1999 are totaled by month and plotted graphically (see Figures 32 and 33).

Aircraft Ground Damage (GD) statistics for Company A

Monthly counts of maintenance-related aircraft damage are also available for 1995 through 1999. The number of cost centers reporting damage incidents are roughly the same as those for LTI. Again, the actual personnel count per month for each line station or heavy maintenance work unit were used to control the error statistics by unit size (i.e., head count). Those corrected aircraft damage data for all months of 1995 through 1999 are totaled by month and plotted graphically (see Figures 31 and 33A).

Company D Paperwork and aircraft damage statistics

Logbook errors as well as total paperwork errors per line station per month have been made available from January 1996 through 1999 in Company D. Actual flight departures per line station per month were also provided by Company D and those data were used to control the error statistics by size (i.e., amount of flight activity) per station. Those corrected paperwork error data for all months of 1996 through 1999 are totaled by month and plotted graphically (see Figure 36). The resulting chart provides for an examination of error trends over 48 months before, during and after Company D MRM training began in January 1998.

Company D has also provided monthly statistics for Aircraft Damage for the years 1997 through 1999 (see Figures 34 and 35). The total of all 2,600 respondents in the phase 1 post-test Company D sample includes the members of three maintenance stations.

MRM ATTITUDE AND OPINION BENCHMARK PROFILE

Profiles and profile analysis are highly useful ways to study and audit group scores (Kerlinger, 1979, pp. 272-274). A profile is a set of scores from a set of measures. Group profiles used here are average scores obtained from a group of scores. The profiles created in the present case are converted to standard scores and are referred to as “benchmark” measures. The distance between the individual means for a specific group, as adjusted for the variability of the population through the use of standard scores, and the standard score of the normative profile (“benchmark”) for that larger population are expressed as percentile ranks.

Percentile ranks are easy to compute, widely used and understood, suitable as a means of displaying information, and make it possible to compare the scores of groups that are unequal in size. When the distribution of underlying scores approximates that of a normal distribution, derivative percentile ranks will give more weight to scores on the extremes of their distribution and less weight to the scores which pile up in the middle (Brown, 1991). For our purposes, this is an advantage in highlighting those MRM unit scores that are substantially above or below the benchmark population.

The MRM Benchmark profile panels are created by calculating standard scores ("Z") for each of the six MRM surveys –Base, Pre-, Post-training, 2-, 6-, and 12-month. Next, percentile scores for any maintenance work unit or site are calculated for the appropriate survey panel using that panel's "Z" scores. These profile panels, derived from standard scores ("Z"), can be used to compare the percentile rank of MRM attitudes and opinions in any given company at any stage in its MRM program with like measures from the large database of like employees during a similar period in their MRM involvement. This panel of Benchmark comparison profiles for attitudes and opinions forms one tool to help audit the effectiveness of a maintenance human factors program.

The Normative Profiles

The appropriate responses from all of the companies measured using the MRM/TOQ were combined by each period in the data collection time series. Those raw data were then transformed into "standard scores" which provide and prescribe the norm (or standard for comparison) for each time period. The data in each time period thus provides its own profile:

- a "baseline" profile (a comparison standard to use before an intervention begins for a given sample population);
- a "pretraining" profile (a standard for comparison immediately prior to a planned MRM intervention);
- a "post-training" profile (for comparison of MRM/TOQ results immediately following intervention);
- two-month and six month profiles.
- The12-month "stand alone" profiles (MRM intervention comparison profiles for responses gathered up to a year afterwards).

In the eight "Survey Results" sections that follow, comparisons are made with population profiles only for a selected subset of the raw data – six attitude and opinion scores from the Likert scales derived from some 26 individual measures, plus three "enthusiasm items." These six scales and three separate items are presented as percentile scores. Each section contains the profile information for a time period: Baseline, Pre-training, Post-training, 2- & 6-Month, and 12-Month data sets.

RESULTS

MRM/TOQ SURVEY RESULTS

The order of presentation for survey results in this section is as follows. The mean scores for the total data set (the "Benchmark") are displayed first, followed by the separate mean scores for the companies "A," "D," "E," "F," "G," and "H," as available. Following these mean scores, the percentile ranks for each of the companies are shown, comparing each company with all of the others included in the Benchmark database.

ENTHUSIASM FOR MRM

Three questions examine and compare participants' general reactions and enthusiasm to their MRM program. One of these questions measures the degree the "training increases safety and teamwork" (*cf.*, Appendix A, p.4, Q III-1). The second question asks about the "usefulness of the training for others" (Appendix A, p.4, Q III-2). The third question measures the "training's effects on my on-the-job behavior" (Appendix A, p.4, Q III-3). In immediate "post-training" questionnaires these three questions are worded to measure expectations (e.g., "this training will be useful to others"). In the questionnaires used 2, 6, or 12 months later (*cf.*, Appendix A, p.6, Q III 1-3) the three questions are worded to measure actual experience (e.g., "this training has been useful to others").

That third question ("training's effects on my on-the-job behavior") uses a four-point, forced-choice scale, instead of the more usual 5-point scale. This current and past use of the "forced choice" method with this question has been to avoid neutral or ambiguous answers. In continuing this practice we prevent respondents from straddling the fence and avoid committing to action or inaction. The lower mean scores for this question are the result of its "shorter" (i.e., 4-point) scale.

THE BENCHMARK PROFILE: COMPARISON OF ALL PARTICIPANTS IN THE DATA SET

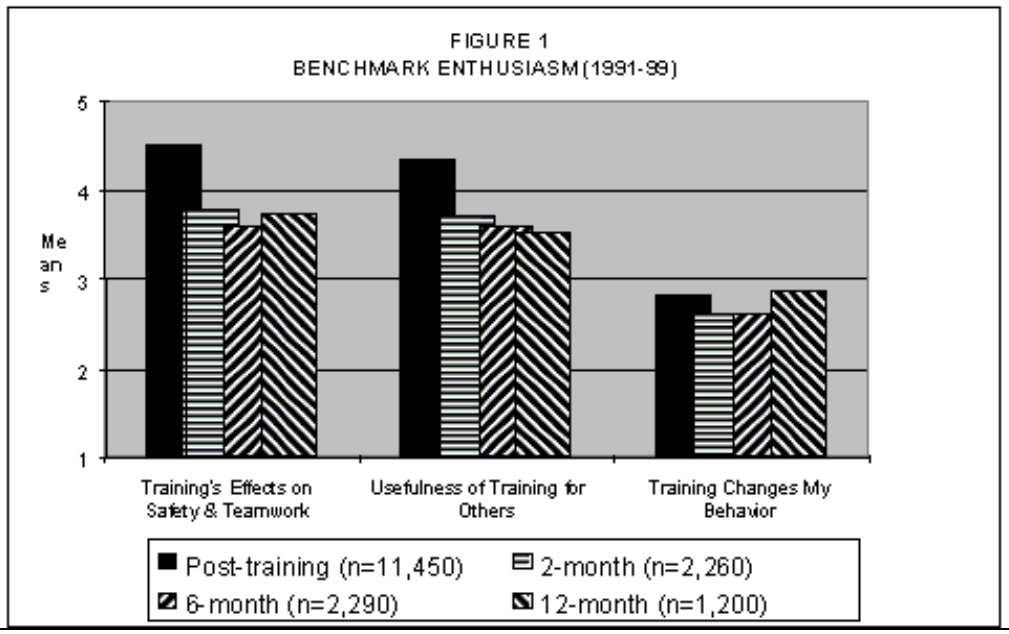

Figure 1

Figure 1 shows the Benchmark's mean scores results for each of the four survey waves that followed MRM training. Clear from Figure 1 is that respondents dating back to 1991 are strongly optimistic about the promise of the MRM program immediately after they have experienced it (i.e., post-training). Their expectations for the training are very high for its general outcomes and for its effect on others. Respondents are less enthusiastic about the training's effect on themselves.

Figure 1 also shows that Benchmark results for surveys taken some months after the training show decidedly lower scores for the general effects of the training than that immediately following the training. This confirms earlier conclusions (Taylor & Robertson, 1995) that positive energy for MRM can dampen with time, and that this may be discouragement with a lack of program follow-through (Taylor, 1998). In assessing the personal impact of the training ("Training's effects on my behavior"), Benchmark results for the 12 month surveys are more positive than their 2 month and 6 month counterparts and these differences are statistically significant ($p<.05$). This may be the result of a biased sample – those who answer and return the 12 month survey may be more positively disposed to the training and its concepts than those who do not. Alternatively the very diminished scores two months and six months afterward (especially for the first two questions) probably reflect lapsed activity in program implementation; or the effect of inflated expectations for MRM program's effects on others that are subsequently realistically revised by the 12 month respondents.

Figure 1 shows very high expectations for the MRM program in general, and for its effects on others immediately after the training. The reality of the situation is assessed somewhat lower in the months following that initial enthusiasm.

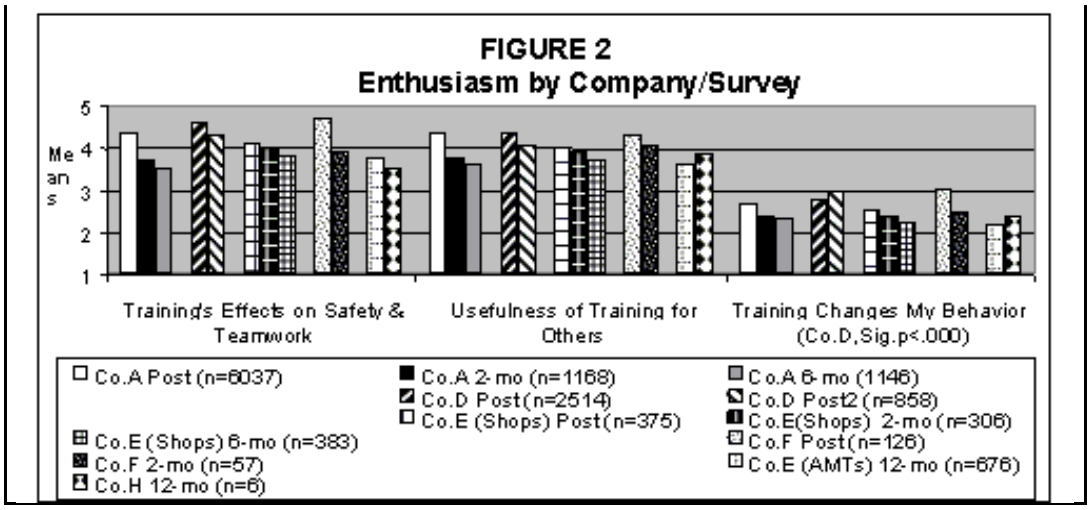

Figure 2

Enthusiasm Scores for the Five Companies

Mean scores for the three enthusiasm items for the five companies employing MRM training are shown in Figure 2. Companies A, D and F all show very high mean scores immediately following their training. All four companies show a decrease in mean scores over time for the first two questions. For the third question, "Training changes my behavior," only Company D shows an increase over the six months between its phase I and phase II training -- and this difference is statistically significant ($p<.000$).

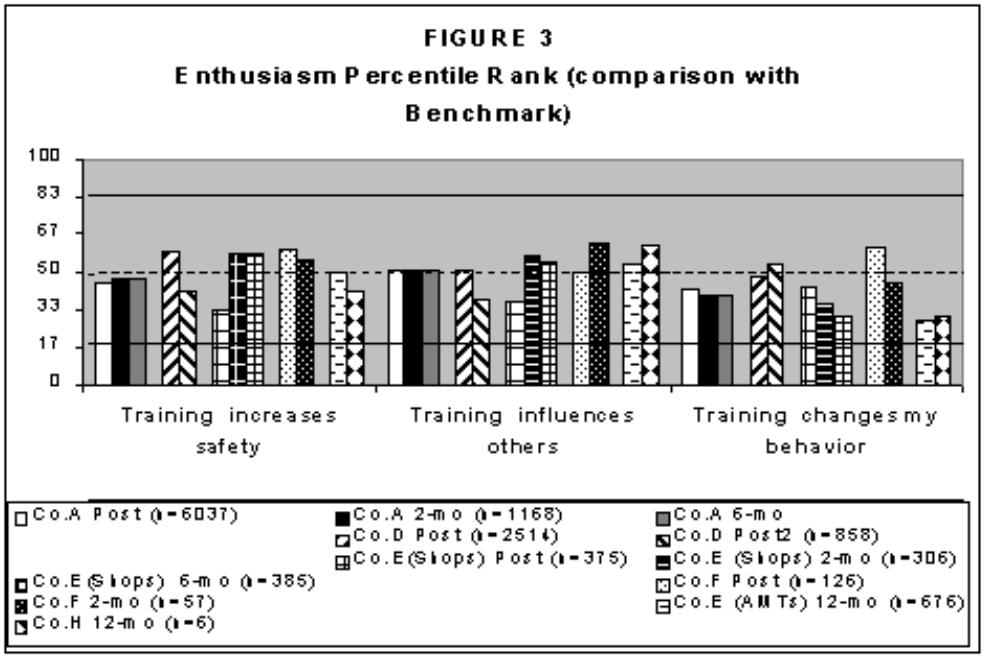

Figure 3

Figure 3 shows the percentile ranks comparing the enthusiasm scores for each company with the Benchmark. Company A is commensurate with the Benchmark norm for the first two questions, and below the norm for the third question ("Training changes my behavior"). Company E's 2-month and 6-month enthusiasm are both slightly above the appropriate Benchmark profiles for the first two items. Company E's post-training enthusiasm for personal change (the third question) is quite low. Company F's post-training enthusiasm is at or above the Benchmark norm for all three questions, while its 2-month survey shows a decrease in enthusiasm for personal change.

Company D's initial (Phase 1) post-training enthusiasm too, ranks above the Benchmark norm. Company D's phase 2 post-training enthusiasm cannot be ranked or compared with the post-training Benchmark because it is a unique measure of people who have already had several months to think about and use the ideas presented in the phase 1 MRM training.

MRM ATTITUDE AND OPINIONS PROFILES

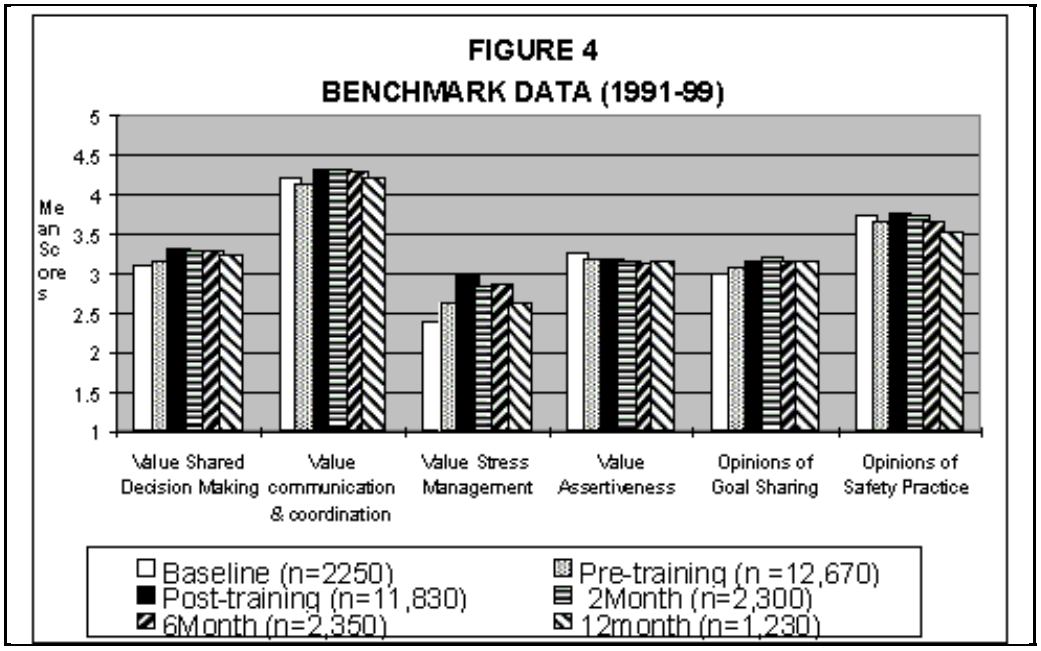

Figure 4

MRM Attitudes and Opinions: Benchmark Profile

Figure 4 displays the mean scores for the four attitude scales and the two opinion scales for all respondents (the "Benchmark" dataset) over all surveys. It captures the accumulated evidence collected over the past decade. This evidence collected from over thirty thousand respondents reveals a very regular profile of attitudes and opinions in maintenance operations.

These results show the Baseline survey's results are at least as high as those of the Pre-training survey are. The only exception is the Stress Management attitude scale, in which the baseline mean is lower than the pre-training mean.

The stress management scale also differs from the other scale results by showing the largest increase between pre- and post-training surveys. Stress management is also different in a nearly symmetrical diminution for the 2, 6, and 12 month surveys ($F=265$, $p<.000$).

Both the "Shared Decision Making" scale, and the "Value Communication & Coordination" scales show the "training effect" -- an increase between pre- and post-training surveys, followed by more stable mean scores for the 2, 6, and 12 month surveys ($62>F>73$, $p<.000$).

The "Assertiveness" scale shows little difference among the five surveys. The two opinion scales reveal only slight variation among the surveys.

Communication & Coordination is highest. Attitudes about coordination and meetings are clearly highest (within .75 points of the top of the five point scale). Employees in airline maintenance uniformly value meetings, communication and coordination. MRM training uniformly increases these scores a small but significant amount.

Opinions of company safety practice is next highest. Employees clearly rate safety management highly. There is a slight boost to these scores immediately after training, but then a slight and slow decline. It's as if this climate doesn't improve in line with respondent expectations.

Assertiveness is third highest. The value of being assertive and speaking-up is very stable over time at 3.10 - 3.25 on the five point scale. For this overall Benchmark population it is not much changed by the training.

The value of shared decision making, and opinions about goal setting & sharing are tied. These two profile scores are closely behind assertiveness. Unlike assertiveness, these two scales do show slight improvement following training and remain stable thereafter.

Value of stress management is the lowest score on the profile. All stress management scores for the Benchmark are below three on the five point scale. This scale also shows the greatest variation over time. After a marked increase after training, attitudes toward stress management fall to the pre-training level.

Attitude and Opinion Means and Percentile Profiles by Company

In Figures 5 through 15 to follow, each survey (Baseline, Pre-, Post-, etc.) are presented separately for the five companies. The first figure presentation for each survey contains the attitude and opinion mean scores for all five companies. The second figure for each survey presents the percentile comparison for the same data.

Baseline Survey Results

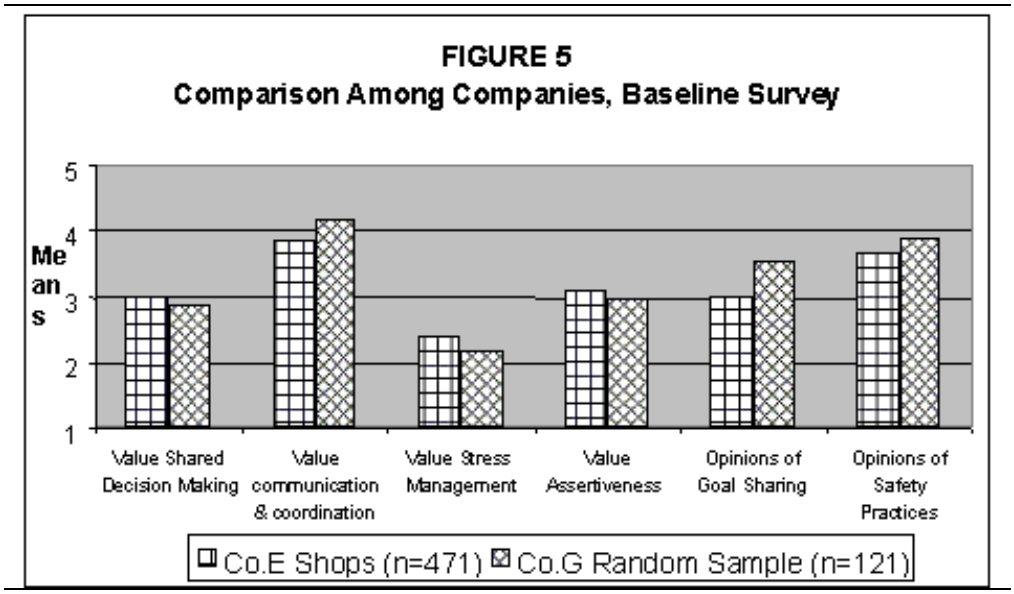

Figure 5

Figure 5 shows the baseline mean scores for Companies E and G. Company E's population is Component Shop Personnel only, while Company G sampled from their total maintenance workforce population. Figure 5 shows Company G has clearly higher attitudes toward "Communication & coordination," and higher opinion of their "Goal Setting & Sharing." The different populations from which these two samples were drawn may account for these differences. However, differences in the organizational cultures of the two companies may also explain the pattern.

Figure 6 displays the percentile scores that compare the two company's baseline results with the Benchmark. Results at or around the 50th percentile indicate a company's similarity to the Benchmark dataset. Only two percentile ranks are noteworthy deviations from the norm -- they are the 30th percentile score for Company E's Communication & coordination baseline attitudes, and Company G's 72nd percentile rank for Goal Setting & Sharing.

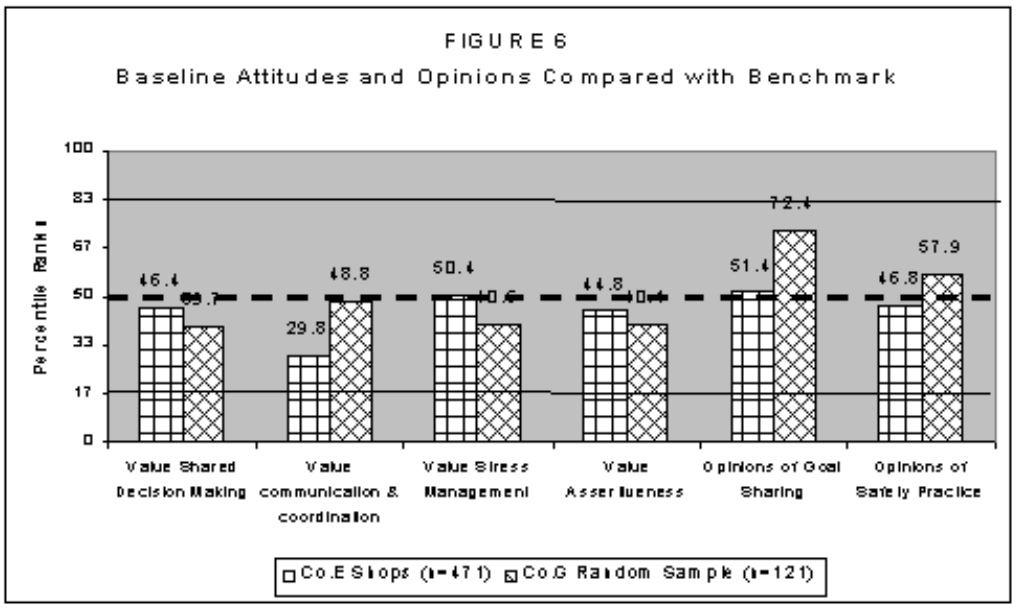

Figure 6

Company E's low baseline percentile rank suggests that these employees -- for whom the MRM training is intended, but who have not yet been informed thereof -- do not much value meetings as ways to coordinate work. The following sections include Company E's results following the MRM training and this low attitude towards meetings continues.

Company G's MRM program had not begun by the end of 1999 so a similar longitudinal comparison is not yet available. However Company G's rather high percentile rank for goal sharing is consistent with that Company's organizational culture -- which obviously carries over to its maintenance personnel.

Pre- and Post-training Survey Results

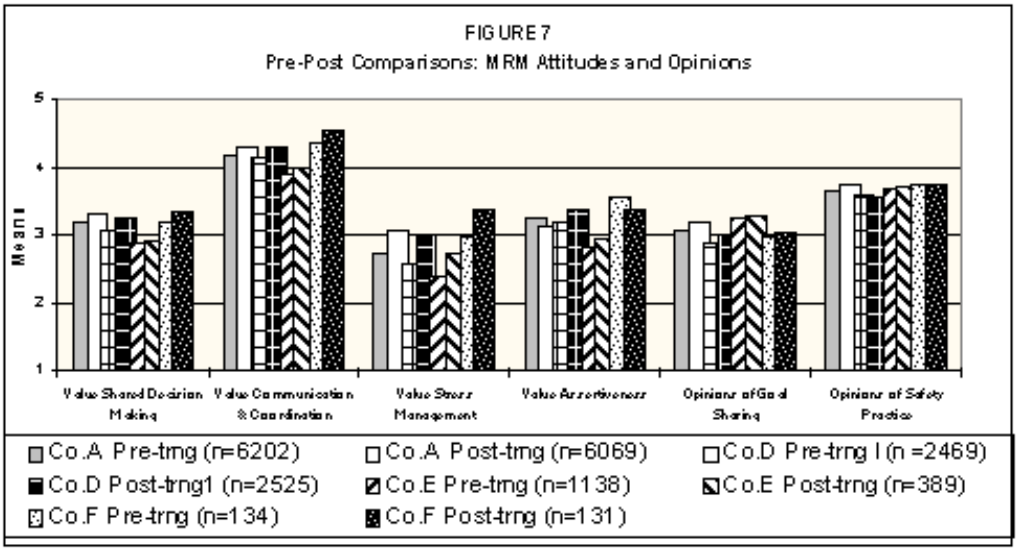

Figure 7

Figure 7 shows the mean scores for each company (A, D, E, and F). These results clearly demonstrate the first part of the "training effect" – the increase in the first three attitude scales between pre- and post-training mean scores – for each company. The assertiveness attitude scale shows mixed results depending on the company – two of these companies' MRM programs (Companies D and F) focused on or emphasized assertive communications. The two opinion scales show few pre-post differences and the safety practice scale shows remarkable similarity among companies.

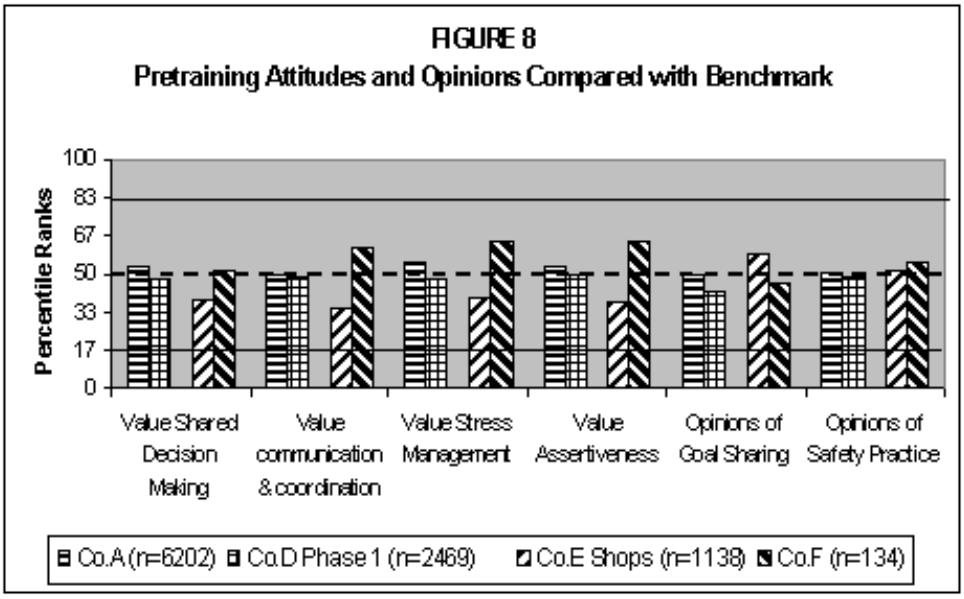

Figure 8

The pre-training percentile ranks for the four companies are shown in Figure 8. Company E now ranks somewhat lower than the norm for all the attitude scales, and slightly above for the goal sharing opinion scale. Given the tense labor relations climate at Company E at the time of the training, these low attitudes and the sense of high common goals among participants immediately prior to the training suggests a defensiveness or "circling the wagons" by participants in face of a company training program.

Company F, on the other hand, is well above the pre-training Benchmark norm for three of the four attitude scales. This above average value for communication, stress management and assertiveness should make the Company F participants especially sensitive to the message of the MRM training to follow.

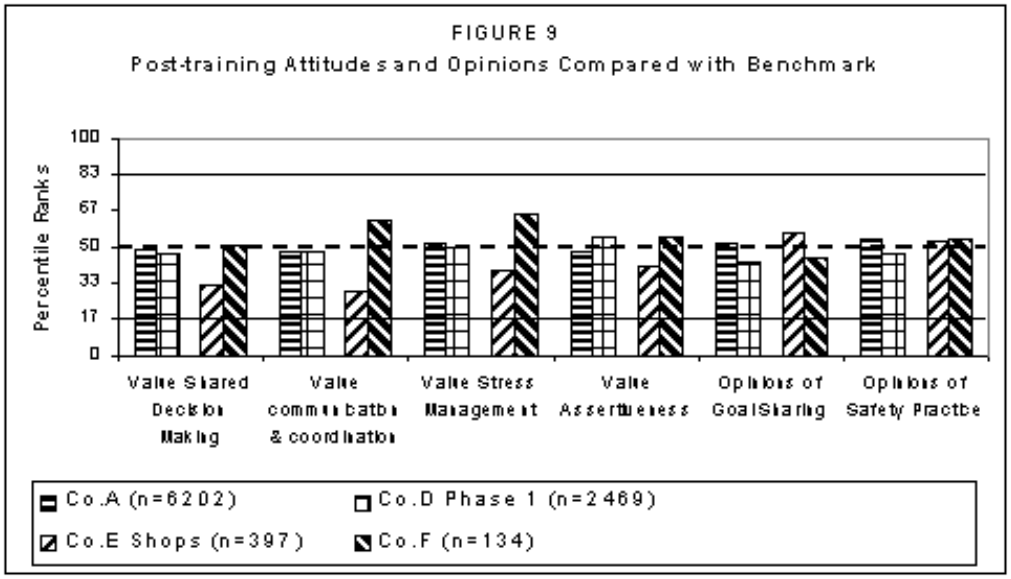

Figure 9

Post-training percentile ranks, displayed in Figure 9, continue to show Company F above the norm for Communication & coordination, and Stress Management. Likewise all of Company E's post-training attitude scores are substantially below the post-training norm. The former company's results show that it has benefited more than the average from the training, while the latter company shows that its poorer attitudes relative to the other companies studied remain low in the profile following the training. Despite its training's positive effects on Company E's stress management (as shown in Figure 7) participants' attitudes on that factor are still lower than most other companies.

Two-month and Six-month Follow-up Surveys

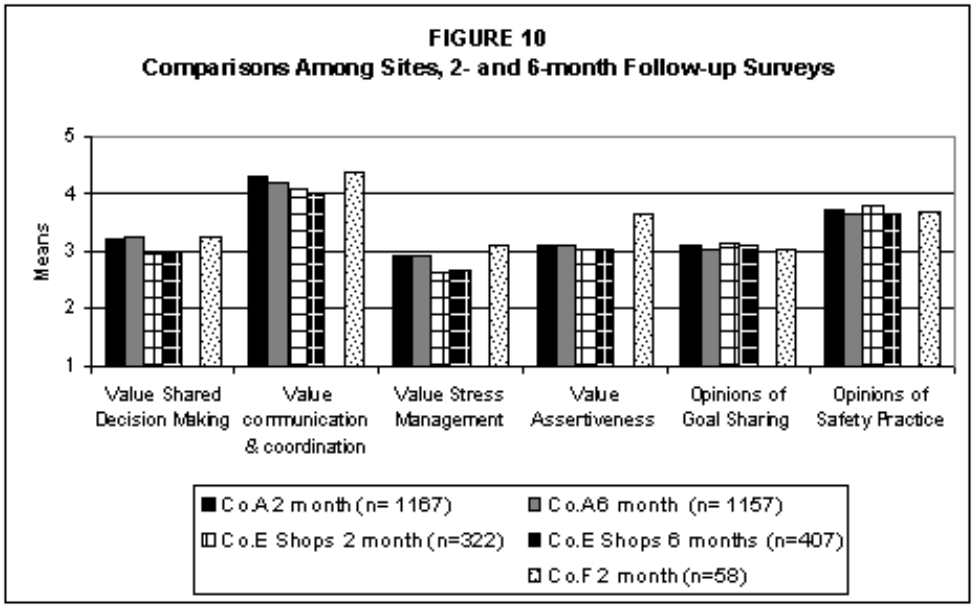

Figure 10

Figures 10-11 present the mean score and percentile results for the two- and six-month follow-up survey results.

The 2- and 6-month mean scores shown in Figure 10 reveal some differences among the three companies (A, E, and F), but very little difference between the two surveys for the same companies (i.e., A & E).

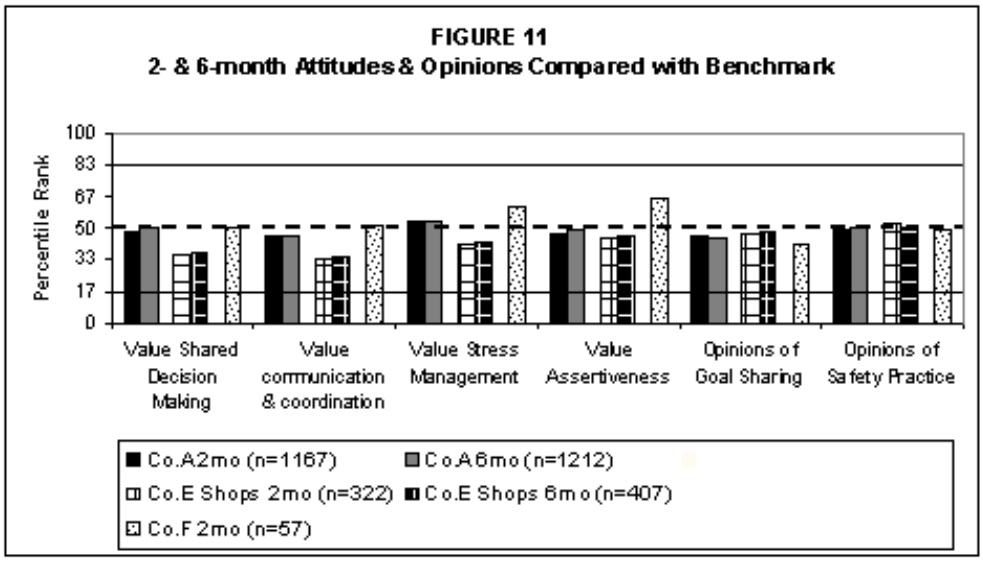

Figure 11

Their percentile ranks (Figure 11) place these differences among the companies in stronger relief. Company E is substantially lower on sharing command, communication & coordination, and stress management. Company F shows two-month survey scores for stress management and assertiveness that are substantially above the norm for all two-month surveys in the Benchmark database. Both of these MRM topics are emphasized in the training program used by Company F. This rise of assertiveness attitude strength for Company F between post-training and 2-month surveys is not statistically significant, but the effect is reminiscent of the delayed reaction and "positive transfer" of learned skills noted in an early MRM program (Taylor, 1995). Designing MRM programs to achieve such an effect is becoming a popular idea and one which is beginning to show positive results (Patankar & Taylor, 2000a). When MRM programs are targeted toward different purposes they will achieve different results.

Twelve-month Stand-alone Surveys

Figures 12 and 13 present the mean scores and percentile ranks for the 12-month stand-alone survey of AMTs in Company E and Company H. In both companies, these mechanics had participated in an MRM program more than a year before, but they had not completed either pre- or post-training survey at that time. This 12-month survey is the first time they have been asked to describe their views and feelings about MRM topics.

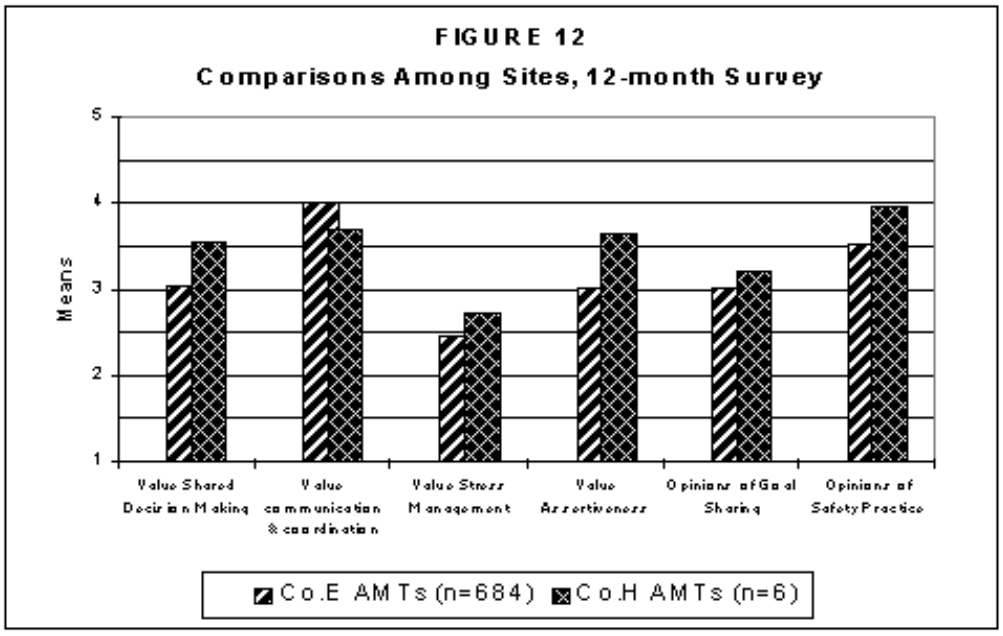

Figure 12

The mean scores in Figure 12 display a marked difference between the two company samples.

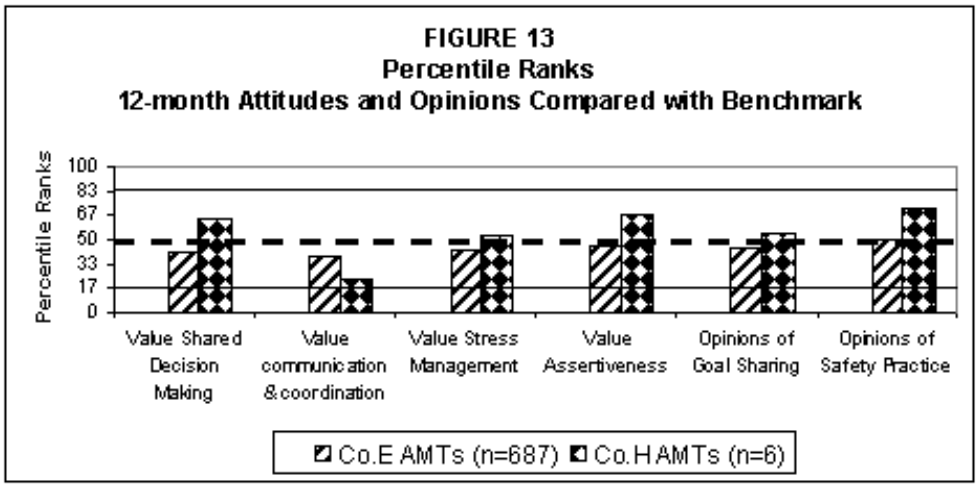

Figure 13

Examination of the percentile ranks in Figure 13 show that in fact Company E scores at a low to normal level compared with the Benchmark, and that Company H scores are substantially above the norm for sharing command responsibility, assertiveness, and safety practice.

Table 5A summarizes the results of the Benchmark comparison for all companies subsequent to their training.

TABLE 5A. COMPANY PERCENTILE RANKS SUBSEQUENT TO MRM TRAINING

SITE	Shared Decision Making	Communication & Coordination	Stress Management	Assertiveness	Goal Sharing & Setting	Maintenance Safety Practices
A	Normal- Norm-Normal	Normal- Norm-Normal	Normal- Norm-Normal	Normal- Norm-Normal	Normal-Norm-Normal	Normal-Norm-Normal
D	Normal	Normal	Normal	Normal	Normal	Normal

E-Shops	Low-Low- Low	Low-Low- Low	Low-Low-Low	Low-Low- Low	High-Norm-Normal	Normal-Norm-Normal
E-AMTs	Norm/low	Low/norm	Norm/low	Normal	Normal	Normal
F	Normal-Normal	High-Normal	High-High	High-High	Normal-Low	Normal
H	High	Low	Normal	High	Normal	High

OPEN-ENDED QUESTIONS

Benchmark Results

A generalized, but descriptive profile of respondents' reactions to their MRM training is obtained by combining benchmark respondents' initial answer to each of the four open-ended questions used in all of the post, two-, six-, and 12-month MRM/TOQ surveys. Figures 14-17 display the distribution of coded answers of all respondents for each of the four questions. The Benchmark database includes all MRM training participants between 1991 and 1999.

Benchmark: Reactions to the Training

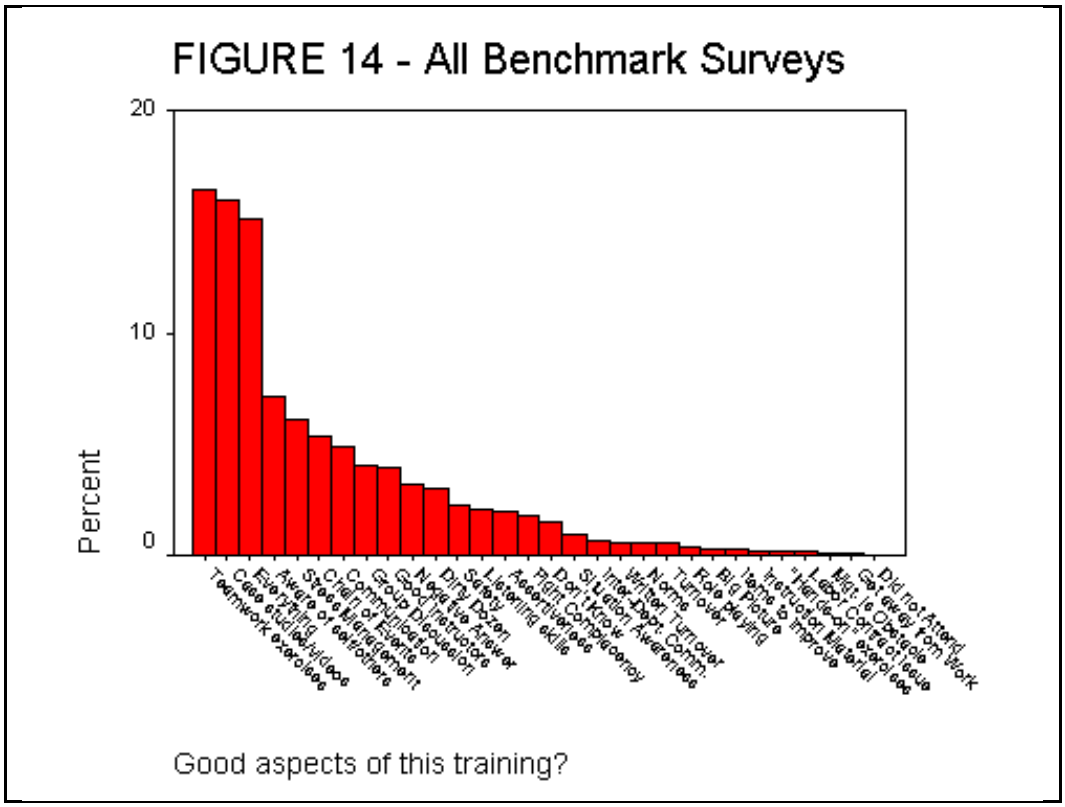

Figure 14

"Good aspects of MRM training?" Figure 14 shows the percentages for all Benchmark respondents for all topics they mentioned, for the question, "what was good about the MRM training?" The first three topics in this preference list – with over 15% each – are teamwork exercises, accident case study videos, and "everything." A distant fourth (7%) is the topic "awareness of self and others," following in fifth place is stress management (6%), and sixth is "chain of events in accidents" (5%).

"What would improve MRM training?" Figure 15 shows the Benchmark percent for all topics volunteered to the question, "What would improve the training?" Here an overwhelming one-third said that the training needed nothing and it was fine the way it was. Another 10% said they wanted more cases and videos, and slightly smaller percents said they wanted the program to continue as recurrent training, they wanted it lengthened, they wanted more managers trained, or they wanted more time for teamwork exercises. Maintenance personnel like MRM and want more of it. Over 40% mentioned some aspects of enlarging the length or the scope of the training.

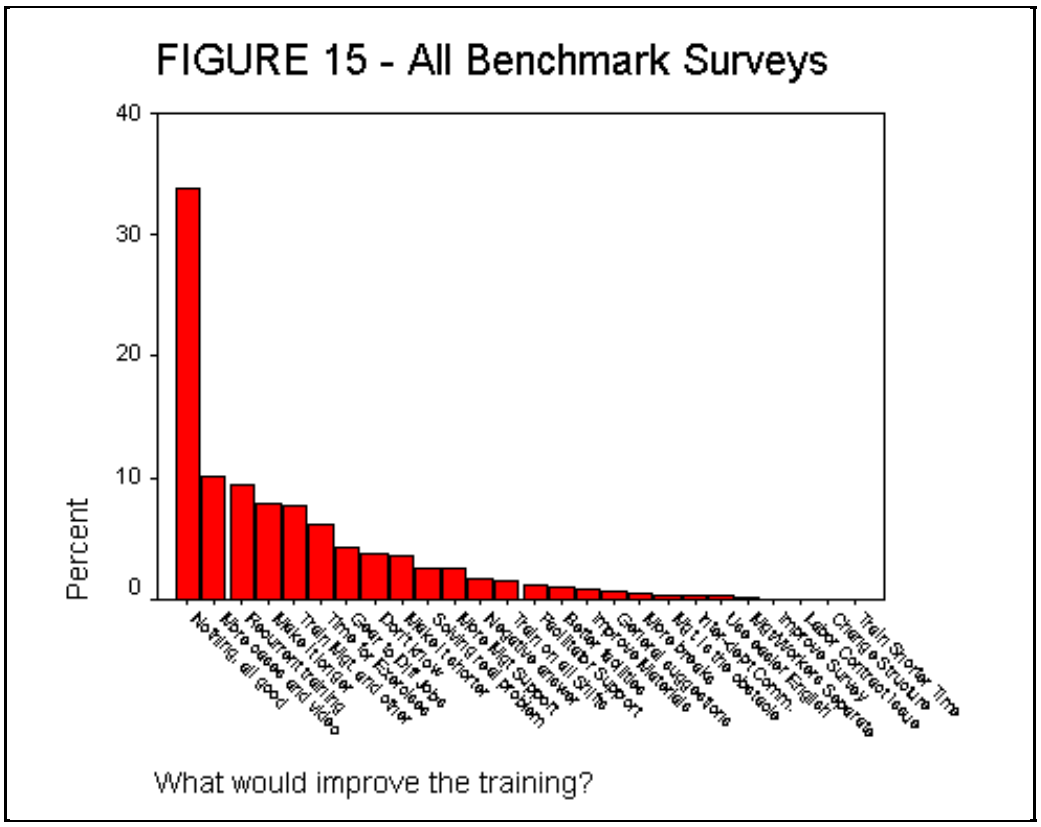

Figure 15

Benchmark: Behaviors Caused by Training

"How will you use MRM training on the job?" Figure 16 displays the answers volunteered to the question, "how will you use the training of the job?" For the Benchmark database – containing training participants from 1991 through 1999 – the first two topics say it all. The first (over 15% of all respondents) is to be thorough and to fight complacency in themselves. Another, slightly smaller percent say that they will be more interactive and promote teamwork. These two themes – passive coping and active communication – are the substance of MRM training.

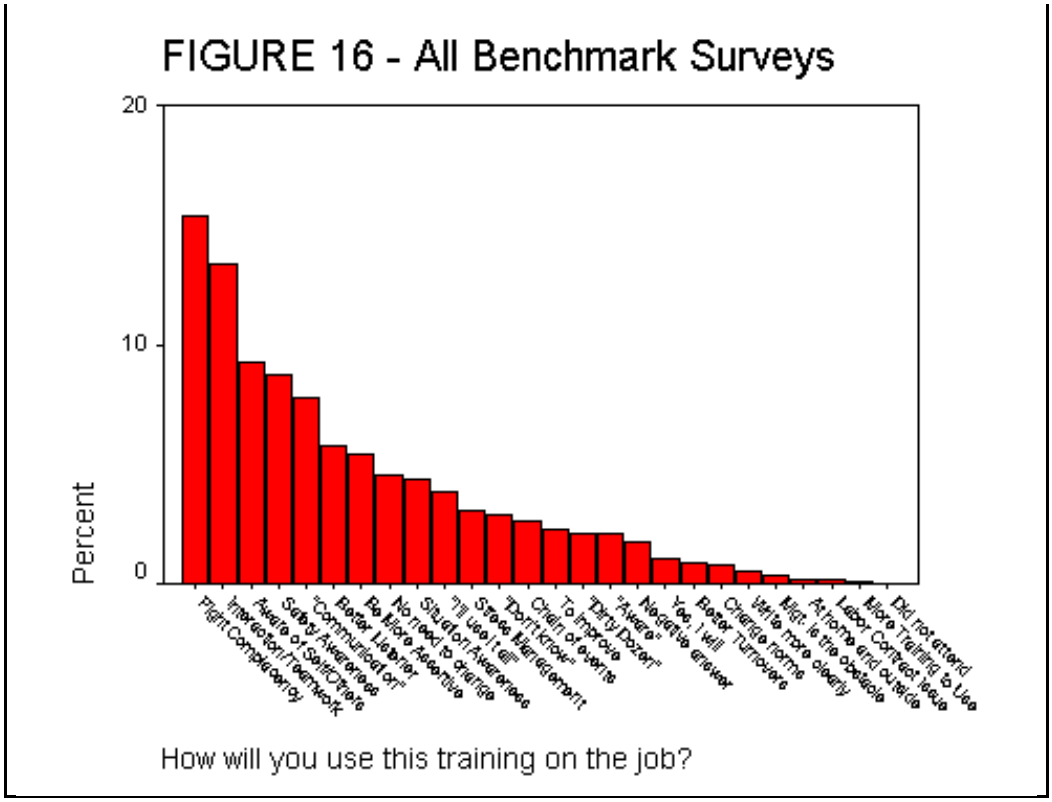

Figure 16

Note that stress management is ranked 11th of the topics to be used. It is a popular idea immediately following training, but not as attractive as a passive coping technique to be applied later by this population.

"What changes have you made as a result of attending MRM training?" Figure 17, "how have you used the training?," shows one in five respondents say they didn't change at all. The next two topics in order are awareness of self and of others (13%), and safety awareness (11%). About 8% each say they fought complacency. This absence of effect and the next three passive topics account for fully half of those answering this question. Another 8% say that they promoted teamwork -- an "active" behavior – and together with those who say they will communicate better (5%), or be more assertive (4%), don't add up to 20% of the total. These self-reported behaviors are at odds with the intentions to change that respondents voiced a few months earlier (and just reviewed in Figure 16).

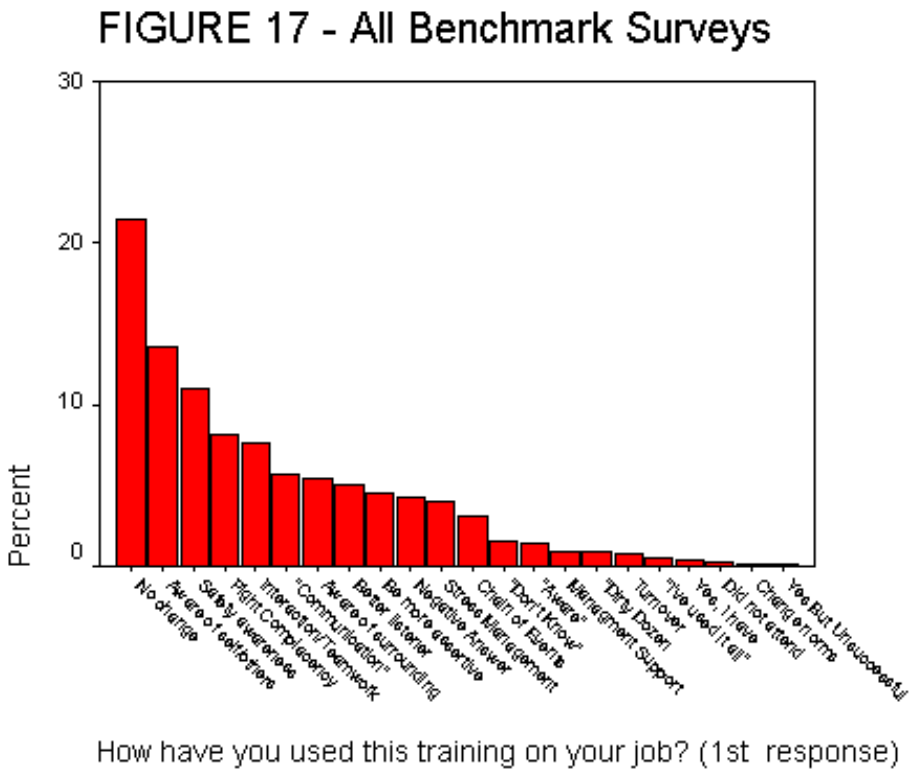

Figure 17

Companies 1998-1999: REACTIONS TO THE TRAINING

Respondents in each of the four companies examined in this report all answered the open-ended questions, and have their own specific reactions to MRM training. This section describes what respondents in companies D,E,F and A volunteered that they liked about their own programs and what changes they suggested.

Company D

Company D developed a two-phase training program for all maintenance employees. At the end of 1999 the program was still in progress. Figures 18 and 19 present what those participants have liked most about the program and how they wanted it improved.

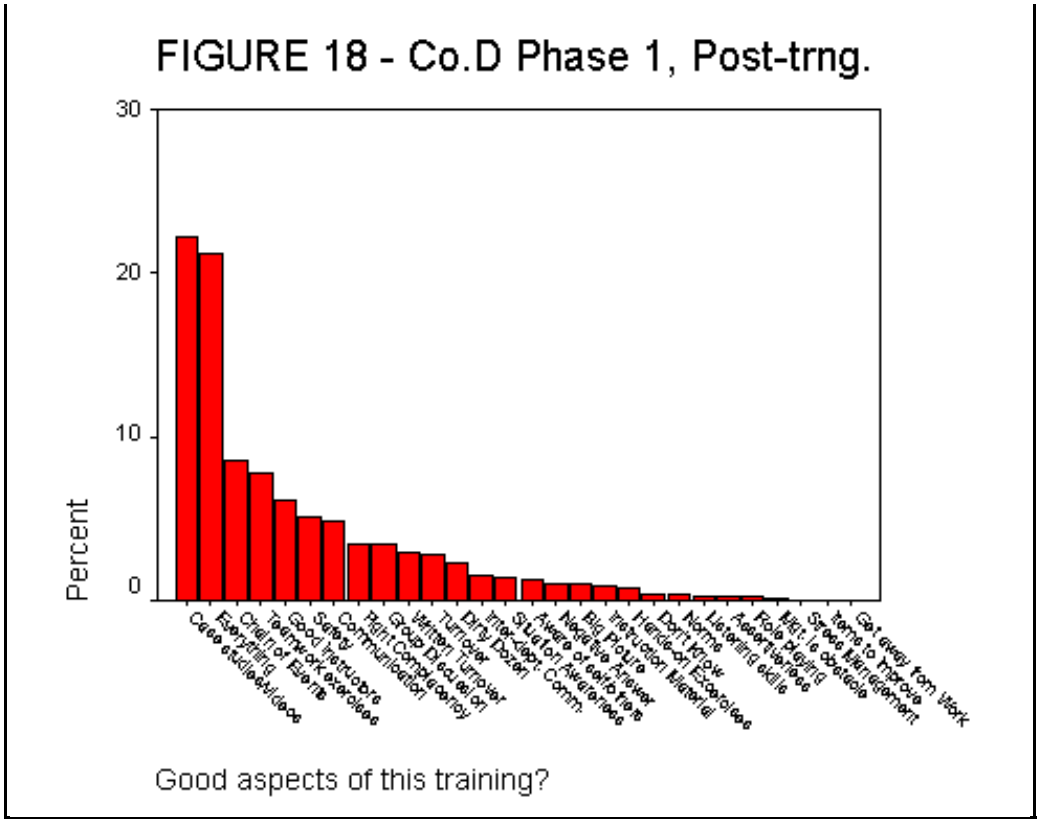

Figure 18

- Figure 18 shows that Company D's enthusiasm for the MRM program is initially very high (and following phase 2 these good feelings remain (*cf.*, Taylor & Patankar, 2000).

- Over 20% of Company D's respondents like everything about the training, and another 20% say they especially valued the accident case studies. In third rank about 8% said they most valued the idea of "chain of events in accidents," followed by a similar percent who liked the teamwork exercises best.

- Figure 19 shows that almost 40% of Company D's respondents have no suggestions to improve it and want it left the way it is. They obviously like the program and want it to continue as to grow.

Figure 19

Company E

During 1998 Company E administered a ½ day MRM awareness training course for component shop mechanics. The training coincided with intensive and emotional labor contract negotiations. Figures 20 and 21 display what was best liked about the course and suggestions its management under those circumstances.

- Figure 20 shows that a clear favorite of Company E's participants were the accident case videos.

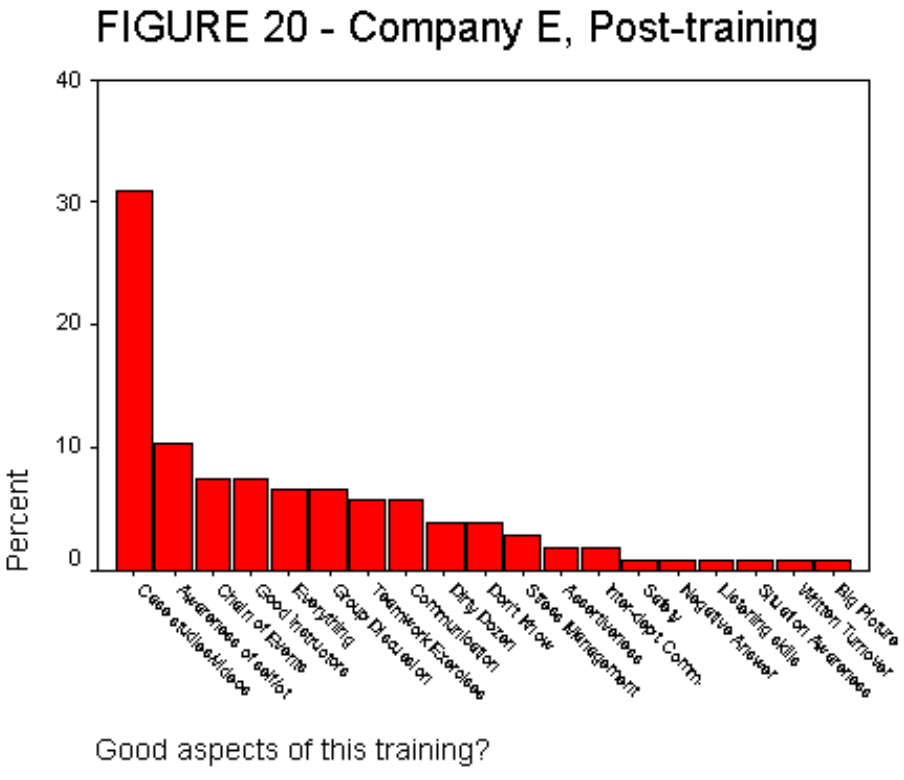

Figure 20

- Figure 21 shows that many of these respondents would also like more case videos.
- But there is little agreement about the program's length – some 12% in Figure 21 wanted the ½ day course made longer, while another 11% wanted it shorter than the four hours that it was.

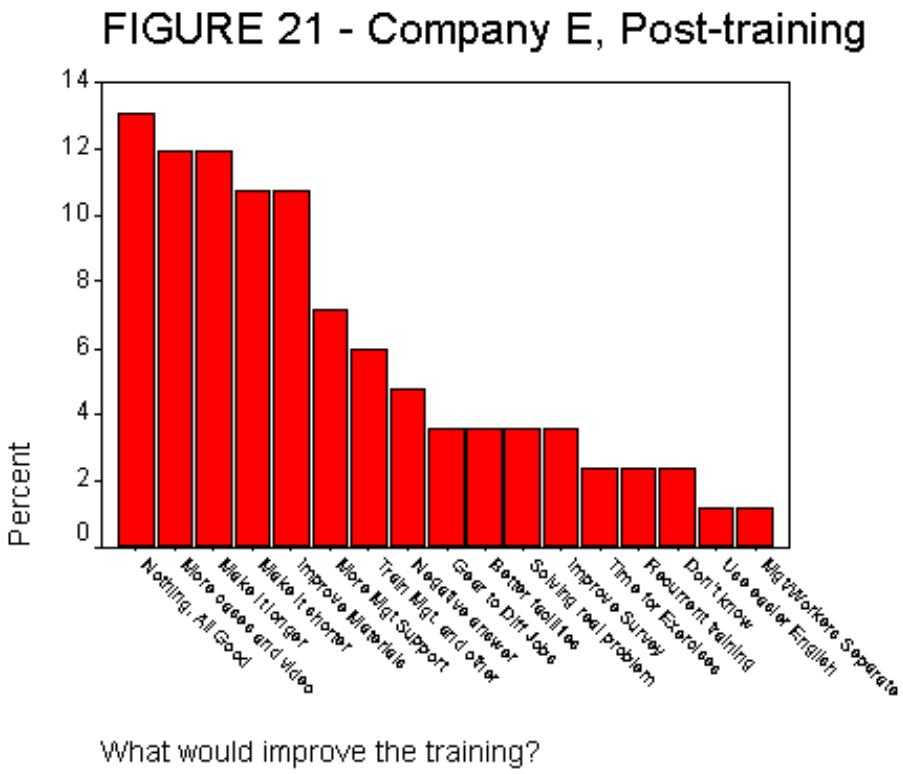

Figure 21

Company F

This is a relatively smaller maintenance operation than companies A, D, or E. All its line maintenance mechanics and foremen attended the two day MRM course presented by an outside vendor. Figures 22 and 23 show what respondents thought was best, and what could be improved.

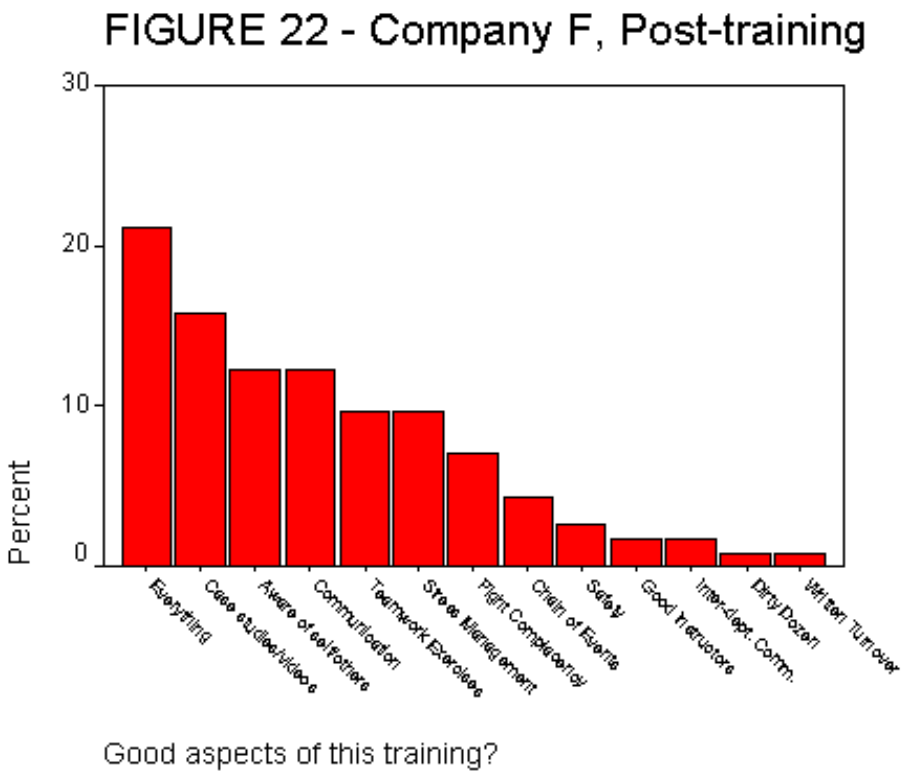

Figure 22

- Figure 22 shows one third of Company F's particpants most liked either "everything" about the course, or case studies; followed by 11% who most valued the self-awareness they learned about. Communication and teamwork exercises were the next most popular.
- Figure 23 shows a remarkable 45% of all Company F respondents said the course was great the way it was – and required no changes. Smaller proportions of participants suggested a variety of improvements which reveal the same sentiment reflected in other companies results – to enlarge the size and scope of the MRM program.

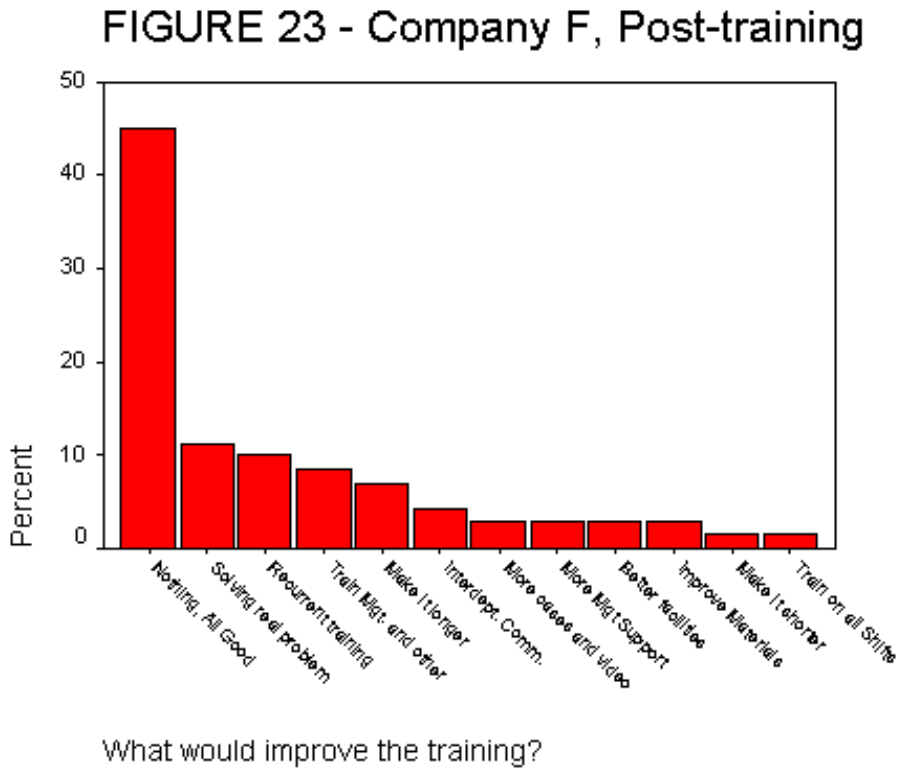

Figure 23

Company A

Company A's two day MRM training course was delivered to most of its AMTs, and to its material services personnel (*cf.*, Appendix B). By late 1998 the training had been concluded. Figures 24 and 25 show for Company A what participants felt was good about the training and what they through could be improved.

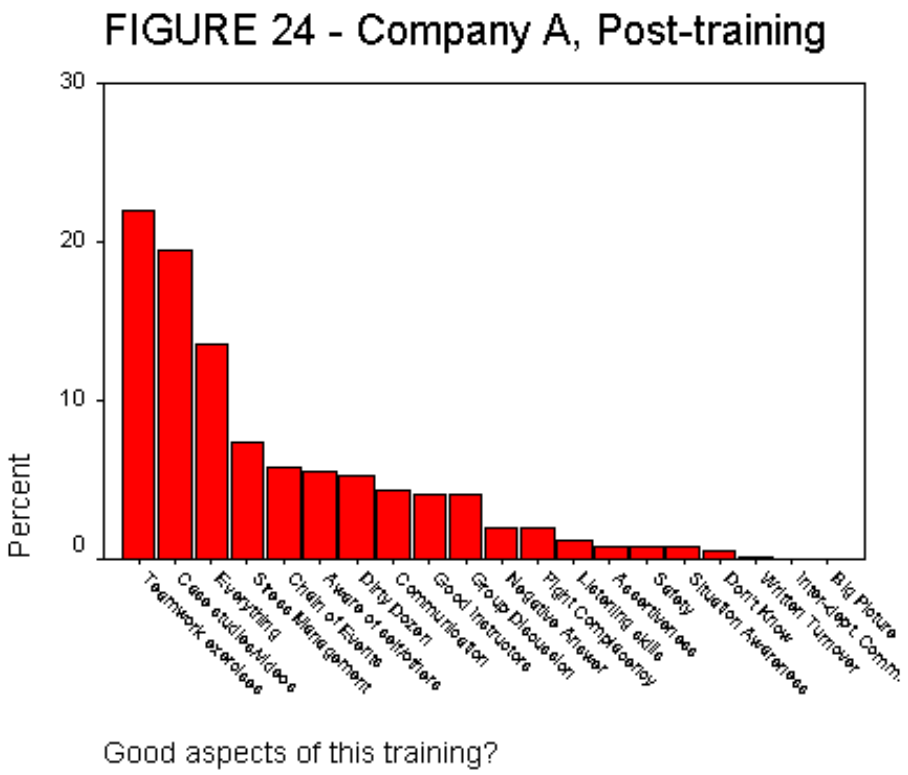

Figure 24

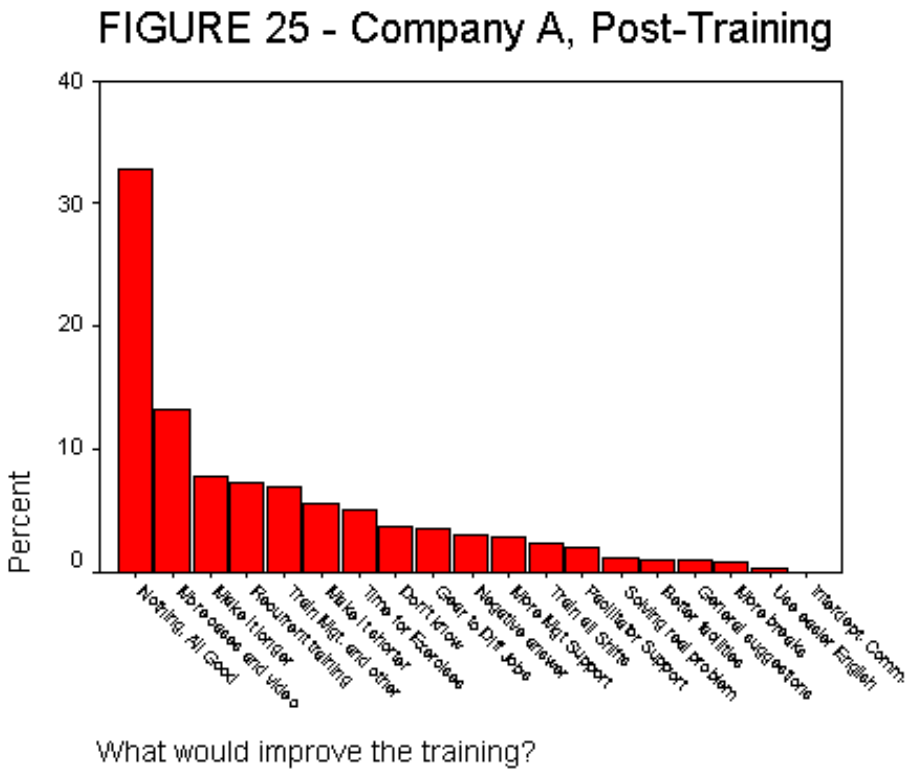

Figure 25

- In Figure 24 , the teamwork exercises were most liked by the largest proportion (21%), followed by those who most valued the accident case studies (20%), then those who said "everything about the course" (13%), while stress management was fourth with 7% who like that topic the most.

- Figure 25 shows that one third of Company A's respondents said no improvements were necessary. Most of those who had suggestions about the training tended to want more of it – "make it longer" 8%, "hold recurrent MRM training" 7%, "train management too" 7%.

Companies 1998-1999: Behaviors Caused by Training

Intentions To Change And Subsequent Change Reported

In addition to eliciting reactions to the MRM training, the open-ended questions help establish a company's direction of change through its respondents' stated intentions to change their behavior following the post-training survey. The follow-up surveys -- two and six months following the training – provide reports of subsequent behavior change. This comparison of intended and realized behavior is available for all four companies A, D, and E and F as well. In this section comparisons will be made between intentions and self-reports of actual behaviors for active and passive behaviors. These comparisons between intended and actual change are viewed in the context of the "purpose" of those companies' MRM programs for example. If a program is designed to heighten individual awareness then individual awareness is what we expect respondents will intend and what they should subsequently report.

Collapsing specific codes into "active" and "passive" categories. An essential difference among MRM programs is the degree to which they encourage active interpersonal behaviors (e.g., speaking up, speaking out) versus more passive individual coping behaviors (e.g., being more aware and fighting complacency). Changes in behaviors should reflect these program differences. The several companies studied during 1998-1999 are compared on the active communication versus the passive coping behaviors reported by their maintenance employees. To aid the presentation of intended behavior changes and subsequent self-reports of behaviors changed, the specific, written responses are divided into "passive" and "active" behavior categories. Table 6 shows the specific codes that are combined into active and passive categories.

TABLE 6. Transforming Answers of Open-Ended Questions Specific Composition of New Change Categories			
"Passive" Category	**"Active" Category**	**"No Change"**	**"Other"**

• **I fight complacency**	• **I interact with others**	• **I won't (or didn't) change**	
• **Aware of self & others**	• **I communicate**	• **I am safe already**	• **All other codes**
• **Aware of safety hazards**	• **Better shift turnover**	• **I don't know how I'll change**	
• **Situation awareness**	• **Being assertive**		
• **Stress management**			
• **Being a good listener**			

Company A is used here in two ways. First, it will be the initial case examined in this section. Company A is also used here as an illustration of this transformation of specific topics from the answers to the open-ended questions to the four categories shown in Table 6.

The other three companies; D,E, and F will be presented only in terms of the four collapsed categories.

Company A MRM: A Single two-day AMT Training Session

Company A's MRM training program was completed for all of its line AMTs and for many (approximately two-thirds) of its base maintenance AMTs. The purpose of the program, stated in the participant's workbook, was to create an awareness of the impact of human performance on maintenance-related errors and personal safety. The learning objectives for the course were as follows:

- Relate how AMT characteristics and personal behavior can impact the maintenance process
- Identify 12 performance factors ["the dirty dozen," *cf.*, Taylor & Christensen, 1998, pp. 145-6] and their role in the chain of events leading to maintenance-related errors
- Develop personal techniques to minimize risk and maximize performance
- Give and receive feedback with coworkers related to personal safety

Company A trained over six thousand employees during a two and a half year period. It addressed its MRM training exclusively to AMTs (supervisors and managers account for less than 1% of the total trained in that company). The AMT's union and the company's management cooperated to initiate the training. Company A's syllabus emphasized the dangers of complacency, the effects of stress and fatigue, and communication in its core curriculum. Training took place at the local level with facilitators coming from the ranks of both AMTs and their first-line supervisors. The training was coordinated and supported by the company's training and education department.

Likelihood of voluntary change. Enthusiasm was positive immediately following Company A's training even if some participants hedged a little on their interpretation of substantial change. Over sixty percent of Company A's participants in the post-training survey said there would be a moderate or large change in their on-the-job behavior (Figure 2 above, shows Company A's post-training mean score for "Training will change my behavior" is 2.69). Although a clear majority believes that the training will affect their actual behavior, this level of enthusiasm does not approach the high ratings -- 80% and 90% (with mean scores of 3.03 and 3.23 respectively)-- reported for the earlier MRM cases (Taylor & Robertson, 1995; Taylor, Robertson & Choi, 1997).

Specific intentions to change. Figure 26 shows the relative percentages of the specific coded topics. The results in Figure 26 show that the six passive intentions in the first dozen topics (Be thorough, 15%; Aware of safety hazards,11%; Aware of self…, 10%; Situation awareness, 5%; Better listening, 4%; and Stress management, 4%) total 50%. Active intentions (Interaction, 14%; Communication, 7%; Assertiveness, 6%), also among the first dozen topics in Figure 26, total only 27%. The code "I'll use it all" is placed in the "Other" category, along with most of the remainder of the topics in Figure 26.

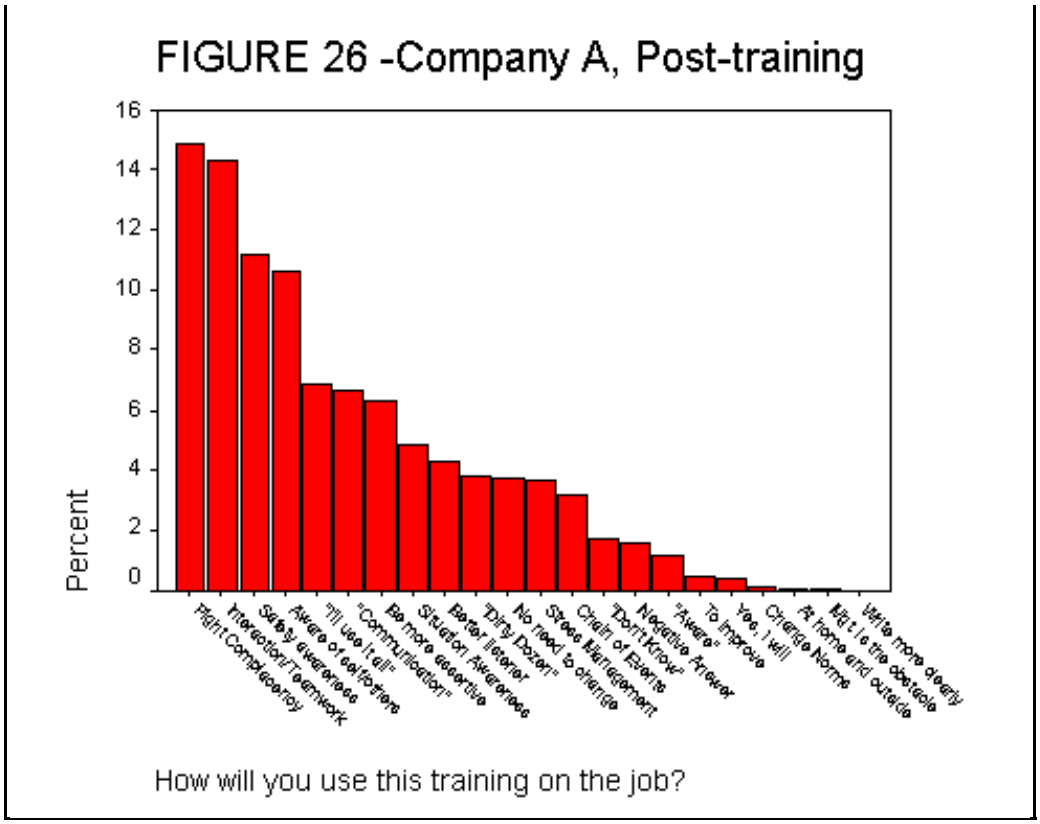

Figure 26

The resulting distribution into those four larger categories – featuring "active communication" intentions and the more passive coping behaviors for comparison – are shown in Table 7.

Table 7 Company A Post-training Behavior Intentions (n=4613)	
Total Passive Intentions	**50%**
Total Active Intentions	**27%**
Other	**19%**
No Change Intended	**4%**
	100%

As shown in Table 7, some 27% of Company A AMTs' responses were coded in the active category while 50% were coded in the passive category.

These post-training results in Table 7 are presented in graphic form (in Figure 27) together with Company A's two and six months intentions, and the self-reports for actual behaviors used.

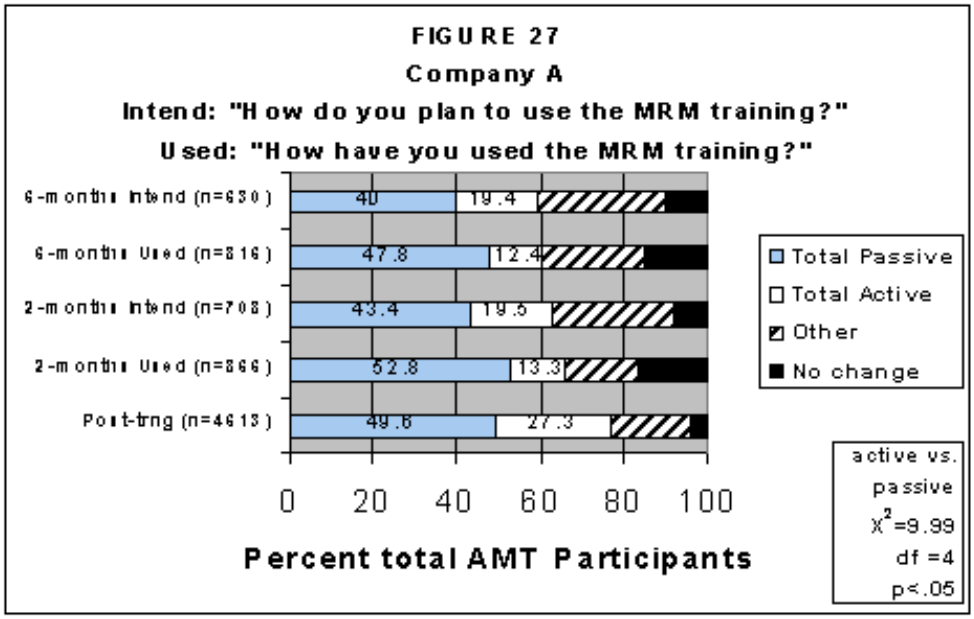

Figure 27

The two-month reports of actual behavior shown in Figure 27 reveals the proportion of "active" behaviors is half as great as the post-training active intentions (13.3% < 27.3%) and the proportion reporting "no change" is more than four times greater than its earlier counterpart. The proportion of active intentions at two months and six months are similar to one another at about 19.5%, but substantially smaller than immediately following training (27.3%). Reports of actually using passive coping behaviors are similar two and six months after training (52.8% and 47.8%) and they are quite consistent with preceding intentions. Only a little over a quarter of Company A's AMTs initially intend to actively engage others, and only about half of that number subsequently report behaving that way.

This result for Company A is substantially lower than earlier MRM programs in active communication. One of the early programs involved a Maintenance management sample (Taylor & Robertson, 1995). The other was an AMT sample (Taylor, et. al, 1997). However, this Company A tendency toward passive coping behaviors is consistent with the purpose and objectives of their MRM program. Results of the Chi-Square test show that the variations between active and passive behaviors over time are statistically significant.

Company D MRM: Distributing MRM training over several months.

Company D adapted its MRM training by dividing two days of training over several months. The first day of training, called "Phase 1," is followed two to six months later with "Phase 2," a second full day of training. Company D created its own MRM training after reviewing the one-shot training used in earlier programs. The reasons for this adaptation were 1) to avoid overloading maintenance personnel with information in a single large session, 2) to provide concepts in the first phase and skills training in the second phase, and 3) to illustrate by example the importance of recurrent training in its MRM program. The AMT's union and the company's management cooperated to initiate and design the training. Training materials were inspired by the earlier programs, but the most of the exercises and cases were created specifically for this application.

Program purpose. Company D's purpose for MRM, as stated in their participant's workbook is, "To provide participants with specific human factors principles and techniques to help them work more safely." The definition of MRM, also included in the participant's workbook, "…is the process where we work together, using available resources, to reduce errors and to promote safety." The statement goes on to say, "MRM addresses human factor errors and problem resolution through open and honest communication between all maintenance operations personnel, and with the FAA."

The training topics for the first day are:

- Identify human factors elements
- Recognize the "dirty dozen" error causes

- Identify the chain of events in accidents
- Effective written communication
- Identify norms
- Establish safety nets
- Recognize safety mechanisms

Although the MRM definition quoted above is more active and interpersonal than Company A's, the supporting topics are largely "awareness" or conceptual issues --with "written communication" as the "active skill" exception.

At the beginning of the second (Phase 2) training day the definition of MRM is reiterated. The training topics in the participants' workbook for the second day are as follows:

- Recognize the nature of errors and how they affect participants
- Focus on how to manage errors
- Introduce tools to use in error reduction

Phase two also places emphasis on "dirty dozen" topics, "lack of assertiveness" and "lack of awareness" as well as situation awareness. As in phase 1, these topics for phase 2 training seem more conceptual than behavioral. The module on lack of assertiveness is, however, focusing on active communication. On the other hand, the main "tool" in the final phase 2 topic list, situation awareness, is an individual, passive mechanism. Company D's MRM program appears to be bridging between the Company A's model of individual AMTs coping with safety hazards and issues, and the interpersonal communication techniques of the original maintenance safety training (Taylor & Robertson, 1995; Taylor & Patankar, 2000).

By design, phase 2 (the second day of training) is conducted as close as possible to two months after the first one.

The two Company D city stations reported here. The course is designed for all maintenance employees (including managers and support staff) and each session is expected to include management and hourly employees from a variety of functions within maintenance. Initially, the training took place in a large line station and both phases 1 and 2 were completed there before the program was moved to two cities containing both base and line maintenance stations. Company D expects that all 8,000 maintenance employees throughout the system will eventually attend the training.

Phase 1 training for city 2 (the large line maintenance station) was 85 percent completed between January and March 1998 and the remainder (for a total of some 500 maintenance employees) was finished in July. City 2's phase 2 training was completed during August and September 1998. Company D City 1's MRM training included both a large line station and a major heavy maintenance base. City 1 began phase 1 training in September1998 and completed it with about 1,000 maintenance personnel in April 1999. Phase 2 began in city 1 during June 1999 and was about 50 percent completed by December 1999. A third city (also both a large line station and a major heavy maintenance base) began phase 1 training in July 1999 and, with over 900 employees attending, had not yet been completed by December 1999. Results from cities 1 and 2 will be used below to illustrate the effects of distributed training and the modified course purpose and topics.

Likelihood of voluntary change. Enthusiasm for all of Company D is moderate when compared with past MRM experience described for the earlier cases (*cf.*, Taylor & Robertson, 1995; Taylor, et al., 1997; Taylor & Patankar, 2000). Sixty-eight percent of the participants following phase 1 said there would be a moderate to large change in their on-the-job behavior. Figure 2 above shows the Company D post-training mean for "Training will change my behavior is 2.80. Following phase 2, 75% Company D participants said there would be at least a moderate increase in their at-work activities and the phase 2 mean score is 3.00. This modest increase is encouraging and statistical tests of this difference are significant. Furthermore, this Phase 2 mean score of 3.0 is commensurate with the successful early AMT program referred to above (Taylor, et al., 1997).

Attitude changes. Figure 7 in an earlier section shows that immediately following the Phase 1 training, Company D participants' attitudes reveal significant improvement in attitudes toward communication, stress management and assertiveness. Following Phase two training all three attitudes increased again significantly. Although attitudes toward sharing command responsibility increase slightly over this time, the differences are not statistically significant.

Opinion changes. Figure 7 also shows Company D participants' evaluation of their station's goal setting and sharing remained unchanged between phase 1 pre and post-training surveys. However their evaluations of the station's safety climate decreased significantly ($F=8.29$, $p<.001$) between phases 1 and 2 (Taylor & Patankar, 2000). Field observation at Company D's city 2 some 60 days after phase 1 training and again four months after phase 2 no resolution. On one hand our field observations revealed that current ground damage statistics for city 2 and the total maintenance system were readily available to all employees. This information, in the form of monthly posters, should heighten safety awareness, but that information did not emphasize longer run trends and it was not always current. Our field interviews did confirm the survey results -- AMTs, leads, and foremen reported that safety standards and programs seemed to be deteriorating. Apart from their own individual care and awareness, they said, little was being done to support maintenance safety in the station.

Specific intentions to change. The questions, "how will you use this training on the job?" and "What changes did you make?" were included in the surveys that followed the phased training in Company D. Results from those questions were transformed into the four categories shown in Table 6 above (p. 46). These data are presently available for city 2 in its entirety as well as for first half of the city 1 participants who have completed phase 2 training.

Figure 28 presents the expected behaviors for both cities at the end of both phase 1 and 2 training. The figure also shows the actual behaviors reported by participants at the time of the phase 2 training. Although only11.3% in city 2 said they intended to communicate actively following phase 1, an even smaller 8.2% reported having done so when they returned for phase 2 training. City one results show consistent proportions (almost 9%) reported for using active communication following phase 2 and what respondents said they intended immediately following phase 1. Also, that nine percent in city 1 who reported actively communicating is virtually the same as city 2 who reported having been more active communicators when they returned for phase 2. These results in figure 28 do not favorably compare with the proportions of active to passive intentions found in cases in the first generation of MRM training (cf., Taylor & Robertson, 1995; Taylor, et al., 1997). But the total of behaviors (passive + active) which range between 60% and 80% in Figure 28 do exceed those same totals for company A (Figure 27). Results of the Chi-Square test for Company D show that the proportional variations in active and passive behaviors over time are statistically significant.

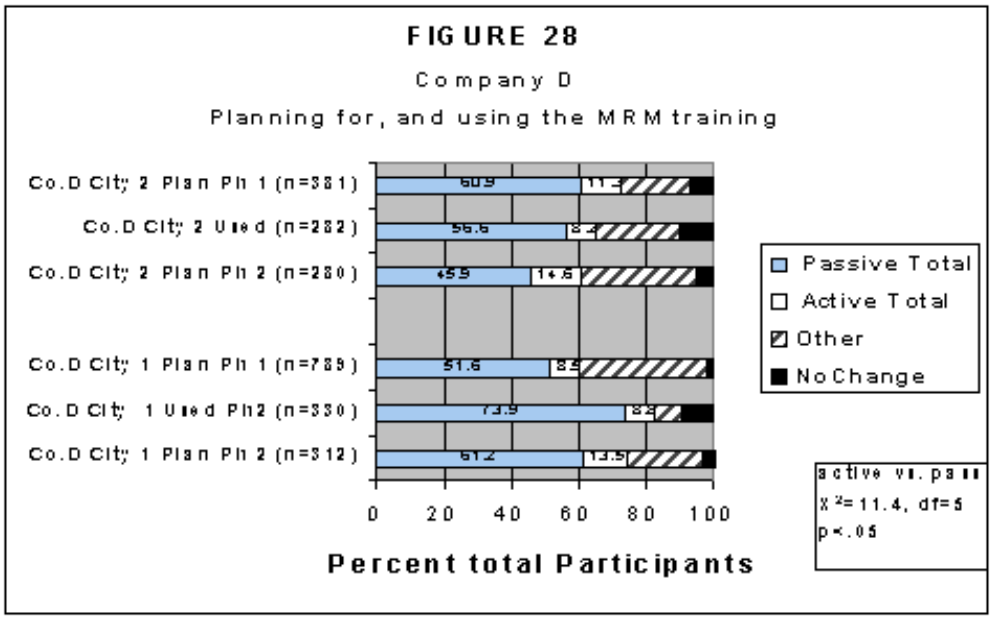

Figure 28

Company E MRM: A Single ½ Day Training Session For Component Shop Mechanics

Purpose. Company E's MRM program focused on awareness of the ways human factors lead to safe operations. Emphasis is placed on four of the "dirty dozen" -- lack of communication, complacency, distraction, and pressure -- three of which are primarily personal issues that can be best managed by the individual. The communication focus is a balance of "listening skills" and "assertiveness" -- both passive and active.

Company E trained over 1,200 employees during a six month period in 1998. It addressed its MRM training exclusively to shop mechanics (supervisors and managers account for less than 1% of the total trained in that company). The AMT's union and the company's management cooperated to initiate the training, but the two parties were engaged in heated contract negotiations during the time the training. Company E's syllabus emphasized the dangers of time constraints and interrupted (or "broken") work. Training facilitators were employees of the company's maintenance training department.

Likelihood of voluntary change. Enthusiasm was positive immediately following the training. Eighty-eight percent of Company E's participants in the post-training survey said there would be a moderate or large change in their on-the-job behavior (Figure 2 above, shows that Company E's post-training mean score for "Training will change my behavior" is 3.15). This level of enthusiasm falls within the high mean scores of 3.03 and 3.23 reported for the earlier MRM cases (Taylor & Robertson, 1995; Taylor, el al., 1997 respectively).

Specific intentions to change. As with companies A and D, the questions "How will you use this training on the job?" and "What changes did you make?" were used for Company E. Also as above, the results from those questions were collapsed into active communication and passive behaviors.

Figure 29 presents the expected behaviors at the end of post-training, and actual self-reports from the two- and six-month follow-up surveys.

These Company E results for expected and actual behaviors following the training are consistent with that company's enthusiasm for change shown in Figure 3. Only 60% total passive and active change is intended immediately after the training ("post plan" in Figure 29).. Furthermore that 60% total is not subsequently exceeded in Company E's 2-month or 6-month samples. Finally the percentages of respondents who say that they will not and did not change is very large compared with companies A and D.

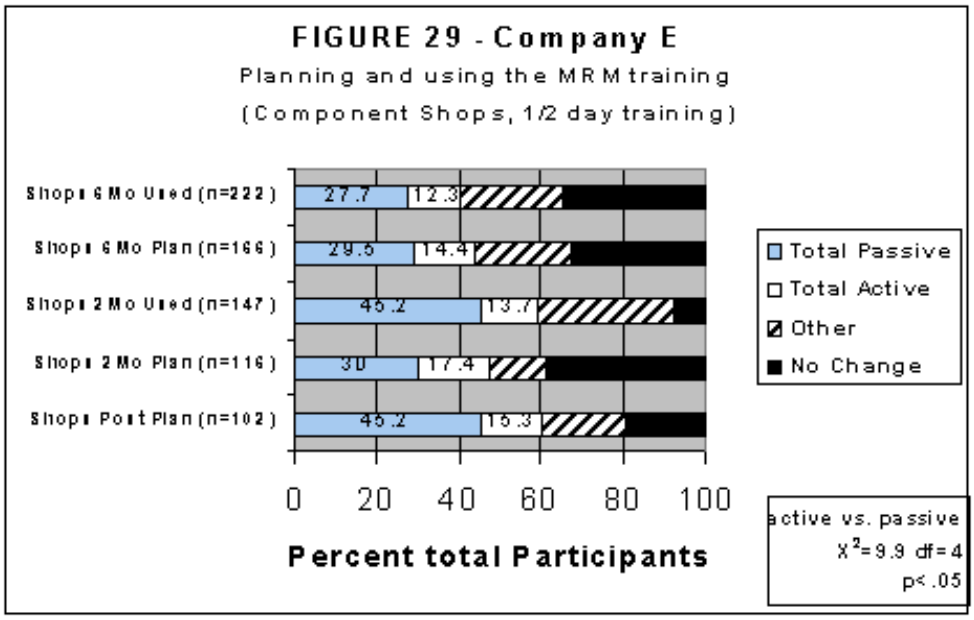

Figure 29

Company F MRM: A Single Two Day Training Session for Line Mechanics

Company F's MRM training program was completed for all of its line and line-hangar mechanics. The program focused on awareness and coping mechanisms or safeguards. The purpose of the training model was "To eliminate or reduce the causes of technician related errors, to enhance safety, and advance professionalism of the technician." Topical targets included human role in the chain of events in accidents, and safety nets for breaking the chain. The five safety nets emphasized human factors aspects of assertiveness, self-awareness, stress management, enhanced problem solving and decision making.

Company F trained some 135 AMTs and Foremen during a six month period in 1998-1999. Training was designed and conducted by an external MRM training vendor.

Likelihood of voluntary change. Enthusiasm was very positive immediately following the training. Eighty percent of Company F's participants in the post-training survey said there would be a moderate or large change in their on-the-job behavior (Figure 2 above, shows Company F's post-training mean score for "Training will change my behavior" is 3.06). A clear majority believes that the training will affect their actual behavior, which mirrors the high rating -- 80% with mean = 3.03) -- reported for the earliest AMT MRM case (Taylor, Robertson & Choi, 1997).

Specific intentions to change. The question, "how will you use this training on the job?" was included in the Company F post-training survey. The questions, "how will you use this training on the job?" and "What changes did you make?" were included in the 2-month surveys that followed the training in Company F. The results from those questions were collapsed into active communication and passive behaviors.

Figure 30 presents the expected behaviors at the end of both post-training and two-month follow-up surveys.

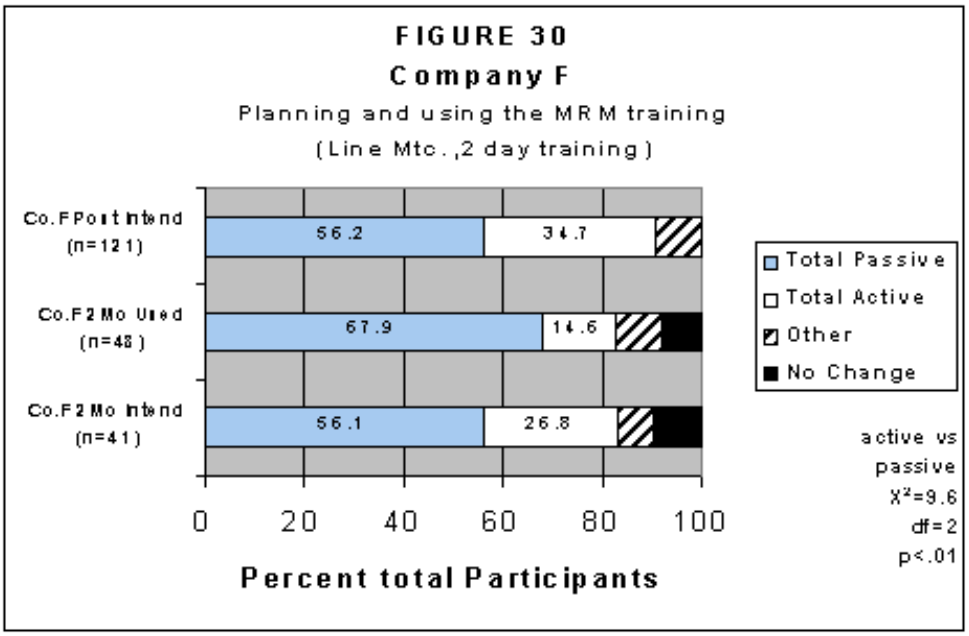

Figure 30

Figure 30 shows that a sizable 34.7% said they intended to communicate actively following the training, but a smaller 14.6 % reported having done so when asked in the two-month survey. Respondent intentions expressed in the two month survey are substantially lower than intentions expressed earlier. This ratio in figure 30 (generally ranging 15 to 35 % active to 55-65 % passive) is rather less than the 40 % active to 45 % passive intentions in the earlier programs (Taylor & Robertson, 1995; Taylor, et al., 1997). The total passive + active behaviors intended and subsequently reported are all above 80% -- and that is very high for the1998- 1998-1999 samples examined here. Results of the Chi-Square test for Company F show that the differences between intended and realized active and passive behaviors over time are statistically significant.

TRENDS IN MAINTENANCE SAFETY PERFORMANCE BEFORE AND AFTER THE ONSET OF MRM TRAINING

Company A Safety Performance Trends

In Company A two measures of Maintenance Department performance have been used. These measures are the frequency of lost time injuries (LTI), and the frequency of "ground damage" -- maintenance-related aircraft damage incidents -- ("GD"). Both measures are now available by work unit by month for the five years 1995-1999. The statistics plotted in the following charts 31-33 are average incident rate of all work units that participated in company A's MRM training. These monthly performance data are plotted in series "before," "during," and "after" the MRM training. All figures show "before," "during" and "after-training" linear trend lines (obtained using the method of "Least Squares") superimposed over the actual monthly data points. In practical terms the before-training series for Line Maintenance stations spans the period January 1995 through June 1996. The "during-training" data series for Line Maintenance runs July 1996 through late 1997. The "after" period for Company A line maintenance is the 24 months of 1998-1999. Because Base Maintenance didn't begin its MRM training until later, the "before" series for that group runs through March 1997. "During training" period runs from March 1997 through mid 1998 and the "after period" for base maintenance covers the 18 months to the end of 1999.

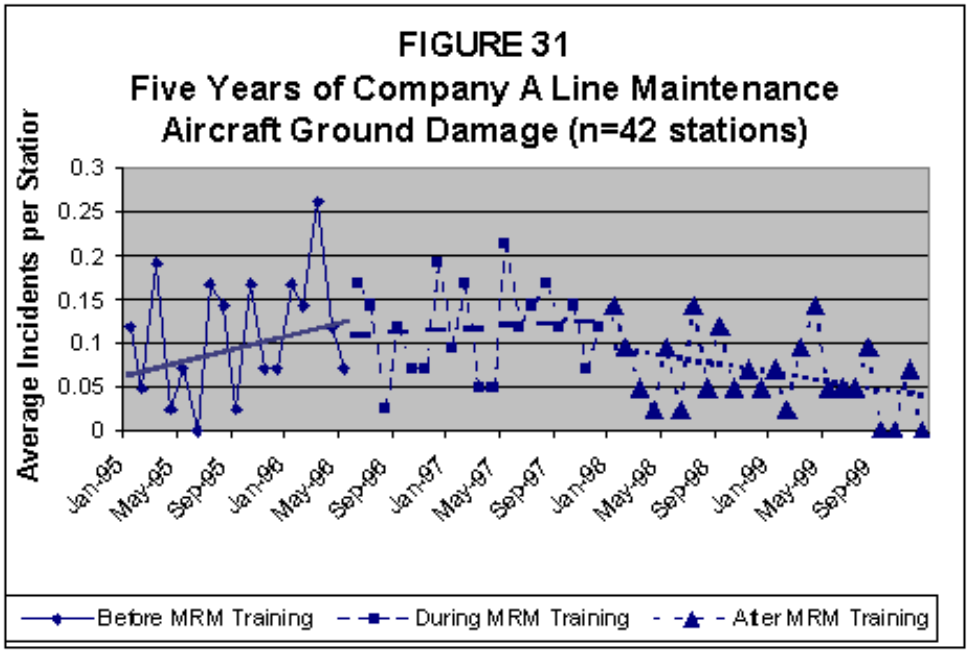

Figure 31

Figures 31 and 32 show the trends for aircraft ground damage and occupational injuries for line maintenance performance. Figures 33 and 33a shows similar trends for occupational injuries for base maintenance.

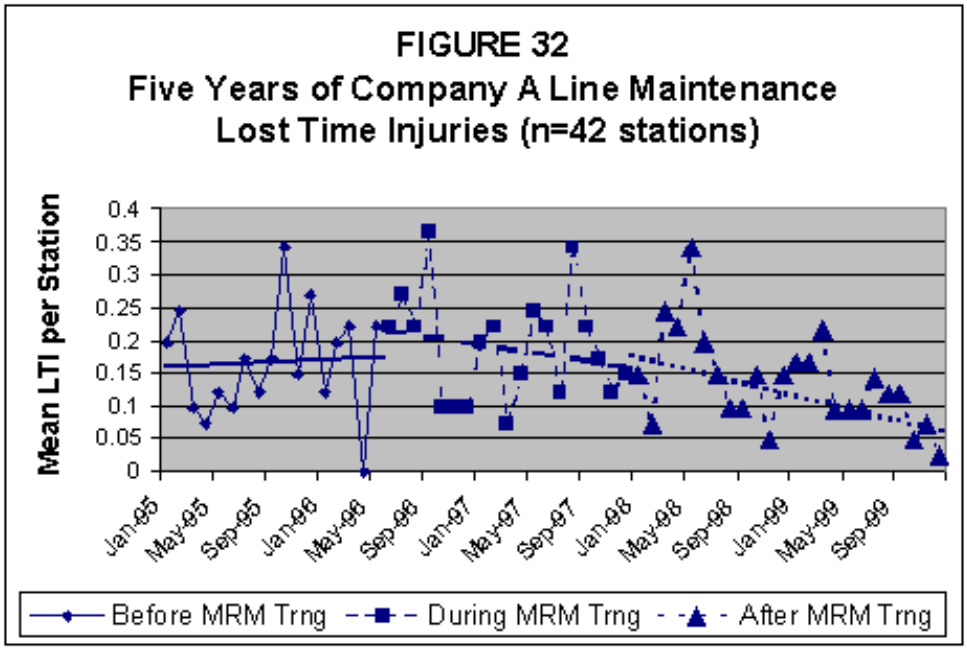

Figure 32

Line maintenance results. Clear from the trends in Figure 31 and 32 is that a dramatic improvement occurs for the line stations taken together. Furthermore, this improvement occurs directly after the onset of the MRM program and its rate of change continues in the two years following the completion of the MRM training. This strongly suggests that the "awareness" program works through its effect on individual actions over a substantial period of time – at least in this company's line maintenance organization.

Base maintenance results. In the previous milestone report from this research project (Taylor,1998), the performance trends for 1995 through 1997 for Company A were also promising. However, at least for base maintenance AMTs who hadn't yet completed their MRM training, more time would be necessary to observe performance. Now the Base Maintenance has concluded its MRM program, and for those units that participated, the effects are also encouraging.

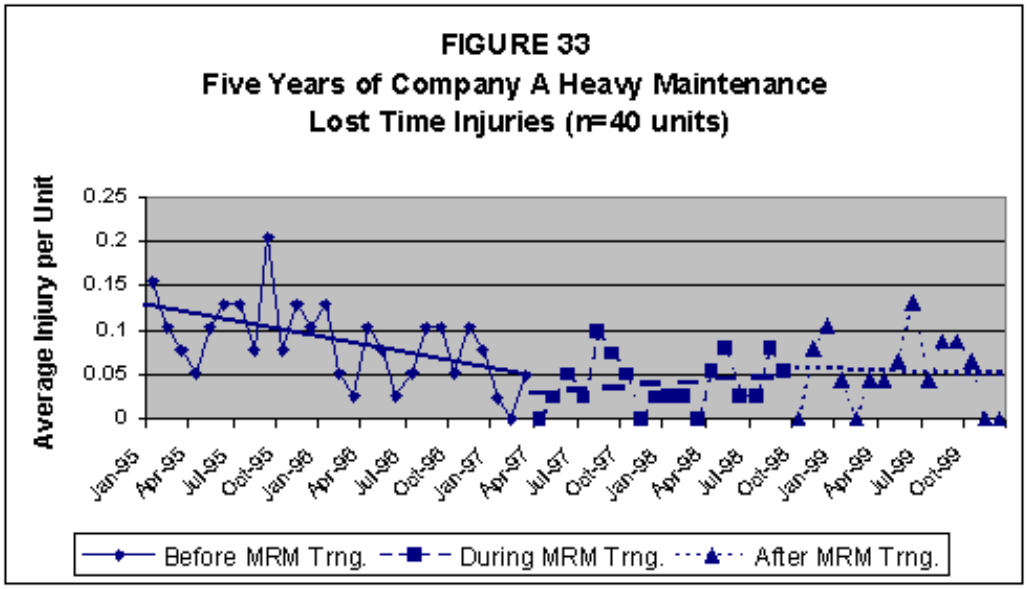

Figure 33

Figure 33 shows that the trend for Co. A lost time injuries remains low during the period of base maintenance MRM training and that the trend falls gradually in the 15 months after the training was concluded. However, because of the sharply downward trend before the training began we must question whether the lower rates during training and after are a continuation of some previous program to lessen injuries in the hangars or are the result of the MRM training.

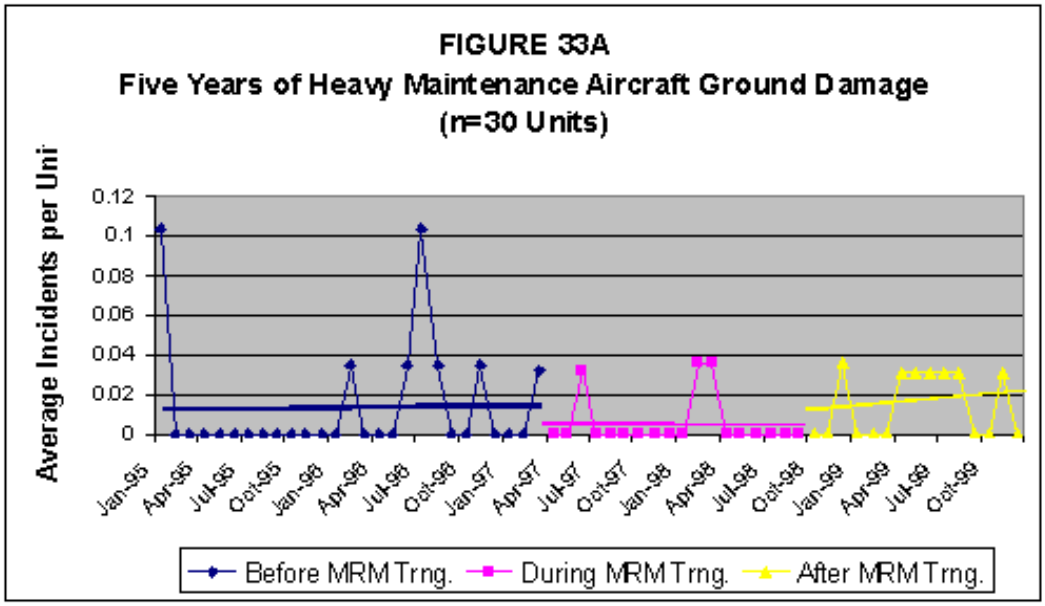

Figure 33A

Figure 33A shows a reduction in Co. A Base maintenance aircraft damage incidents during the time of their MRM training, but a slight increase in the months since. There are no data available to explain this reversion.

Company D Error Performance Trends

Company D have reports of aircraft ground damage data since 1996 and paperwork/document errors since 1992. The following figures (34-36) are arranged similar to figures 31, 32, and 33 for company A. Unlike the data for company A, company D's ground damage (GD) data are weighted for flight departures (a measure of station activity and work volume).

Line maintenance Ground Damage. Figure 34 shows four years (1996-1999) of aircraft damage incidents charged to city D-2 maintenance, compared with all line stations (n=45). The overall pattern of ground damage incidents for all company D line stations in the system remains steady with a flat trend line during this four year period. The results for city D-2, however, show a slightly increasing incident rate before the MRM training began. That trend reverses following the phase 1 training and it continues downward for 16 months after the second training phase concluded. After the first initial months after the MRM program began there, city 2's GD incidents increase coincident with those of the larger system. This is further evidence that city 2's results are as good or better than company D's total maintenance system after their MRM program began.

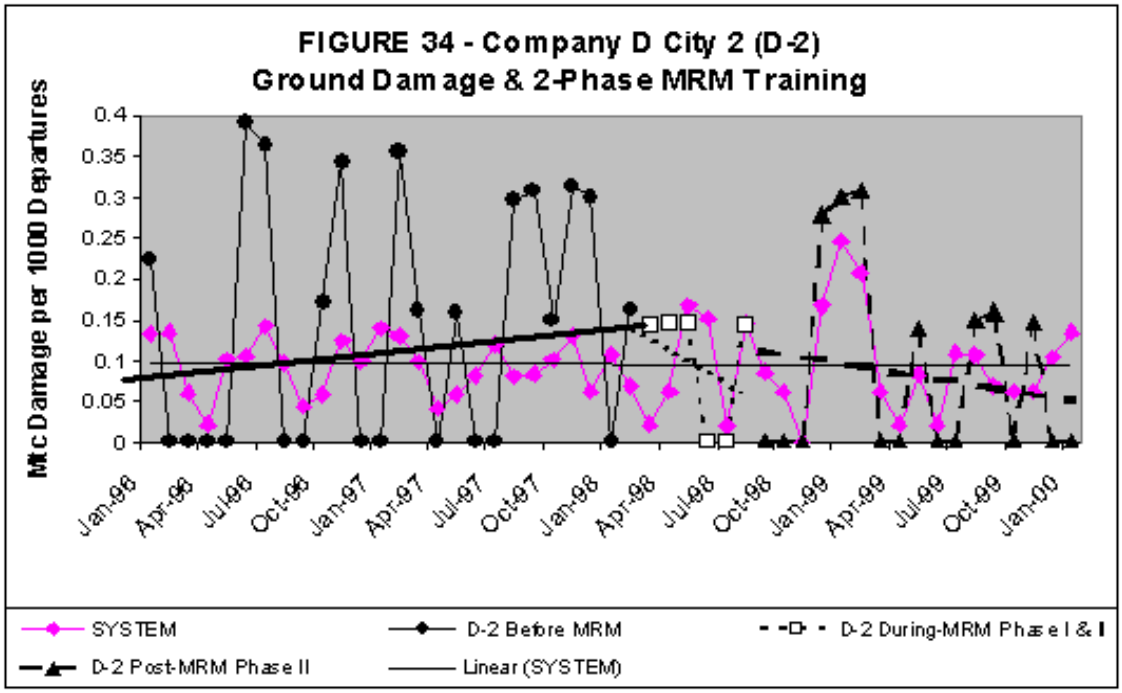

Figure 34

This improvement in safety results shown in Figure 34 is further evidence for the effect that MRM awareness instruction can have on maintenance performance. This is supported by the evidence for sustained enthusiasm to change behavior described above for Co D, Phase 2, (Figure 2). This sustained effect is reinforced in Figure 34 in the period following more than a year from the completion of the training. It seem clear that the distributed, two-phase training program developed by Company D may avoid some of the frustration and anger caused by a perceived lack of support by their managers and co-workers to improve the safety climate (Taylor, 1998).

Base maintenance ground damage. City 1 in company D is both a heavy maintenance base and a large line station. The initial ground damage results for city D-1 are portrayed in Figure 35 above. They track a similar level of incidents as the total system before the onset of the MRM training. During January to April 1998 (their period of MRM training), the city 1 month-by-month results track the same "spikes" as the total system results (representing bad weather and high traffic volume). Despite a clear downward trend in the months after the training, city 1's GD results fluctuate and there are several "spikes" above the system's total. No additional data are available to explain the increased variation in ground damage incidents for Company D's City 1 beginning in May 1998.

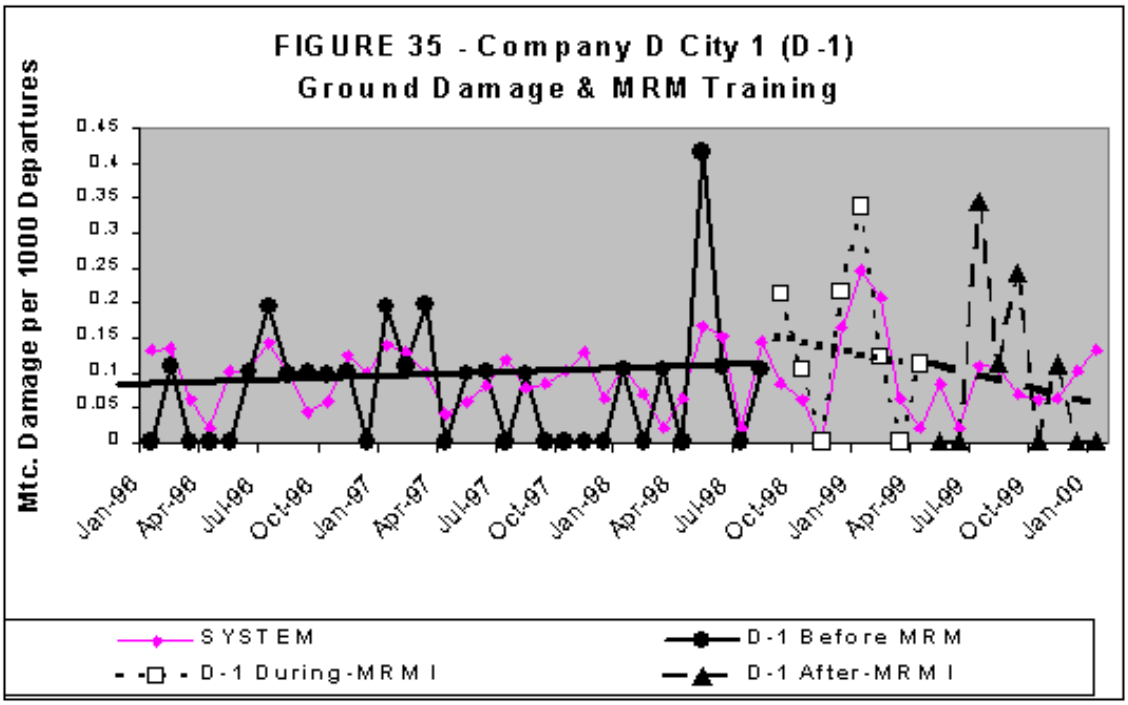

Figure 35

In a previous report (Taylor, 1998) Company D's errors for logbooks only were reported as low and stable. Since 1997 Company D's logbook error records have radically changed, consistent with substantial changes in their logbook entry system. As a result it is no longer possible to track and compare logbook errors separately from total errors.

Line maintenance document errors. For Company D, reducing maintenance paperwork errors has been a priority factor leading to improved safety since 1992 (Taylor, 1994). Barring a few lapses since 1992, company D has collected document error performance data on a monthly basis. The combined total of errors per line station per month has been analyzed through mid 1997 and reported in previous reports (*cf.*, Taylor, 1994, 1998). Company D's MRM training program includes several modules on written communication and the importance of correctly completing job cards, written turnovers, and other work documents. Figure 10 presents the per-month results of total document errors for the 36 months 1997-1999. Figure 36 plots both the line system total and the performance for city 2 (D-2) for this measure.

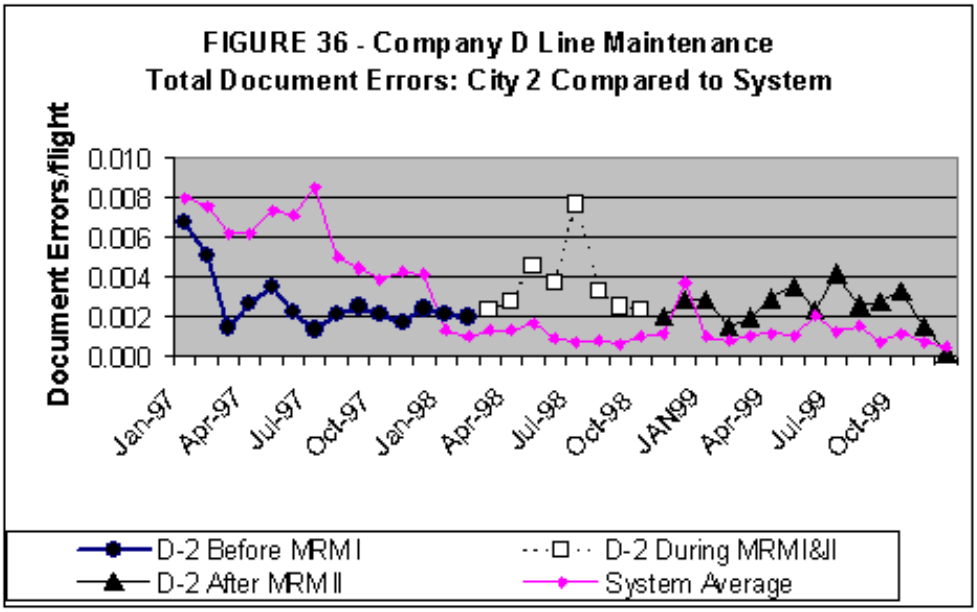

Figure 36

After mid 1997 Figure 36 clearly shows the system's rate for total document errors drops through a plateau and remains stable and low (less than 2 errors per 1,000 flights) for 1998-1999. The 1997 pre-MRM performance for city 2 is substantially below total system performance, but it fluctuates during the period of MRM training and appears to settle back into an error rate somewhat above it's immediate pre-training level. These results might be explained by a coincidentally increased diligence of the clerical auditors, but there is no record of a "crackdown" during that period.

TESTING RELATIONSHIPS BETWEEN ATTITUDES AND PERFORMANCE

A most important result is whether the attitudes and opinions affected by the MRM training are, in turn, a cause of subsequent and expected performance outcomes. Correlating the safety performance of maintenance work units with the attitudes of their members can test this effect.

The results for line and base maintenance work groups are calculated and presented separately. To accomplish this for Company A the performance results are available for a sizable number of work units over a large number of months. Attitudes and opinions of company A's work unit members are combined into the larger group's average scores in order to have commensurate units of analysis. Table 8 contains the correlations for lost time injuries (LTI) and Table 9 shows the correlations for aircraft ground damage (GD).

Company D does not have a large enough number of maintenance work units' performance data available (only two city stations at present) to be able to conduct this correlation analysis at this time.

The correlation statistics used in the present analysis. The relationships between the attitude indices and the performance measures, as presented in Tables 8 and 9 below, were calculated using the Spearman Rank Order Correlation statistic ("Rho"). The use of Rho is advisable in this case because the characteristics of small data sets (i.e., smaller than n=30) narrow the analytic power of most other statistical tests. This analysis is treated as descriptive, not predictive, thus 2-tailed tests of significance are used.

In order to simplify the analysis the 60 months of Lost Time Injury (LTI), and Aircraft Ground Damage (GD) performance data were clustered into six month groups and average scores were calculated. Thus average safety performance scores were created for the ten periods: January-June 1995, July-December 1995 and so on through December 1999. These ten performance periods were correlated with the six MRM attitude and opinion scales (*cf.*, Table 3), averaged by AMT work units, from the pre-training and post-training surveys, as well as from the 2 month and 6 month "follow-up" surveys.

The results of the two performance measures, LTI and GD, are "improving" when they decline numerically (i.e., the absence or lowering of occupational injuries and/or ground damage incidents). To keep the presentation of findings consistent all results in Tables 8 and 9 are described as negative coefficients when the correlations are in the expected direction (i.e., favorable attitudes equal better subsequent performance). Predictions about the effects of prior performance on subsequent attitudes were not made.

The previous report for 1996-1997 reported that neither the pre-training attitudes nor the immediate post-training attitudes were significantly correlated with safety performance before or after the onset of training (Taylor, 1998). Analyses of the present data confirm that lack of significance once again. The two-month and six-month follow-up surveys showed a number of interesting and expected correlations with both safety measures. With the increased data set presently available, the results for six-month follow-up survey are shown to have the strongest effect on subsequent safety performance. To simplify the present discussion of the overall correlations analysis only the correlation coefficients using the six-month follow-up survey, are presented in Tables 8 and 9 for further discussion here. Tables 8 and 9 each compare the line and base maintenance results for LTI and GD respectively. The correlations with the 2-month follow-up surveys are included in Appendix C.

Each of the tables 8a and 8b and 9a and 9b present the correlations between the six month follow-up attitudes and opinions and two measures of safety performance for line and base maintenance stations respectively. For line stations, over 75% of those surveys were completed between mid 1997 and mid 1998. Those two six month periods are shown shaded. The line maintenance unit performance data in each of the two Tables range from 30 months before the completion of most of the six-month surveys, to 18 months after their collection. For base maintenance units, over 75% of those surveys were completed during 1998. Those two six month periods of 1998 are shown shaded for the base maintenance correlations in each table. The base maintenance unit performance data in the two Tables range from three years before the completion of most of their six-month surveys, to one year after their collection. The number of work units ("n") in each column for line and base maintenance differ from n's in Table Y above, due to missing data in the survey or in the performance data for the six month period shown.

Line and Base Maintenance Relationships With Lost Time Injuries (LTI).

Line maintenance LTI results. For the four attitude scales only a few (5 of 40 possible, or 12.5%) statistically significant line maintenance correlations are found in Table 8a – and they are in both positive and negative directions and in random pattern. On the other hand, with the two opinion scales, six significant correlations following the survey are to be seen. This result means that the higher the line mechanics evaluated goal sharing and safety practice in their units six months after their MRM training, the lower the incidence of LTI in those units in the 18 months thereafter. These results suggest "if we share goals and experience a responsive safety practice, our future injuries rates will be lower." Some three periods of LTI data prior to the six month survey are also significantly correlated (for a total of 9 of 20 significant correlations possible, or 45%), and all coefficients for the two opinion scales are in the expected (negative) direction.

TABLE 8a - Correlations Between LOST TIME INJURIES (LTI) and MRM Attitudes & Opinions

Spearman Rank Order Coefficients (rho) Between Company A Attitude & Opinion Scales' 6mo Follow-up Surveys, and Lost Time Injuries (corrected for # of employees) for Line Stations Only

	Jan-June '95	*July-Dec '95*	*Jan-June '96*	*July-Dec '96*	*Jan-June '97*	*July-Nov '97*	*Jan-June '98*	*July-Dec '98*	*Jan-June '99*	*July-Dec '99*
	n= 24 units	n= 24 units	n=25 units	n=25 units	n=25 units	n=25 units	n=24 units	n=25 units	n=26 units	n=26 units
Share Command Responsibility	.25	.30	.07	.06	.18	.16	.02	.59**	-.01	.49**
Communication and Coordination	-.31	-.61**	.10	.02	.30	-.22	.04	.10	.06	.26
Manage Stress Effects	-.05	-.14	-.13	-.08	.22	-.07	-.18	.25	.02	.40*
Assertiveness	-.14	.14	.37*	.31	.28	.07	.21	.28	.27	.30
Goal Sharing	-.21	-.38*	-.31	-.30	.16	-.31	.07	-.51**	-.34*	-.46**
Safety Practice	-.33	-.22	-.38*	-.41*	.20	-.14	.05	-.44*	-.34*	-.42*

*sig .05, 1-tail
**sig .01, 1-tail
Shaded columns represent survey period

Base maintenance LTI results. Table 8b shows that only 2 of the 60 correlations between attitudes and opinions and LTI for base maintenance are statistically significant. This is 3% of the total possible, which is a smaller proportion of significant correlations than the 5% expected by chance alone. We cannot conclude that any Base maintenance attitude is consistently related to subsequent LTI.

TABLE 8b - Correlations Between LOST TIME INJURIES (LTI) and MRM Attitudes & Opinions

Spearman Rank Order Coefficients (rho) Between Company A Attitude & Opinion Scales' 6mo Follow-up Surveys, and Lost Time Injuries (corrected for # of employees) for Base maintenance Only

	Jan-June '95	*July-Dec '95*	*Jan-June '96*	*July-Dec '96*	*Jan-June '97*	*July-Nov '97*	*Jan-June '98*	*July-Dec '98*	*Jan-June '99*	*July-Dec '99*
	n=23 units	n=23 units	n=24 units	n=24 units	n=25 units	n=25 units	n=22 units	n=22 units	n=27 units	n=27 units
Share Command Responsibility	.10	**-.39***	.14	-.13	-.09	.01	.33	.12	-.13	.19

Communication and Coordination	.22	-.34	-.14	.11	-.02	.25	-.14	-.29	.07	.02
Manage Stress Effects	.22	-.15	-.08	.01	-.27	-.31	-.09	-.15	-.29	**.39***
Assertiveness	.21	-.34	.00	-.31	-.16	.15	.11	.04	-.11	-.22
Goal Sharing	.23	.02	-.10	-.30	.02	.01	-.28	.08	.08	-.19
Safety Practice	-.01	-.06	.00	-.26	-.01	.03	.02	.12	.01	.11

*sig .05, 1-tail
**sig .01, 1-tail
Shaded columns represent survey period

Tables 9a and 9b present the correlations between the six month follow-up attitudes and opinions and performance for line and base maintenance stations .

Line and Base Maintenance Relationships With Aircraft Ground Damage (GD).

Line maintenance GD results. Table 9a shows that for the four attitude scales, eight correlations with ground damage (8 of 40 possible, or 20%) were statistically significant, but three of them were not in the expected direction. Specifically the two scales dealing with assertiveness and autonomy account for those positive correlations. The stress management scale however shows three periods in which lower GD incidents are correlated with it – two of which precede the survey measurement and one is coincident with it.

TABLE 9a - Correlations Between AIRCRAFT GROUND DAMAGE (GD) and MRM Attitudes

Spearman Rank Order Coefficients (rho) Between Company A Attitude Scales' 6mo Follow-up Surveys, and Aircraft Ground Damage (Corrected for # of Employees) for Line Stations Only

	Jan-June '95	*July-Dec '95*	*Jan-June '96*	*July-Dec '96*	*Jan-June '97*	*July-Nov '97*	*Jan-June '98*	*July-Dec '98*	*Jan-June '99*	*July-Dec '99*
	n=25 units	**n=25 units**	**n=25 units**	**n=25 units**	**n-24 units**	**n=24 units**	**n=25 units**	**n=25 units**	**n=25 units**	**n=25 units**
Share Command Responsibility	**.20**	**.01**	**.42***	**.11**	**.25**	**.47***	**.19**	**.26**	**.12**	**.08**
Communication and Coordination	**-.30**	**-.10**	**.08**	**-.26**	**-.20**	**-.52****	**-.05**	**-.21**	**-.22**	**-.03**
Manage Stress Effects	**.00**	**-.45***	**-.42***	**.05**	**-.24**	**-.46***	**-.20**	**-.26**	**.05**	**-.02**
Assertiveness	**.19**	**.02**	**.20**	**.18**	**.37***	**.41***	**.31**	**.16**	**.26**	**.08**

Goal Sharing	**-.22**	**-.02**	**.28**	**-.07**	**-.17**	**-.37***	**-.08**	**-.21**	**-.04**	**.13**
Safety Practice	**-.37***	**-.07**	**.33**	**-.13**	**-.19**	**-.22**	**-.14**	**-.13**	**.00**	**-.04**

***sig .05, 1-tail**
****sig .01, 1-tail**
Shaded columns represent survey period

Base maintenance GD results. It will be noted in Table 9b that correlation coefficients are absent for two periods in the base maintenance portion of table 9 (the last half years of both 1995 and 1997). This lack of correlations involving damage rate for base maintenance is explained by (the admittedly happy circumstance of) the zero damage rate (and thus zero variance for calculating correlations) for all base maintenance during those two periods. The four significant correlations (of 48 possible, or 8%) in that base maintenance table are all in the expected direction and are all with the "communication" attitude scale. Furthermore, one of these correlations is coincident with the base maintenance six-month survey and two of them are subsequent to it. This pattern is strong and regular – it suggests that the more value that is placed on meetings and briefings and other communication, six months after training, the better is ground damage performance in the following year.

TABLE 9b - Correlations Between AIRCRAFT GROUND DAMAGE (GD) and MRM Attitudes

	Jan-June '95	*July-Dec '95*	*Jan-June '96*	*July-Dec '96*	*Jan-June '97*	*July-Nov '97*	*Jan-June '98*	*July-Dec '98*	*Jan-June '99*	*July-Dec '99*
	n=25 units	n=25 units	n=25 units	n=25 units	n=25 units	n=25 units	n=22 units	n=22 units	n=25 units	n=25 units
Share Command Responsibility	**-.07**	**-**	**.05**	**-.08**	**.05**	**-**	**.00**	**.08**	**-.07**	**-.07**
Communication and Coordination	**-.50****	**-**	**.10**	**-.11**	**.10**	**-**	**-.44***	**.10**	**-.49****	**-.49****
Manage Stress Effects	**-.05**	**-**	**.05**	**-.06**	**.05**	**-**	**-.13**	**.13**	**-.05**	**-.05**
Assertiveness	**.26**	**-**	**.09**	**.09**	**.09**	**-**	**.21**	**.10**	**.26**	**.26**
Goal Sharing	**-.18**	**-**	**.04**	**.07**	**.04**	**-**	**-.07**	**-.24**	**-.18**	**-.18**
Safety Practice	**.00**	**-**	**-.13**	**-.01**	**-.13**	**-**	**.13**	**-.11**	**-.01**	**-.01**

***sig .05, 1-tail**
****sig .01, 1-tail**
Shaded columns represent survey period

Discussion

Correlations with Performance Changes. Company A's line maintenance MRM attitudes are no longer seen in these data as major or consistent correlates of subsequent line maintenance safety performance. Where the line mechanics' survey results following MRM training are correlated with enhanced performance it is with their positive evaluation (opinions) of management practices. In particular, positive opinions of present safety practice and goal setting practices, some six months after training, are associated with subsequent improvements in their lost time injury rates. Their attitudes toward stress management – formerly found to be correlated with lower subsequent ground damage incidents – now lag behind the positive safety performance and do not precede or predict lower ground damage in 1998-1999.

For Base maintenance, attitudes toward communication are correlated with subsequent ground damage. This suggests the positive effect of management communication evident in some Company A heavy maintenance bases since 1997. No similar pattern of consistent correlations is seen for lost time injuries for Company A's base maintenance sample.

The Role of Stress Management in Safety Improvement

Understanding and acceptance of stress management is clearly a bona fide result of MRM training for all the companies measured here (*cf.*, Figures 4 and 7). Stress management is a topic MRM training programs all include and respondent attitude changes show that it "takes." In Company A the pre-training and post-training training comparisons clearly show the MRM program's statistically significant impact on feelings about managing stress (Figure 7). Those heightened post-training feelings fall back very little in the months following training and the differences are statistically non-significant (Figures 7 and 10). Despite this widespread effect on attitude change resulting from the training, few Company A AMT respondents specifically state that they will subsequently apply the lessons learned about stress management (Table 4; Figures 16, 17, 26).

The 1997 correlations of GD with six month surveys (Table 9) show convergence with the two month survey data reported in the last milestone report (Taylor, 1998). Whereas the earlier results showed the attitudes were coincident or prior to the safety performance, the present results show those 1997 performance results are correlated with subsequent attitudes. The increase in appreciation of stress management six months after training (Table 9) followed low rates of aircraft damage prior to that survey. Stress management is primarily a passive coping activity and its improvement following the training and its relationship to safety performance improvements is entirely consistent with company A's MRM purpose. For base maintenance nothing measured in the six month follow-up MRM/TOQ seems associated with LTI, but favorable attitudes toward communication are related to subsequent improvement in the incidence of aircraft damage.

Stress management is an activity that maintenance personnel can do by themselves and which does not require the involvement of others (although cooperation may benefit all parties in this regard). The training helps Company A mechanics and their Leads improve their individual attitudes about stress and its management.

Despite finding early success in applying stress management awareness to subsequent safety (Taylor, 1998); now, with nearly two years experience following MRM training in Company A, any quantitative effect of stress management attitudes on safety improvements is nearly non existent. Although Table 9 shows that there does seem to be some effect of reduced ground damage incidents on coincident and subsequent feeling about stress management, but there is no correlative evidence that the effect goes the other way around.

The Role of Safety Awareness in Safety Improvement

Information reported in the open ended questions and field interviews shows that many respondents intend to – and actually do – perform more carefully and self-consciously in the months following MRM training. Table 4 also shows being "more careful" as a result of the training is frequently mentioned as intention or accomplished fact in the maintenance line stations we observed. As it does so, "being careful" appears also to lead to improved safety (Table 8). But this continued emphasis on working alone may be placing AMTs in the position of not knowing whether or how much the MRM program is working, or whether other people actually value the lessons of the training as they did. This uncertainty may lead to frustration.

An unplanned liability of the individual change model

It is ironic -- given the apparent success of Company A's MRM program as expressed in long-term safety outcomes and the direct relationships with several of the survey measures – that mechanic's enthusiasm for the program turned from positive to negative. The questionnaire and interview data above (Table 4; Figures 2 and 3) and earlier (Taylor, 1998) examining the attitudes and opinions of line maintenance employees in the months following their MRM training show the apparent frustration and anger these individuals voiced. They expected more support by their managers and co-workers in fulfilling the promise of the MRM program to improve communication and collaboration (Taylor, 1998). Subsequent interviews and observations in one of Company A's repair hangars did not confirm this "backlash" exists – at least in that part of heavy maintenance.

This individual-based awareness training, with its emphasis on building individual's coping skills, appears to give AMTs little subsequent information about whether or how much the MRM program is working, or whether other people value the lessons of the training like they themselves do. Months after the training many AMTs reported still being careful, fighting complacency, and managing their own stress levels. But many also didn't think the MRM program would be very useful in the future (Taylor & Christensen, 1998, pp. 152-160). Many said they didn't know or couldn't tell if others were using the lessons learned from the training – they rarely talked about MRM informally and were typically not encouraged to do so by their leaders.

The Role of Communication in Safety Improvement

Open communication is an idea at the heart of MRM programs and of the Crew Resource Management Training Programs which preceded and inspired it. Due to both the individual characteristics of those who choose to become airline mechanics (Taylor & Christensen, 1998, Chapter 2), and the occupational culture they are further shaped by (Taylor, 1999), poor communication has proven to be a difficult behavior to change and improve.

Some airline companies have long held pre-shift meeting and briefings in maintenance (Taylor, 1991) to communicate mission and work goals, as well as progress and changes in work flow and priorities. Success follows where such meetings commonly and consistently emphasize some work goal – whether it be error reduction (Taylor & Christensen, 1998, pp.112-113), or increased production (Taylor & Christensen, 1998, pp.135-137).

In this regard, Table 4 above shows intentions and actual reports for the Heavy Maintenance station which included communication and interaction in a relatively high proportion when compared with the other (Line Maintenance) sites. The shift briefings and management climate at that Heavy Base could well explain the high correlations between favorable attitudes toward Communication & Coordination and subsequent improvements in aircraft ground damage (Table 9).

MRM RETURN ON INVESTMENT (ROI)[3]

Maintenance Resource Management programs have various effects on operations and safety in the aviation industry. There is ample demonstration of the effect of employee awareness and enthusiasm for safety issues on reduction in personal injuries and aircraft ground damage incidents (Taylor & Patankar, 2000). Behavior training programs at some corporate aviation departments have shown an increase in the pilot-technician communication, improvements in technical support from third-party service providers, and better cooperation with the local FSDO (Patankar & Taylor, 1999). Although no single return-on-investment (ROI) formula will apply to all possible configurations of MRM programs, a general model of calculating the ROI has been developed and presented (Taylor, 2000a). This model uses the correlations between safety performance indicators and the MRM/TOQ measures following classroom instruction to account for realistic contributions to success while giving due consideration to other safety efforts in progress simultaneously. Given that MRM programs could be configured to achieve a variety of results, Patankar and& Taylor (2000) recommend that the MRM program managers identify specific targets during the planning stages so that they will have more realistic means of evaluating the effectively of their programs.

Absence of accidents does not equate to presence of safety, at least not safety by design. However, the cost of safety programs is often regarded as an expense without specific measurable returns. When maintenance managers have limited financial or technical resources and are under strong operating pressures, there may be pressures to reduce safety programs. Under these circumstances or when airlines have conducted large-scale safety training, but not measured its financial benefits, the airline management may demand return-on-investment analyses of safety programs. Patankar and Taylor (2000) have introduced the concept of "targeted MRM programs." This concept acknowledges that MRM training could be tailored in several different ways and therefore would result in a variety of benefits. If the goals of the training were identified and set as targets, appropriate ROI measurement techniques could be applied.

Return On Investment. Profits are derived from earnings. The rate at which earnings grow is a function of the company's return on investment: net income as a percentage of investment costs. Although ROI competes with other financial indicators (e.g., return on equity, return on assets) for an executive's attention, it is the longest lived and most robust of the evaluation tools for management decisions. Evaluating the benefits of training has been long admired, but little practiced (Kirkpatrick, 1975; Phillips, 1997). Evaluating the effects of specific change efforts like MRM is likewise underdeveloped. Training and other MRM interventions, especially for safety improvement, are rarely treated as investments and are usually just considered necessary costs of doing business or worse yet, expendable activities. Little wonder then that converting MRM benefits into a standardized and comparable format, such as "return on investment," is so little in evidence and has only lately been discussed and understood within the training and organizational effectiveness community.

Costs and benefits. Assessment of costs and benefit mark an important step in measuring ROI . It is essential that true and accurate costs of any "organization effectiveness" (OE) intervention (whether training, or structure/process, or a combination) be specified and calculated. Likewise, calculating the cash benefits resulting from an OE intervention is important. Rules for listing, collecting and calculating those costs and benefits can be found elsewhere (*cf.*, Phillips, 1997). Where cost or benefit data are available alone, they are too often presented as evidence that an intervention was a success – "it came in below budget" (low cost) or "it saved (or produced) a substantial amount of cash" (high cash benefit). Most managers and executives familiar with financial analysis would consider such direct statements to be without reference and therefore without much meaningful information for decision making. When both cost and benefit data are available they are also, and all too frequently, combined by placing them in direct comparison with one another – in the familiar "cost-benefit differences" or "cost-benefit ratios." These combinations cannot correspond with other efforts to justify the economic success of an intervention, nor are they a standardized measure to be understood in implied comparison with other results. These benefit (-) cost "differences," or benefit (÷) cost "ratios," cannot be considered effective outcome measures by themselves because the actual practical effect may be magnified or otherwise skewed by the absolute size of the effort and its budget.

In conventional terms, a company's "earnings" are its "income" minus its "expenses" for some fixed period of time. Given that definition of "earnings," ROI is traditionally reported as "earnings" divided by "investment." To further standardize the ROI expression, the resulting quotient is multiplied by 100 to convert it to a percent expression.

$$ROI = \left(\frac{Earnings}{Investment} \right) \times 100$$

Equation 1

As an illustration, ROI of "25%" means that the investment costs are recovered – and an additional 25% of the cost amount is reported as earnings. This definition of ROI is in contrast to the direct cost-benefit ratio, since "earnings" are not a direct equivalent to "benefits." The concept of "benefits" is more similar to the "income" in traditional ROI calculations. In light of this, the MRM equivalent to earnings would be benefits minus costs, or "net program benefits."[4] Thus for the calculation of MRM ROI, the numerator of the equation is the net program benefits, or "Net MRM Benefits." The denominator, "program costs," likewise compares to "investment" in traditional ROI.

Using this formula, ROI calculations for OE are thus commensurate with ROI calculations for more typical applications of the concept (i.e., efforts to increase productivity) because they are calculated to the same basic formula. Because ROI calculations are expressed as percentages they are standardized to the same scale. Executives and other policy makers who are accustomed to thinking about ROI for earnings are likely to dismiss cost and/or benefit statements in favor of the same data transformed through the ROI formula. This is because they are better able to understand the implications of ROI for an MRM intervention and will be less likely to discount the results.

Obstacles to the use of MRM ROI. There are two major obstacles to overcome in developing a realistic and appropriate model of ROI for safety. The first obstacle is lack of experience. The second obstacle is difficulty in assessing causality. The new ROI model to follow addresses both of them.

1) Lack of experience. The industry's use of broad OE tools for safety improvement-- which includes training, among others – is just beginning. The application of appropriate ROI models to this larger class of organizational intervention has not yet been attempted. The new model is shown below, followed by an example using data collected from several real MRM interventions. This illustration of the new ROI model for MRM provides the evidence (as well as the process) for any company wishing to assess the ROI as part of a planned intervention to improve "and prove" organizational effectiveness.

2) Assessing causality. In most airline companies, everyone is focused on safety. This means that there are usually many initiatives to improve a particular safety outcome and if improvement is achieved, many will want to take credit for it. The use of bivariate correlations between MRM outcomes (new knowledge, attitudes and behaviors) and subsequent safety results provides a way to conservatively estimate the degree of impact MRM has on safety, as well as providing a numerical value ("coefficient of determination," *cf.* Taylor, 2000a) to use as a "causal operator" term in a new ROI equation.

A New Model of ROI In MRM

A new model of ROI for aviation safety interventions builds upon the traditional model (equation 1, above). Not only does this new model accept training interventions; it also uses measures of changes in organization structure and work process as causes for behavioral improvements. This "MRM ROI" model's features build on the annualized percentage ratio of net program benefits to program costs – the ROI formula familiar to operating managers and their financial counterparts. But it also includes a novel component designed to account for the degree of effect the targeted MRM intervention has had on net program benefits (the ROI formula's numerator).

$$MRM\ ROI = \left(\frac{[Net MRM\ Benefits] \times Causal Operator}{MRM\ Costs} \right) \times 100$$

Equation 2

Equation 2 is similar to equation 1 above, with the addition of the "causal operator" term as a multiplier of the net program benefits in the numerator. This "causal operator" term represents the variance explained by the prior variable (the human MRM results of attitudes, knowledge and/or behaviors) in the subsequent safety outcomes. In estimating cause-effect relations between two variables separated in time, the coefficient of determination was defined as measuring the variability in the later variable that is accounted for by the prior one. Thus this "causal operator" can act as a quantitative measure for the contribution of MRM to safety in a given period of time.

The effect of this modification to the traditional ROI equation is to reduce the size of net program benefit by a positive factor between zero and one and thus change the benefit outcome downward to a level that acknowledges the residual as potential effects on that benefit belonging to other interventions.

Return on investment (ROI) is a simple but powerful idea. It can be applied to the evaluation of organizational effectiveness initiatives including management awareness instruction and communication training. The method for measuring ROI requires quantification of several variables, but appropriate measurements are usually available or can be readily developed from past work by others. The usefulness of ROI calculation has been dramatically illustrated in the two examples described, both to question the financial viability of certain assumed "benefits" as well as to justify a program in standardized terms when the net program benefits are high.

AVIATION MAINTENANCE AND NATIONAL CULTURE[5]

The effect of the larger environment on organizational performance is beginning to be recognized in the commercial aviation industry. Both systems thinking (Maurino, et. al, 1995), and national culture (Johnston, 1993; Helmreich & Merritt, 1998), are becoming ideas in "good currency" in addressing the twin issues of risk management and flight safety. Both whole system thinking and examination of national culture are ways of addressing the "people side" of the safety equation, now that the technical side has become so reliable. Today, the challenge of improving safety can be seen as improvements in managing the "human factor". Understanding culture is part of understanding human behavior. Company A, with line maintenance operations throughout the Pacific and South America, collected MRM/TOQ data from a wide variety of national cultures. From these Company A worldwide data (both pre- and post-training) some effects of national occupational culture on maintenance personnel can be examined.

According to organizational anthropologist Geert Hofstede, the culture of a country is not a combination of properties of the "average citizen," nor a "modal personality;" but is a set of likely reactions of citizens with a common mental programming (Hofstede 1991, p. 112). The full set of reactions, he continues, need not be found within the same persons, but only statistically more often in the same society. Hofstede thus cautions against confusing an individual's personality with his national culture.

National culture and the dimension of "authority." Managing the human factor includes understanding the expectations of those to be led, as well as understanding the practices of the leader. Hofstede (1984, 1991) defines "high power distance" cultures such as China and many Latin American countries as stressing the absolute authority of leaders. Following the example set by Helmreich & Merritt (1998), who applied Hofstede's ideas to flight operations; the "Command Responsibility" and "Assertiveness" scales, from the MRM/TOQ, were used to measure Hofstede's "power distance" concept. The low ends of both scales, "command responsibility" and "assertiveness," represent high "power-distance." That is, respondents who place lower value on subordinate employees sometimes taking command, or for employees speaking up in possible criticism of one another, are included in Hofstede's high power-distance category. For the present purposes, the MRM/TOQ "command responsibility" and "assertiveness" scales are used to measure Hofstede's "power-distance" dimension.

The cultural dimension of "collectivism." Understanding the human factor also includes understanding the individual in the group setting. Johnston (1993) has observed that Hofstede's data reveal a strong, inverse correlation between the "power distance" and the "individualist/collectivist" dimension – countries with large power distances tend to be collectivist, and vice versa. In Hofstede's scheme, "collectivism" is characterized by interpersonal interdependence with a priority on group goals. The MRM/TOQ includes two individual questions on communication and coordination. One deals with the start-of-the-shift meeting's importance for safety and coordination, and the other asks about debriefing and critique as an important part of team coordination. Following Helmreich & Merritt, these two questions will be used to gauge the degree of Hofstede's "collectivism." The MRM/TOQ also contains a set of questions dealing with group goal sharing (*cf.*, Taylor & Christensen, 1998, pp. 135-137). These items, combined into a "goal sharing" scale, will also be used as an additional, reliable, measure of "collectivism" to help examine differences in national culture among airline mechanics.

Transforming the survey data for neutralizing the effects of culturally based response bias. Helmreich & Merritt, who reviewed the literature on cultural bias in survey methodology (1998, p.238), report that Japanese and Korean respondents tend to overuse the mid-range of a scale and that South Americans appear to overuse the extremes of the scale. To draw culturally neutral conclusions from their index mean scores, Helmreich & Merritt collapsed the 'slightly agree' and 'strongly agree' categories into one category, and did the same with the 'slightly disagree' and strongly disagree' categories. The maintenance mean score results reported here for the 'Command Responsibility,' 'Assertiveness,' and 'Goal Sharing' multi-item scales are based on that data transformation, as are the two separate communication questions.

Company A used the pre-training MRM/TOQ to measure some 2,350 line maintenance personnel in 48 cities throughout the world, prior to any effect of their MRM program. Although MRM/TOQ was administered at 32 cities in the continental U.S. and 16 cities in the rest of the world, only those cities with more than five respondents (n>5) were included in the sample reported here. That smaller number of sites includes some 1,800 maintenance workers in 25 cities. About 1,600 of those maintenance personnel were employed in 15 stations in the U.S. and the remainder were local employees at line stations in 10 foreign cities in Asia and Latin America.

Table 10. CULTURAL REGIONS, NUMBERS OF RESPONDENTS, AND CONSTITUENT LINE STATIONS

U.S. Regions	**Foreign Regions**
East Coast (n=417)	**SE Asia (n=95)**
Boston	**Singapore**
New York – Kennedy	**Bangkok**
New York – La Guardia	**Taipei**
Washington – National	**Hong Kong**
Washington – Dulles	**Delhi**
Miami	
Midwest (n=735)	**East Asia (n=55)**
Minneapolis	**Tokyo**
Chicago	**Osaka**
Indianapolis	**Seoul**
Denver	
West Coast (n=466)	**Latin America (n=26)**
Seattle	**Rio de Janeiro**
Portland	**Sao Paulo**
San Francisco	
Los Angeles	

The Cultural Regions. Company A line maintenance respondents are divided into three foreign regions or cultural areas and three U.S. regional areas. The six regions and their constituent city stations are shown in Table 10 . Taiwan and India are included with SE Asia in accordance with Hofstede's analysis and discussion of these two locations (1984; pp. 167, 216). East Asia region includes the line stations in Tokyo, Osaka, and Seoul. Latin America is represented by two locations in Brazil (Rio de Janeiro, and Sao Paulo).

"Power-Distance" and Airline Mechanics

Hofstede uses the term “power distance” to represent a measure of how a national culture handles the fact that people are unequal. He specifically defines power distance as “the extent to which the less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally.” Small power distance countries have limited dependence of subordinates on bosses, while large power distance countries have larger dependence, and deference, toward bosses. We assumed, for the measures used here, that Hofstede’s “power-distance” dimension is an inverse function of the MRM/TOQ “command responsibility” scale– that is, as power-distance increases, command responsibility decreases. This function is also presumed for the MRM/TOQ “assertiveness” scale – higher power distance equals lower assertiveness. Figure 37 shows the pre-training survey results, by region, for the command responsibility and the assertiveness scales. The overall test of differences among regions, the Multivariate “F” Test, was statistically significant for both of the scales measuring power distance (F for Command Responsibility = 20.55, df = 5, p>.000; F for Assertive-ness = 17.48, df = 5, p>.000).

As Figure 37 clearly shows, the higher scores for the U.S. regions and lower scores for the Asian regions are consistent with the expected power distance positions. Although the Latin American region scores lower on power distance than the Asian regions, it is not statistically different from either the U.S. regions or the Asian regions. All six mean score difference comparisons between the lesser power-distance for each of the three U.S. Regions and the greater power-distance for each of the two Asian regions prove statistically significant for both scales (.000<p<.03).

The power distance for East Asia, as measured by mean differences in assertiveness, was found to be significantly greater than the already high power-distance of the SE Asia region (p<.000) -- while the command responsibility difference between the two Asian regions is not large enough to reach significance. What these findings imply is that the even greater power-distance of mechanics in Japan and Korea is evidenced by their being even less willing than their counterparts in Singapore, Thailand, Hong Kong, Taiwan and India to speak up when it may cause conflict or disagreement with others.

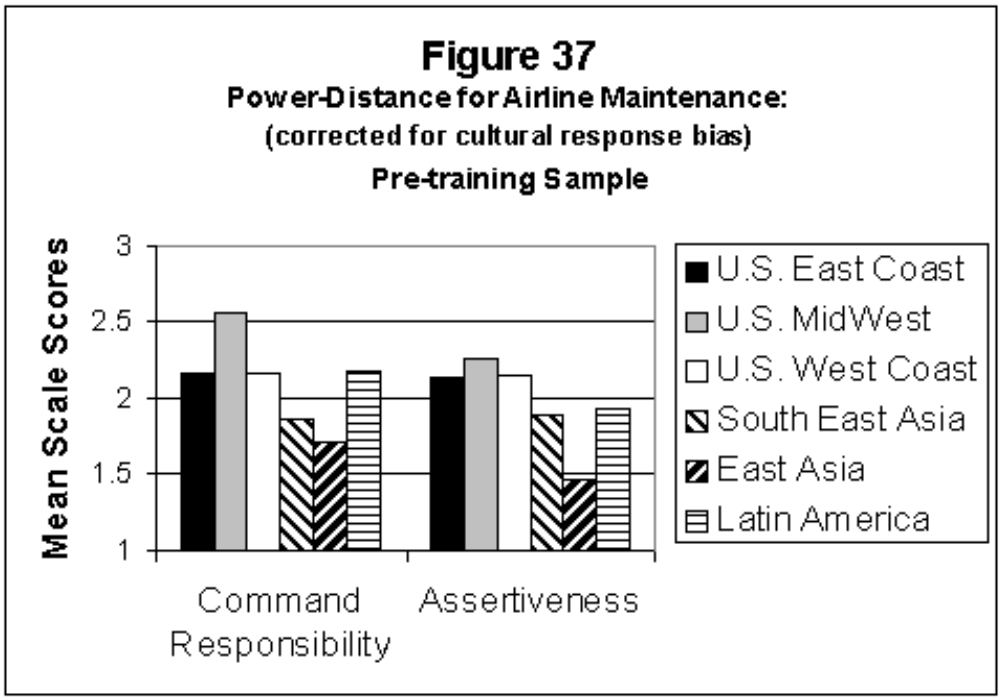

Figure 37

Individualism and Airline Mechanics

Figures 38 and 39 compare the different national regions of mechanics on two questions. One question is worded, “Start-of-the-shift team meetings are important for safety and for effective team management.” A second question states, “A debriefing and critique of procedures and decisions after the completion of each major task is an important part of developing and maintaining effective team coordination.”

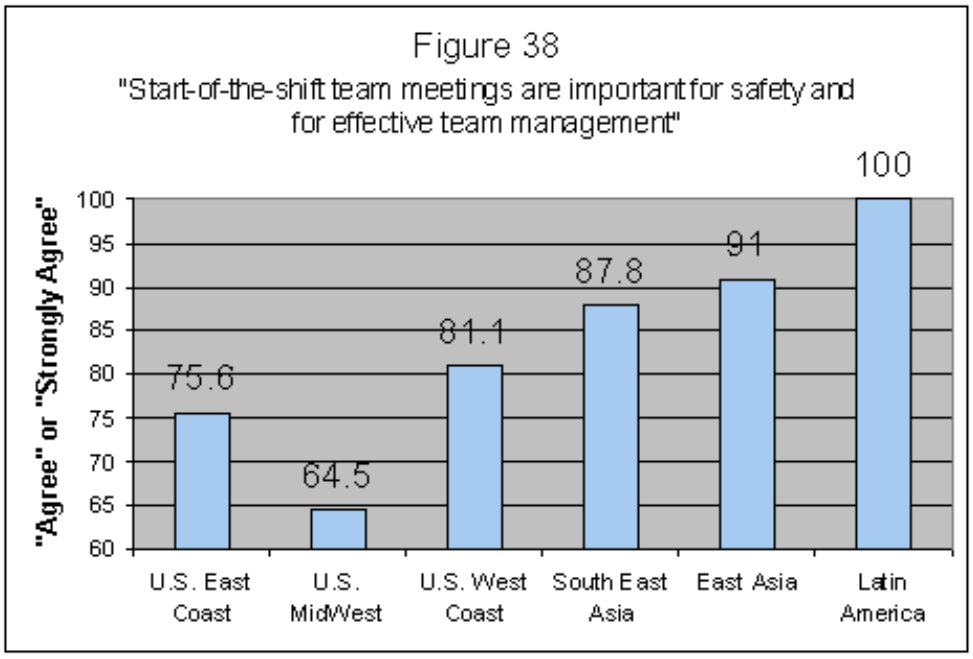

Figure 38

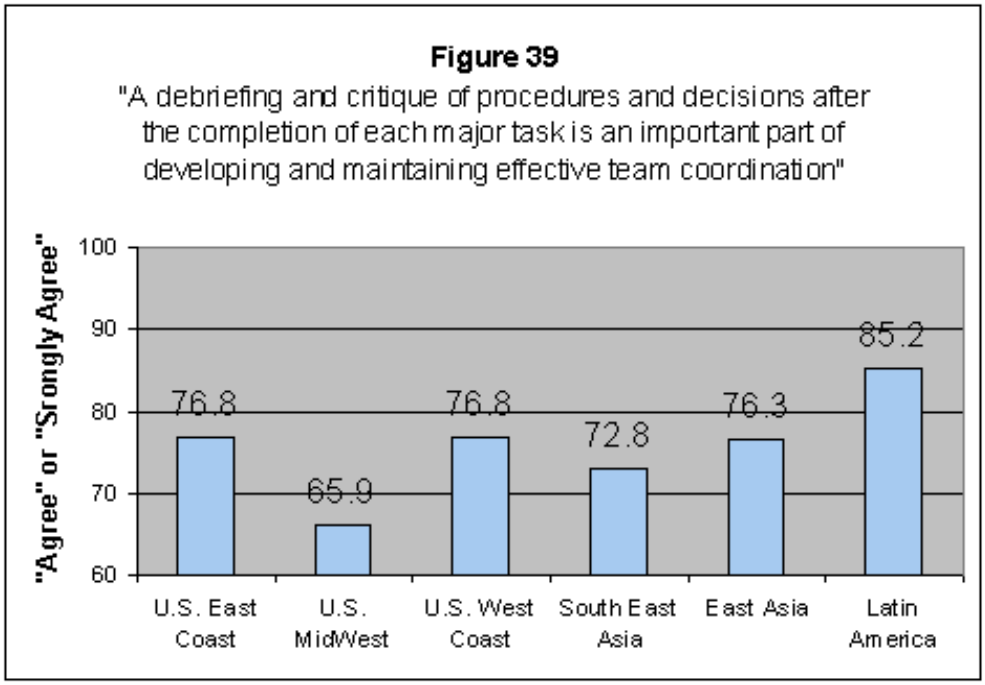

Figure 39

Response to question about start-of-the-shift team meetings importance for safety and coordination

Figure 38 shows, first of all, that in all six regions nearly two-thirds of all respondents say they agree that holding shift meetings are important. But Figure 38 also reveals substantial differences among them in the additional strength of that agreement. It is evidence that mechanics are less collectivist than pilots. Results in Figure 38 are substantially different from the pilots' samples using these questions (Helmreich & Merritt, pp. 67-77) where all responses exceeded 85% agreement.

The differences among the mechanics' in different cultures are significant when tested as expected proportions (X^2=16.7, df=5, p<.005). Furthermore, while Southeast Asia, East Asia, and Latin America all have over 85% agreement to this value of crew meetings at the start of the workday, their U.S. counterparts do not. The mean differences among the six regions are significant overall as well (F=22.3, df=5, p<.000). The mean scores for both East Asia and Latin America are significantly larger than all three U.S. regions (.000<p<.001). Although Southeast Asia is significantly below Latin America (p<.01) there is no significant difference between SE Asia and East Asia. The foreign mechanics are all more collectivist than those in the U.S. Also noteworthy is that the mean score for U.S. Midwest region is significantly smaller than either of the other two U.S. regions (4.06>3.76<4.24; p<.000).

Response to question on debriefing and critique as an important part of team coordination

Although an overwhelming majority of all respondents agree with this value, Figure 39 shows both diversity of agreement and continued differences among the U.S. regions. A test for expected agreement for this item shows statistically significant differences among the regions displayed in Figure 39 (X^2=21.6, df=5, p<.001). A test for differences among the regions' mean scores for this item is also significant (F=7.2, df=5, p<.000). The mean scores for both the U.S. East and West Coast regions (with mean scores of 4.12 and 4.13, respectively) are significantly larger (p<.000) than the U.S. Midwest region's mean score (3.86 on a 5-point scale). Except for the East Asian mean score (4.27) being found to be significantly larger than the U.S. Midwest region, no other significant differences were found for this item. Despite these overall differences, there is less effect between the U.S. and foreign regions than was found for meetings at the start of the work shift (Figure 38). Indeed there is much less diversity among the mechanics' scores on this item.

Mechanics are more uniformly individualistic than pilots. Compared with Helmreich & Merritt's results (Helmreich and Merritt, 1998, p.78) all mechanic's mean scores are substantially larger than the U.S. pilots and substantially smaller than the pilots from Japan and Brazil. Mechanics, it appears, are less diverse overall than the pilot sample and their mean scores on this debriefing item are lower (less collectivist) than two-thirds of the countries in the pilot sample, including all except the Western European, U.S. and Anglo pilots. Both of the items measuring "collectivist" values reveal that mechanics in the U.S. are more individualist than their counterparts in three foreign regions. Compared with the pilots sampled by Helmreich and Merritt, the U.S. mechanics (and many of the foreign mechanics, especially in Figure 39) reported here show lower agreement to the value of work-related crew meetings and briefings .

Collectivism As Measured By The Goal-Sharing Scale

Following Hofstede's definition of collectivism, the MRM/TOQ Goal Setting and Sharing scale's scores should increase as collectivism (or a value for group goals) increases. In the pre-training MRM/TOQ survey, these items form an index of purposeful, collective behavior, prior to human factors concepts being presented to the respondents.

Figure 40 presents the mean scores for each of six regions on the transformed goal sharing scale. The overall test of differences among regions, the Multivariate "F" Test, was statistically significant for the goal sharing scale (F = 10.77, df = 5, p>.000).

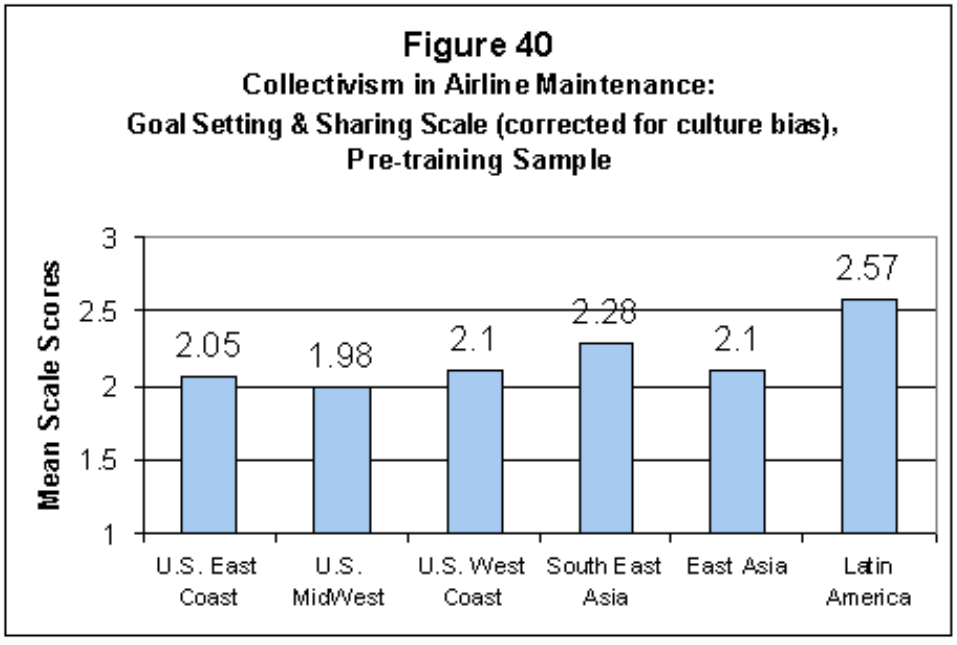

Figure 40

Multiple comparison tests of mean differences reveal that both South East Asia and Latin America are significantly higher on the goal sharing scale than all three U.S. regions ($p<.03$). The US mechanics are more individualistic and the SE Asian and South American mechanics are more collectivist. These differences are confirmed by Hofstede's findings of differences among national cultures (Hofstede, 1984, p. 159). Those original findings graphically depict Japan (a major component of the 'East Asia' region in the present analysis) lying closer to the U.S. on the 'Individualism' dimension, while Hong Kong, Singapore, Thailand and Taiwan (components of the SE Asia region), and Brazil (sole constituent of the Latin American region) lay further distant on that scale. That pattern is replicated in the differences among national cultures shown in Figure 40.

The data in Figures 37 - 40 show, for mechanics, that national culture supercedes organizational and occupational influences. A&P mechanics, working under the same company's rules and policies, but originating from markedly different national cultures will demonstrate their national differences. For example, individuals from a high "power-distance" culture such as East Asia (see Figure 37) might be less likely to assert themselves regarding management practices than their low "power-distance" cultural counterparts. Further, the concept may be westernized to an extent that alternative training approaches to the concept would be useful for different cultures.

Effect Of National Culture On MRM Training

Attempts, such as MRM training, to change organizational culture in a multinational company will need to take national differences into account. This would predict that strongly individualistic national cultures (such as the U.S.) should create some difficulty for achieving the operational MRM objectives of collaboration, open communication, and the pursuit of and actively shared safety goals. This assertion was tested using the same international Company A line maintenance sample.

Differences between Company A's pre-training and post-training scores on the MRM/TOQ are compared among the several national cultures described above.

Line maintenance respondents for the present analysis are divided into the same three foreign regions or cultural areas described in Table A, and they are compared with U.S. mechanics now combined into a single sample. The U.S. combined sample contains all of city stations shown in the U.S. column in Table A.

Stress recognition. Figure 41 shows a dramatic increase for all four regions from before the training to after it. This strong pre-post effect in line maintenance is typical for MRM results throughout the company's ground operations work force – including hangar maintenance, engineering, and ramp services. National differences are statistically significant in Figure 41, but all four regions improve over time to the same degree. Thus, we may say the training worked as intended, with respect to beliefs and values about stress management, and that differences of national culture do not affect the training's effects in this content area.

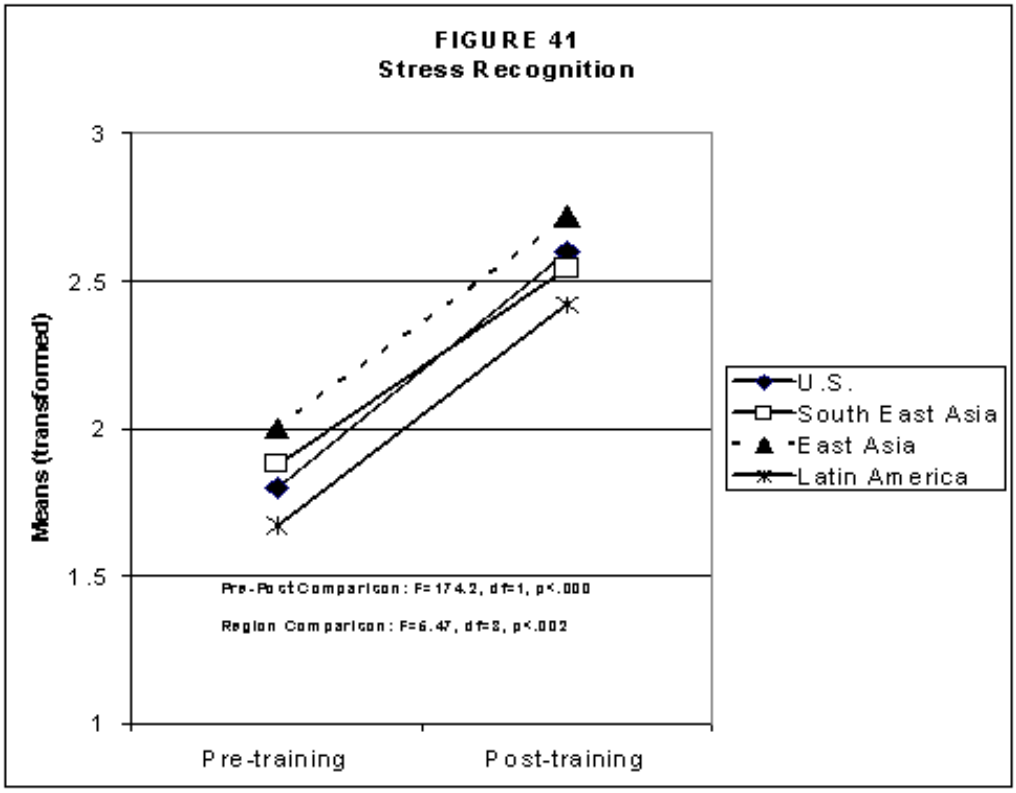

Figure 41

"Power distance." Figure 42, on the other hand, does show an interesting interaction effect between national culture and the training. In Hofstede's study "high power distance" cultures such as China and many Latin American countries stress the absolute authority of leaders. As noted above, the "Command Responsibility" index from the MRM/TOQ is very similar to other measures of power distance -- apart from being reversed in scale value. The initial (pre-training) results in Figure 42 reiterate the differences in national cultures on the power distance value previously reported.

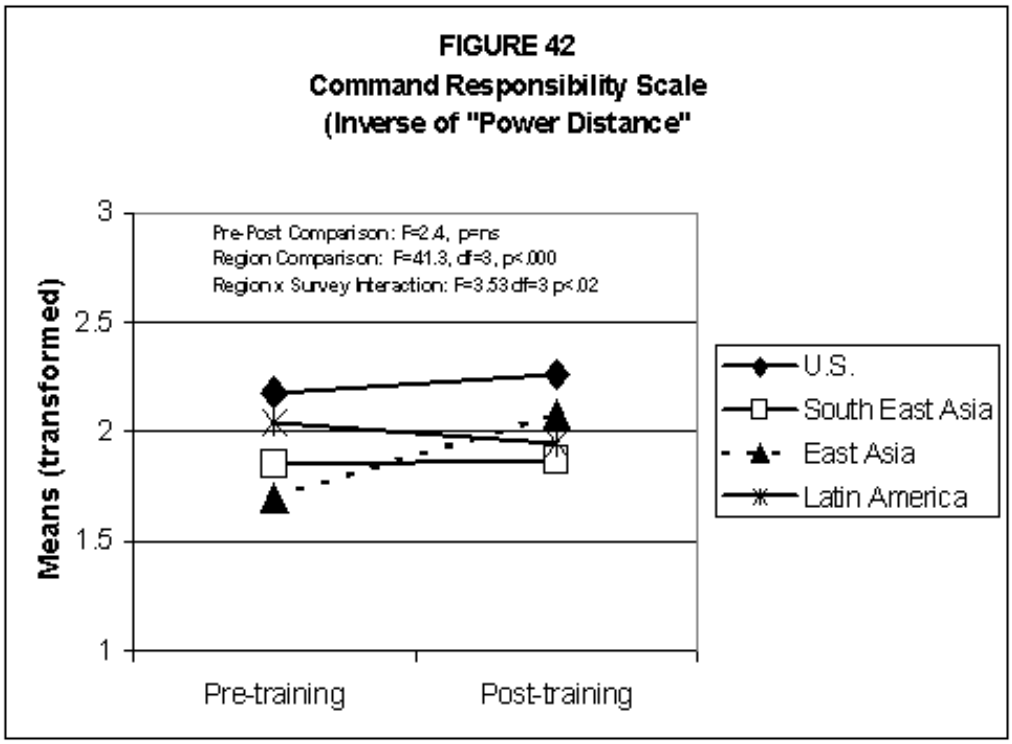

Figure 42

Sharing command responsibility is not a direct focus for Company A's MRM training except that assertiveness (with one's management as well as with coworkers) is described and discussed. Little skill training in participation or assertiveness is provided to participants. The post-training results for "command responsibility" results in this company are a slight increase in favorable attitudes. Figure 42 reveals that while three of the four national regions do not increase significantly over time, the East Asia group (Japan and Korea) starts lowest before the training and improves a great deal. This East Asian cross-over strongly suggests that their values for lessening power distance by sharing command responsibilities was affected during the two day training program. The U.S. mechanics, on the other hand, show a high initial value for sharing command (lower power distance) and although largely unchanged by the training, they remain high at the end.

Figure 42 shows that Company A's MRM awareness training appears to lessen power distance in some Asian cultures where feelings about sharing command responsibility were initially low. Whether or not this effect of national culture can be replicated, we find persistence of the major benefits of a consistently applied human factors awareness training curriculum across cultures. As MRM evolves from mere awareness training to directly influencing changes in safety practices and beliefs, the next generation MRM should be designed to compensate for cultural barriers by providing specific structures and processes that provide for "operational" impact from classroom instruction. Depending on their respective culture, MRM program participants should have the opportunity to practice and learn the program's expected interpersonal behaviors "their way."

Organizational Culture and "Revolving Door Management"<A HREF='/HFAMI/lpext.dll?f=FifLink&t=document-frame.htm&l=popup&did=F