CREW RESOURCE MANAGEMENT TRAINING'S EFFECT ON RAILROAD CREWS' PERCEPTIONS OF TASK INTERDEPENDENCE AND TEAMWORK

A Dissertation

by

TOBIN BRUCE KYTE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2008

Major Subject: Psychology

CREW RESOURCE MANAGEMENT TRAINING'S EFFECT ON RAILROAD CREWS' PERCEPTIONS OF TASK INTERDEPENDENCE AND TEAMWORK

A Dissertation

by

TOBIN BRUCE KYTE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Winfred Arthur, Jr.
Committee Members,	Mindy Bergman
	Stephanie Payne
	David Martin
Head of Department,	Les Morey

August 2008

Major Subject: Psychology

ABSTRACT

Crew Resource Management Training's Effect on Railroad Crews' Perceptions of Task Interdependence and Teamwork. (August 2008)

Tobin Bruce Kyte, B.A., Northern Arizona University;

M.A., New York University

Chair of Advisory Committee: Dr. Winfred Arthur, Jr.

The accuracy and similarity of team members' perceptions regarding the interdependencies of their task as well as the criticality of teamwork behaviors is essential to team performance. Unfortunately, these perceptions are not always accurate or similar, which has led to calls for research evaluating the effectiveness of interventions aimed at improving these perceptions. The present study evaluated the accuracy and similarity of crew members' perceptions of task interdependence and teamwork in the U.S. railroad industry. Specifically, this study assessed (1) the effect of Crew Resource Management (CRM) training on the accuracy and similarity of locomotive and maintenance of way (MOW) crew members' perceptions and (2), the extent to which the accuracy and similarity of those perceptions are retained 2-years after training.

The overall results of the present study suggests that CRM training is effective in increasing the accuracy and similarity of crew members' perceptions of team-relatedness (amount of task interdependence) and perceptions regarding the importance of

teamwork. However, the effectiveness is often dependent on the metric used (i.e., accuracy vs. similarity), and the specific characteristics of the crew members (i.e., locomotive vs. MOW, higher vs. lower interpositional experience). Furthermore, the results suggest that training did not increase the accuracy or similarity of crew members' perceptions of team workflow pattern (form of task interdependence). Lastly, a small sample size and low power precluded the running of quantitative statistical analysis assessing the long-term retention of the accuracy and similarity of participants' perceptions of task interdependence or teamwork. However, for the sake of completeness, the means, standard deviations, and effect sizes are presented in the Appendix.

DEDICATION

To my mom and dad, for all their support and guidance

ACKNOWLEDGEMENTS

I would like to express my deepest thanks to the chair of my committee, Winfred Arthur, Jr., for his assistance in the completion of this dissertation. His expertise, comments, suggestions, time and support were greatly appreciated. I would also like to thank Stephanie Payne, Mindy Bergman, and David Martin for serving as members of my committee. Their time and support were greatly appreciated. In addition, a special thanks to the Multi-Modal Freight Transportation Team at the Texas Transportation Institute for all their help and support. Special thanks goes to Curtis Morgan and Leslie Olson for efforts in facilitating the training program, collection of data, and constant support throughout this entire project. I would also like to thank Scott Kaye and Ralph Elston from the Federal Railroad Administration, whose support of CRM made this project possible. In addition, I would like to thank all the participants and management at the host railroad for their support of this project. Finally, I would like to thank Celia Pennington for her support and endless patience.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Today, much of the work in organizations is accomplished not by individuals working in isolation, but in teams (Devine, Clayton, Philips, Dunford, & Melner, 1999). It has been suggested that teams offer benefits exceeding individuals working alone, including greater adaptability, flexibility, and creativity (Paulus, 2000; Salas, Kosarzycki, Tannenbaum, & Carnegie, 2004). Furthermore, as a result of technological developments, many of today's tasks are too physically and mentally demanding and complex for a single individual to complete alone, making collaborative work a necessity for most tasks (Salas, Dickinson, Converse, & Tannenbaum, 1992; Sundstrom, De Meuse, & Futrell, 1990). For example, technology drives operations in high-reliability industries (i.e., commercial aviation, railroad, military) making the critical operational tasks (flying airplanes, driving locomotives, engaging enemy tanks) highly interdependent and complex, and thus requiring a team effort. During crew tasks, crew members share the workload, contribute specific expertise to subtasks, as well as monitor and perform back-up behaviors for other team members (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000).

Resulting directly from the interdependence of team tasks, coordination among the inputs of team members is required, and most of this coordination is the responsibility of the team members themselves (Van de Ven, Delbecq, & Koenig, 1976). There has been empirical evidence suggesting that specific behaviors related to this

---

This dissertation follows the style and format of the *Journal of Applied Psychology.*

coordination, although not relevant to independent work, are critical when working within a team (Foushee, 1984; Stout, Salas, & Carson, 1994). These critical team process or teamwork behaviors include adapting, communicating, coordinating, decision-making, interpersonal relations, leadership, shared situational awareness, and assertiveness (Bowers, Morgan, Salas, & Prince, 1993; Burke, Wilson, & Salas, 2003; Prince, Brannick, Prince, & Salas, 1992). Moreover, the performance of highly complex tasks can require specific behaviors such as performance monitoring and feedback, closed-loop communication, and back-up behaviors (McIntyre & Salas, 1995).

However, it is apparent that teams and crews do not always perform the necessary team behaviors essential for safe and efficient task performance. Accidents with primary human factors causes constitute the largest category of accidents in many high-reliability industries (Foushee, 1984; Helmreich & Foushee, 1993; Howard, Gaba, Fish, Yang, & Sarnquist, 1992; Roop, Morgan, Kyte, Arthur, & Villado, 2007). For instance, in the railroad industry, primary human factors causes have constituted approximately 38% of all train accidents over the last five years (Federal Railroad Administration [FRA], 2004). The majority of these human factor accidents are directly or indirectly related to breakdowns in coordination and communication (i.e., teamwork) within the operating crew itself (Mearns, Flin, Fleming, & Gordon, 1997; Von Thaden & Steelman, 2005). Furthermore, many high profile accidents in commercial aviation (Helmreich & Foushee, 1993), healthcare (Cooper, Newbower, & Kitz, 1984), military (Anderson & Sandza, 1987), oil and gas (Cullen, 1990), and railroad (FRA, 1999) industries have also been caused by breakdowns in coordination (assertiveness, leadership, situational awareness) and communication of the operating crew.

In order to gain a better understanding of team processes and performance, a great deal of research has assessed team member skill or ability factors that contribute to teamwork behaviors and subsequent performance (Guzzo & Dickinson, 1996). Recently however, there has been increased interest in cognitive and affective variables and their relationship to team processes and performance (Helmreich, Foushee, Benson, & Russini, 1986; Klimoski & Mohammed, 1994; Kraiger, Ford, & Salas 1993; Marks, Zaccaro, & Mathieu, 2000; Mathieu et al., 2000; Stout, Salas, & Kraiger, 1997; Tjosvold, 1986). For example, team member knowledge or perceptions of task interdependence, a key cognitive variable in several theories of team behavior and performance (Cannon-Bowers, Salas, & Converse, 1993; Tjosvold, 1986), has been the subject of recent research efforts (Anderson & Williams, 1996; Arthur, Edwards, Bell, Villado, & Bennett 2005; Bishop & Scott, 2000; Haines III, Merrheim, & Roy, 2001; Pearce & Gregerson, 1991; Shaw, Duffy, & Stark, 2000; Van der Vegt, Emans, & Van de Vliert, 1998; 2001; Ven de Ven et al., 1976; Wageman, 2001). Perceptions of task interdependence is a type of "taskwork" knowledge (Arthur et al., 2005; Cannon-Bowers & Salas, 2001) and is defined as a team member's perception of (1) the extent to which successful team performance requires the individual to work with members of the team in order to optimally perform the specified task or job and (2) the path by which work and/or information flows through the team in order to optimally perform the specified task or job (Arthur, Villado, & Bennett, in press).

Theories of team mental models suggest that the accuracy and similarity of team members' knowledge or perceptions of task interdependence are essential to team performance (Cannon-Bowers et al., 1993; Rentsch & Woehr, 2004) because interdependence is a defining characteristic of team tasks (Cannon-Bowers, Salas, Blickensderfer, & Bowers, 1998; Day, Fein, & Arthur, 2000; Tziner & Edens, 1985). Accurate and similar perceptions of task interdependence ensure behavior is functional and task relevant (Arthur et al., 2005; Rentsch & Hall, 1994) and allows team members to predict other members' behaviors and actions, and better coordinate their own (Salas et al., 1992). Using theories of team mental models as a basis, research has found positive relationships between the accuracy and similarity of team members' knowledge or perceptions of the task and team performance (Arthur et al., 2005; Cooke et al., 2003; Edwards, Day, Arthur, & Bell, 2006; Marks et al., 2000; Marks, Sabella, Burke, & Zaccaro, 2002; Mathieu et al., 2000; Mathieu, Heffner, Goodwin, Cannon-Bowers, & Salas, 2005; Smith-Jentsch, Mathieu, & Kraiger, 2005).

The above research, in conjunction with findings indicating within-group variability in team members' perceptions of task interdependence (Comeau & Griffith, 2005; Wageman, 1995) suggest that researchers should look for ways to improve the accuracy and similarity of team members' perceptions or knowledge of the task (Arthur et al., 2005; Brun et al., 2005; Cooke et al., 2003; Edwards et al., 2006; Webber, Chen, Payne, Marsh, & Zaccaro, 2000). Training interventions have been found to be one of the most effective human resource interventions for improving performance and attitudes (Arthur, Bennett, Edens, & Bell, 2003; Guzzo, Jette, & Katzell, 1985; Neuman, Edwards, & Raju, 1989) and it has been suggested that one purpose of team training should be to improve the accuracy and similarity of team members' knowledge regarding the interdependence requirements of the task and team (Arthur et al., 2005; Cannon-Bowers et al., 1993; Cooke, Kiekel, & Helm, 2001; Cooke et al., 2003; Edwards et al., 2006; Marks et al., 2002; Marks et al., 2000; Rouse & Morris, 1986; Tesluk, Matheiu, Zaccaro, & Marks, 1997). However, there are only a handful of studies assessing the impact of training on criteria related to team members' knowledge and perceptions of task interdependence (i.e., team-interaction models; see Brun et al., 2005; Cooke et al., 2003; Marks et al. 2002; Marks et al., 2000), and an extensive and comprehensive literature search did not locate any studies assessing the impact of training interventions on perceptions of task interdependence directly.

The content covered and methods used in Crew Resource Management (CRM) training seem particularly relevant for increasing the accuracy and similarity of team member perceptions of interdependence (see Salas et al., 1992). CRM is widely used in the commercial aviation industry to teach cockpit crews how to work together as a team and relies heavily on analyzing and discussing case studies of real life aviation accidents that were caused by teamwork failures (Federal Aviation Administration [FAA], 2004). Thus, the first objective of the present study was to examine the impact of CRM training on railroad crews' perceptions of task interdependence.

A second objective of the present study was to investigate the effect of railroad CRM training on the accuracy and similarity of participants' perceptions of teamwork and outcomes that are traditionally used to evaluate CRM in other high-reliability industries. For example, studies in commercial aviation (Salas, Burke, Bowers, & Wilson, 2001), off-shore oil drilling (Flin & O'Connor, 2001), and healthcare (Howard et al., 1992) have shown that not only does CRM increase knowledge and improve trainees' attitudes toward teamwork, but crew members react positively to the training as well. Although there is a high degree of similarity in the tasks, team processes, and accidents attributed to the failure of the crew coordination in the airline and railroad industries (Roop et al., 2007), the implementation and practice of CRM in the railroad industry is limited and there are no known studies of its effectiveness in this environment (Morgan, Olson, Kyte, & Roop, 2003). Thus, I conceptualize crew members' perceptions regarding teamwork as a type of teamwork knowledge (Cannon-Bowers & Salas, 2001; Rentch & Hall, 1994) and posit that these perceptions should be seen as a key affective training criterion in the railroad environment.

Furthermore, although critical in determining the time frame for refresher training (Roop et al., 2007), I was unable to locate any research assessing the degree to which participants' CRM knowledge, or accuracy and similarity of attitudes toward teamwork and perceptions of task interdependence degrade over time in the railroad industry. Thus, the third objective of the present study was to investigate the extent to which performance on these criteria is retained after training. This is important because as noted by Schmidt and Björk (1992), acquisition and retention are inseparable. Consequently, the evaluation of the effectiveness of any training intervention should be in terms of not only initial acquisition, but also long-term retention.

Task Interdependence

Task interdependence is a critical variable in numerous domains of organizational theory and research including organizational design (Galbraith, 1977; Thompson, 1967; Weick,1979), job design (Brass, 1985; Kiggundu, 1983; Medsker & Campion, 1997), social psychology of groups (Tjosvold, 1986), and team task analysis, design, and performance (Arthur et al., 2005, Campion, Medesker, & Higgs, 1993; Dieterly, 1988; Medsker & Campion, 1997; Salas et al., 2004; Tesluk et al., 1997). However researchers in these domains often conceptualize and measure task interdependence differently because they are interested in specific aspects of interdependence. For example, some researchers focus on task interdependence as the relationship between organizational units (Thompson, 1967; Van de Ven & Ferry, 1980), whereas others suggest it represents the degree of task-driven interaction among work group members (Shea & Guzzo, 1987). In an attempt to clarify the term "interdependence," Wageman (2001) makes a distinction between "task interdependence," a structural characteristic of the task itself, and "behavioral interdependence," the amount of task-related interaction (teamwork type behaviors) actually engaged in by team members. In congruence with Wageman (2001), researchers have recently described task interdependence as a "taskwork" variable which can be conceptualized as (1) the extent to and (2) the manner in which group members must exchange information and resources (i.e., work together) to complete their tasks (Arthur et al., 2005; Arthur et al., in press; Van Der Vegt, Emans, & Van De Vliert, 2000). These two related, yet distinct, aspects of task interdependence are derived directly from the team task analysis literature (see Arthur et al., 2005; Wageman, Hackman, & Lehman, 2005). Similar to other task characteristics, an analysis of a task's interdependence can be used to redesign jobs, create selection systems, identify specific types of teams, choose training interventions, understand performance problems, and better measure specific aspects of performance (Arthur et al., 2005; Dieterly, 1998; Medsker & Campion, 1997; Tesluk et al., 1997).

Team-relatedness. Many authors define and conceptualize task interdependence as the extent to which a task requires the input of resources from multiple individuals for its completion (Pearce & Gregersen, 1991; Van der Vegt et al., 1998; Wageman, 1995). Work and information can include skills, materials, tools, equipment, and instructions (Kiggundu, 1981). Similarly, Arthur et al. (in press) label this conceptualization of interdependence as "team-relatedness" and define it specifically as, "the extent to which successful team performance requires the individual to work with members of the team in order to optimally perform the specified task or job" (p. 3). Although tasks are often referred to as "interdependent" if to some degree they require input from multiple individuals for completion, researchers posit that task interdependence is not a dichotomous variable. Rather, it is a continuous variable (Arthur et al., 2005; Wageman, 2001). Representing the *amount* of task interdependence (Arthur et al., in press), conceptualizing and defining team-relatedness as "the extent to which successful team performance requires the individual to work with members of the team in order to optimally perform the specified task or job" (p. 3), reflects its continuous nature (Wageman, 1995).

Team workflow pattern. Because multiple individuals input resources into an interdependent task, there needs to be an exchange of resources between team members for the task to be completed (Doerr, Mitchell, Schriesheim, Freed, & Zhou, 2002). Considering this, many authors have recently suggested that task interdependence should be conceptualized and operationalized as more than just the amount or magnitude of task interdependence (i.e., team-relatedness), but should also include the manner in which resources move among members of a team (Arthur et al., 2005; Arthur et al., in press; Van Der Vegt et al., 2000). This conceptualization of task interdependence is defined by Shea and Guzzo (1987) as “the order in which materials, information, or other resources are transferred between workers while completing a task” (p. 331). Arthur et al. (in press) label this conceptualization of task interdependence as “team workflow pattern,” and unlike team-relatedness, is conceptualized as a nominal variable representing the type, kind, or form of task interdependence. In fact, team workflow pattern, as a conceptualization of task interdependence, has had a long history starting with Thompson (1967) who suggested that there were three distinct ways in which work flowed through teams: (1) pooled/additive, (2) sequential, and (3) reciprocal. Tesluk et al. (1997), working from Thompson’s (1967) original conceptualization, suggested a fourth workflow pattern, which they titled “intensive.” It is suggested that as workflow patterns move from “pooled/additive” to “intensive,” there is an increased dependence among team members for job performance and thus an increased need for coordination (Saavedra, Early, & Van Dyne, 1993) and research has shown that as a task moves from a pooled to an intensive workflow pattern, team members perceive a greater degree of team-relatedness (Arthur et al., 2005; Arthur, Glaze, Bhupatkar, & Villado, 2008; Comeau & Griffith, 2005).

A *pooled/additive* team workflow pattern is one in which team members perform their activities separately and the work is simply pooled or aggregated to make up team performance (Thompson, 1967). A pooled/additive workflow is distinct from other workflow patterns in that no resources flow between the team members. Tesluk et al. (1997) give the example of a janitorial crew who are given the task of cleaning a hospital. Each member of the janitorial crew works independently on a number of different tasks in a specific section of the hospital. In order for the task of "cleaning the hospital" to be accomplished, each crew member's completed tasks are pooled or added together.

A *sequential* team workflow pattern represents a situation in which work and activities flow from one member of a team to another, but mostly in one direction (Tesluk et al., 1997). Work on an assembly line is a good example of a sequential workflow pattern in that resources move in a unidirectional pattern from team member to team member. Compared to a pooled/additive team workflow pattern, task performance is not simply additive, meaning that if one person on the team does not complete his or her specific task, eventually the work of others down the line will be stymied.

Information and resources can also move between team members in a *reciprocal* workflow pattern. Unlike the unidirectional flow of resources in sequential workflow, reciprocal workflow patterns are more dynamic and involve resources moving bi-directionally between team members. An example given by Tesluk et al. (1997) is the task of developing and running a therapeutic program for a hospital's psychiatric unit. A mental health treatment team consisting of a psychiatrist, psychologist, social worker, nurse, and one or more nursing assistants exchange resources in the performance of this task. Tesluk et al. (1997) further explain the workflow of this team as follows:

> A psychiatrist depends on other team members to provide information about a patient's mental functioning and behavior to make the most appropriate decisions regarding the type of psychotropic medication to prescribe. In a similar manner, other team members conducting rehabilitative activities (e.g., psychologist, social worker) depend on the psychiatrist's prescription practices when they are engaged in counseling activities. Most of these back-and-forth type interactions that occur during treatment team meetings can be anticipated. If during a team meeting the psychiatrist suggests altering a patient's medication, certainly the psychologist and nurses need to know how this may impact the patient's behavior and if they, in turn, need to adjust their treatment strategies. Hence, anticipated coordination and reciprocal interdependencies between mental health care team members are critical points for effective team performance defined in terms of improvement of the patient's condition (p. 203).

As can be seen in the example above, distinct from the assembly line nature of sequential interdependence where the work of one individual becomes the input for another, the back and forth flow of resources featured in reciprocal interdependence can represent a type of feedback loop (Tesluk et al., 1997).

Finally, the *intensive* workflow pattern represents a situation in which resources, materials, and information flow back and forth among all members of the team. This differs from a reciprocal team workflow pattern where the bi-directional exchange of resources is only between particular members of a team and the general order through which work flows is in a particular pattern. Comparatively, Tesluk et al. (1997) emphasize that it can be difficult to predict when and where the sharing and flow of information will occur within the team because the nature of tasks that require an intensive workflow pattern is dynamic and complex. Tasks that require this type of workflow are also prevalent in the hospital setting, for example operations and emergency situations are performed by operating room teams. Tesluk et al. (1997) state that, “because of the often highly stressful, quick-paced, real-time nature of the task and the serious consequences that may result from poor performance, members of operating room teams share a high level of interdependence” (p. 204).

Perceptions of Task Interdependence

Based on the preceding review and the conceptualization of task interdependence by Arthur and associates (Arthur et al., in press; Arthur et al., 2008), perceptions of task interdependence is defined as a team member’s perception of (1) the extent to which successful team performance requires the individual to work with members of the team in order to optimally perform the specified task or job and (2) the path by which work and/or information flows through the team in order to optimally perform the specified task or job. Team members’ perceptions and knowledge regarding the interdependencies of their tasks are critical variables in several theories of group and team behavior (Cannon-Bowers et al., 1993; Deutsch, 1949; 1962; Tjosvold, 1986). For example, Deutsch's (1949) theory of cooperation posits that social interaction can be understood in terms of how individuals perceive their goals to be related to others. According to the theory, individuals' perceptions or beliefs about how they are interdependent directly affect their interaction patterns, feelings, attitudes, learning, and productivity (Tjosvold, 1986). That is, group members' behavior is largely determined by their perceptions that they are linked to others in their group in such a way that they must work together to succeed (Johnson & Johnson, 1989). Although Deutsch's (1949) model was originally intended for non-work groups, Tjosvold (1986) expanded it to teams in organizational settings where there are objective task assignments and task interdependence is often highly structured.

Similarly, theories of team mental models posit that team members' knowledge structures regarding the interdependencies of the team task are important cognitive components of behavior (Cannon-Bowers et al., 1993; Rentsch & Woehr, 2004; Rouse, Cannon-Bowers, & Salas, 1992). A mental model is conceptualized as an organized knowledge structure (Mathieu et al., 2000) and is defined specifically as a "psychological representation of the environment and its expected behavior" (Holyoak, 1984, p. 193). Rouse and Morris (1986) state that a mental model is a "mechanism whereby humans generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future system states" (p. 360). Thus the purpose of a mental model is to allow an individual to describe and explain relationships between variables in their environment, to construct expectations, and to make predictions regarding those variables, which in turn allows them to take appropriate action regarding their interaction with those variables. When working in a team setting, one of the most important variables in the environment is fellow team members. Subsequently, predicting team members' actions, thoughts, and behaviors is critical (Cannon-Bowers et al., 1993; Mathieu et al., 2000; Rouse et al., 1992).

Research suggests that there is not one type of team mental model, but instead several that are critical for team performance (Cannon-Bowers et al., 1993; Klomoski & Mohammed, 1994; Mathieu et al., 2000; Rouse et al., 1992). One team mental model believed to be critical to team processes and performance is team members' "team-interaction" model which represents their knowledge regarding the interdependencies of the task (Brun et al., 2005; Cannon-Bowers et al., 1993; Cannon-Bowers & Salas, 2001; Mathieu et al., 2000; Marks et al., 2002; Rouse et al., 1992). Parallel to the definition and conceptualization of perceptions of task interdependence outlined in this paper, a team-interaction mental model includes knowledge regarding task related activities and action sequences, roles and responsibilities, information sources, interaction patterns, communication channels, role interdependencies, and task specific patterns of information flow required for effective performance (Cannon-Bowers et al., 1993; Cannon-Bowers & Salas, 2001; Klimoski & Mohammed, 1994). This knowledge forms the basis of team members' expectations, which in turn influence behavior (Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995; Tannenbaum, Beard, & Salas, 1992).

Although research has shown a positive relationship between the actual, structured interdependence of a task and team members' perceptions of that interdependence (Arthur et al., 2005; Wageman, 1995), theorists suggest and research has shown that team members can misperceive the structural interdependence of their task and can have different perceptions than their teammates (Arthur et al., 2005; Tjosvold, 1986). Unfortunately, as a theory rooted in non-work settings (Tjosvold, 1986), cooperation theory does not make predictions regarding the effects of misperceiving the interdependence of the task or the effects of within-team variance of perceptions of task interdependence on performance. However, team mental model theory develops the conceptual links between team member knowledge of task interdependence and performance by introducing the concepts of accuracy and similarity (Cannon-Bowers et al., 1993; Cooke, Salas, Cannon-Bowers, & Stout, 2000).

Accuracy. Theory and research suggest that to positively affect team processes and performance, team members' mental models need to be accurate (Cannon-Bowers et al., 1993; Cooke et al., 2000; Edwards et al., 2006; Rentsch & Hall, 1994; Webber et al., 2000). Edwards et al. (2006) stated that team mental model accuracy "refers to the degree to which team members' mental models adequately represent a given knowledge or skill domain" (p. 727). Because team members' cognition, and ultimately behavior, is based on mental models, the accuracy of those mental models is essential for that behavior to be functional, task relevant, and lead to optimal task performance (Rentsch & Hall, 1994).

Research on team-interaction mental models supports the link between the accuracy of perceptions of task interdependence and performance (Edwards et al., 2006). As previously noted, team-interaction models represent knowledge regarding the interdependencies of the task, including knowledge regarding task-related activities and action sequences, and task specific information flow required for effective performance (Cannon-Bowers et al., 1993; Cannon-Bowers & Salas, 2001; Klimoski & Mohammed, 1994). As a result, operationalizations and measures of team-interaction mental models indirectly tap perceptions of task interdependence as conceptualized in this paper (i.e., team-relatedness and team workflow pattern). For example, many studies assessing team mental models take place in the laboratory using 2-3 member plane, tank, or helicopter simulators (see Cooke et al., 2003; Edwards et al., 2006; Marks et al., 2002; Marks et al., 2000; Mathieu et al., 2000; Mathieu et al., 2005). The study protocols divide mission critical subtasks between team members and team-interaction mental models are measured by assessing team members' perceptions regarding the relatedness of those subtasks or perceptions of how the subtasks work together to help the team member optimally perform the overall task or mission (see Edwards et al., 2006). As a result of subtasks being distributed among team members, the relatedness ratings (team-interaction mental models) are indirectly measuring team members' perceptions of team-relatedness as conceptualized in this paper. Research has shown that as team member relatedness ratings (team-interaction mental models) become more accurate, performance increases (Edwards et al., 2006).

Using the same research design, research on team-interaction mental models has also measured team members' perceptions of the type of information that passes between each pair of team members and in what direction. For example, Cooke et al. (2003) found that team members who could accurately identify the type of information that flowed between various team members had higher levels of team performance than those who could not. As an indirect measure of team members' perceptions of team workflow pattern, this research supports the link between the accuracy of perceptions of task interdependence and performance. In congruence with theories of team mental models (Cannon-Bowers et al., 1993; Klimoski & Mohammed, 1994), as a result of correctly perceiving the team workflow pattern, participants approach the task and behave in a manner consistent with that interdependence (i.e., teamwork behaviors). This compatibility of a team members' approach to the task (team processes) and the demands of the task increase the likelihood of successful task performance.

Similarly, Arthur et al. (2005) used a team mental model framework to assess the link between performance and the accuracy of perceptions of task interdependence as conceptualized in this paper. In their study, Arthur et al. specifically manipulated the task interdependence of their computer-based experimental task to reflect differing degrees of team-relatedness and team workflow pattern. Differing by only their task interdependence, the tasks categories served as true scores which were subsequently used to assess the accuracy of team members' perceptions of task interdependence. Arthur et al. found that participants who had more accurate perceptions of a task's team-relatedness and workflow pattern performed better on the task. Supporting theories of team mental models, Arthur et al. noted that team members who work on a highly interdependent task, and recognized it as such, perform better on the task because their behavior and overall approach to the task is compatible with the task-based requirements needed for successful performance. This same reasoning applies to results in Arthur et al.'s (2005) individual task condition. Arthur et al. found that teams that worked on individual-based tasks, and perceived those tasks to be low in task interdependence, performed better than teams that perceived the tasks to be highly interdependent.

Similarity. In addition to being accurate, theory and research suggest that team member mental models need to be similar. Similarity refers to the degree to which team members' mental models are similar or overlapping and it is suggested that this similarity is the key to understanding mental models' effect on team processes and performance in the team environment (Cannon-Bowers, Salas, & Converse, 1990; Cannon-Bowers et al., 1993; Klimoski & Mohammed, 1994; Rouse et al., 1992). When team members have similar mental models, they interpret cues in a similar manner, resulting in team members making compatible decisions and effectively coordinating their activities in difficult and constantly changing situations (Cannon-Bowers & Salas, 2001; Cannon-Bowers et al., 1993). In fact, the importance of team mental model similarity is so prevalent that the majority of the literature uses the term "shared mental models" instead of "team mental model." Although there has been more empirical research on mental model similarity than accuracy (Edwards et al., 2006; Stout et al., 1997), research suggests that the importance of accuracy supersedes that of similarity because by definition, as team members' mental models become more accurate, those mental models also become more similar (Edwards et al., 2006). In contrast, similarity does not imply or guarantee accuracy. That is, although team members may have similar or overlapping mental models, they might be incorrect (Cannon-Bowers et al., 2001; Edwards et al., 2006), leading to failures of team processes and performance. A recent study using a task with a single best way to complete it suggests that team mental model accuracy may be a better predictor of performance than similarity (Edwards et al., 2006).

Related to perceptions of task interdependence, Cannon-Bowers et al. (1993) stated that "effective team performance requires that team members hold common or overlapping cognitive representations of task requirements, procedures, and role responsibilities" (p. 221-222). Much of the research discussed above highlights the importance of accurate "team-interaction" models and also supports the positive correlation between the similarly of team-members' perceptions of task interdependence and performance (Edwards et al., 2006; Marks et al., 2002; Marks et al., 2000; Mathieu et al., 2000; Mathieu et al., 2005; Smith-Jentsch et al., 2005). For example, Mathieu et al. (2000) found that the similarity of team members' perceptions regarding the relatedness of subtasks was positively correlated with performance. As stated earlier, as a result of subtasks being distributed among team members, the team-interaction mental models (i.e., relatedness ratings) are similar to team members' perceptions of team-relatedness as conceptualized in this paper.

In summary, perceptions of task interdependence is conceptualized and defined as a team members' perceptions regarding (1) the extent to which successful team performance requires the individual to work with members of the team in order to optimally perform the specified task or job and (2) the paths by which work and/or information flows through the team in order to optimally perform the specified tasks or job (Arthur et al., in press). Several theories of group behavior suggest that team members' perceptions or knowledge regarding the interdependencies of the task represent a critical cognitive link between features of the task and team member behavior (Deutsch, 1949; 1962; Tjosvold, 1986) and performance (Cannon-Bowers et al., 1993; Rouse et al., 1992). Borrowing from the team mental model framework, research has shown that the extent to which team members' perception or knowledge of a task's interdependence are accurate and similar is related to team performance (Arthur et al., 2005; Cooke et al., 2003; Edwards et al., 2006; Marks et al., 2002; Marks et al., 2000; Mathieu et al., 2000; Mathieu et al., 2005; Smith-Jentsch et al., 2005). Research suggests that in the same way that the characteristics of a task specify the requisite actions and behaviors necessary for task performance (Cannon-Bowers et al., 1995; Fleishman & Zaccaro, 1992), accurate perceptions of the task, including its interdependencies, are necessary for team members to have an accurate understanding of what teamwork behaviors and processes will lead to effective task performance (Arthur et al., 2005). That is, team members make decisions regarding what behaviors are appropriate during a task based on their understanding of the interdependencies of the task.

Team Training

The importance of team members having accurate and similar perceptions of task interdependence, in conjunction with research showing within-group variability in team members' perceptions of task interdependence (Comeau & Griffith, 2005; Wageman, 1995), has led to calls for research investigating ways to improve the accuracy and similarity of team members' perceptions or knowledge regarding the interdependencies of their tasks (Arthur et al., 2005; Brun et al., 2005; Cooke et al., 2003; Edwards et al., 2006). Of the various human resource interventions at their disposal, many suggest that training in particular is likely to be effective in influencing "team-interaction" mental models (Cannon-Bowers et al., 1993; Cannon-Bowers et al., 1995; Cooke et al., 2003; Edwards et al., 2006; Marks et al., 2002; Marks et al., 2000; Rouse et al., 1992; Tesluk et al., 1997) and perceptions of task interdependence specifically (Arthur et al., 2005). Goldstein (1991) states that the training process is "the systematic acquisition of attitudes, concepts, knowledge, rules, or skills that result in improved performance at work" (p. 508). Conceptual models of team effectiveness outline the mechanisms through which team training affects mental models. For example, Tannenbaum et al.'s (1992) conceptual framework of team training and performance suggests that team training is an organizational input that affects team members' task knowledge, mental models, and attitudes. According to the model, team members' knowledge and attitudes affect team performance through their effect on team processes (Gladstein, 1984; Tannenbaum et al., 1992).

Consistent with theories of team training and performance (Rouse et al., 1992; Tannenbaum et al., 1992), recent research suggests that team training can indeed influence the accuracy and similarity of various "team-interaction" mental models. For example, Cooke et al. (2003) found that full cross-training (compared to conceptual cross-training and no training) resulted in participants reporting more accurate knowledge of the type and direction of information that flowed between various team members (i.e., team workflow pattern). Similarly, research has found that both team-interaction training and cross-training lead to increased accuracy and similarity of team members' knowledge regarding the relatedness and sequencing of subtasks and team member actions (Marks et al., 2002; Marks et al., 2000). The present study investigates the effect of CRM training on participants' perceptions of task interdependence.

CRM Training

Originally called "cockpit resource management," CRM training is defined as a "family of instructional strategies that seek to improve teamwork in the cockpit by applying well-tested training tools (e.g., simulators, lectures, videos) targeted at specific content (i.e., teamwork knowledge, skills, and attitudes"; Salas et al., 1999, p. 163). In order to combat the increasing number of aircraft accidents that were attributed to "pilot error," commercial airlines in conjunction with the National Aeronautics and Space Administration (NASA) developed CRM training almost a quarter century ago (Helmreich, Merritt, & Wilhelm, 1999; Salas, Bowers, & Edens, 2001). Eventually, CRM moved outside the cockpit to include flight attendants and maintenance personnel and recently, CRM training has been adopted in other high-consequence team-related industries outside commercial aviation including the military, commercial shipping, and the medical, nuclear power, and maritime industries (Helmreich et al., 1999). The distal goal of CRM training is to improve teamwork behaviors in the work environment (e.g., cockpit, locomotive cab) in an effort to reduce errors related to those behaviors (Stone & Babcock, 1988; Stout, Salas, & Fowlkes, 1997). As reflected in its definition, the more proximal goal of CRM training is to improve teamwork knowledge, attitudes, and skills (Helmreich & Wilhelm, 1991; Salas, Fowlkes, Stout, Milanovich, & Prince, 1999).

Although practitioners suggest that full CRM training initiatives involve practice, feedback, and continued reinforcement on the job (FAA, 2004), the foundation of CRM training is initial awareness training which gives participants introductory knowledge of the concepts related to CRM and thus introduces a common CRM vocabulary. Since its inception, the specific learning domain sampled by CRM training has not been consistent (Salas et al., 1999). However the FAA (2004) currently suggests that CRM content should reflect (1) communication processes and decision behavior, and (2) team building and maintenance. Content covered under "communication processes and decision behavior" include briefings, assertion, crew self-critique of work processes, conflict resolution, communication, and decision-making. It should also include discussions on effective leadership and followership including staying focused on the goal of achieving safe and efficient flight. Content covered in "team building and maintenance" should include interpersonal relationships, group climate, individual factors and stress reduction, workload management and distribution, and situational awareness (FAA, 2004).

Although the methods used in the delivery of CRM awareness training have not been uniform in the airline industry (Helmreich & Wilhelm, 1987; Salas et al., 1999), CRM training usually takes place in a one- or two-day course, with 8-20 participants in a class, facilitated by instructors who are knowledgeable of CRM principles and human error. The FAA (2004) suggests that initial CRM training should involve a combination of several training methods including lectures, audiovisual presentations, discussion groups, role-playing exercises, and videotaped examples of good and poor team behavior. The awareness program typically uses case studies, showing real life examples of accidents caused by lack of teamwork in the cockpit, or examples of CRM "saves."

Effect of CRM on perceptions of task interdependence. Research suggests that the effect of training on team-member knowledge and perceptions depends on the type of training, team, and knowledge (Marks et al., 2002; Rouse et al., 1992) and mismatches have been suggested as reasons for disappointing results (see Brun et al., 2005). Because researchers have broadly suggested that CRM might be useful for creating and reinforcing team mental models (Rouse et al., 1992), CRM training seems particularly relevant and would most likely be effective for increasing the accuracy and similarity of railroad crew members' perceptions of task interdependence. First, as previously noted, human factors caused accidents constituted approximately 38% of all train accidents over the last five years (FRA, 2004), and like other high-reliability industries (Mearns et al., 1997; Von Thaden & Steelman, 2005), the majority of these were directly or indirectly related to breakdowns in coordination and communication (i.e., teamwork) within the operating crew itself (FRA, 1999; Roop et al., 2007). Because perceptions of task interdependence are important for team processes and performance (including safety), training interventions aimed at increasing the accuracy and similarity of those perceptions seem particularly important for railroad crews.

Furthermore, there is reason to suggest that similar to other teams (Arthur et al., 2005; Tjosvold, 1986), there is a lack of convergence between a railroad task's structural interdependence (Wageman, 2001) and railroad crew members' perceptions of that task's interdependence (i.e., perceptions are inaccurate and unshared). First, research has shown that there is within-group variability in team members' perceptions of task interdependence (Comeau & Griffith, 2005; Wageman, 1995). Moreover, although the tasks railroad crews perform are highly team-related and require an intensive workflow pattern (Morgan et al., 2003), various factors specific to the railroad environment could lead crew members to perceive their jobs as less team-related and involving a less intensive workflow pattern than is actually the case. For example, during some critical job tasks (e.g., taking a train order), there are strict rules and regulations regarding communication procedures between locomotive crew members and dispatchers. Although critical, the highlighting of these strict "sequential" communication procedures could gradually result in the locomotive crew failing to recognize other instances where less proceduralized, more reciprocal or intensive workflow patterns occur. Moreover, several factors specific to the railroad industry, such as the crew members belonging to different unions because they occupy different jobs, could plausibly result in crew members perceiving their specific jobs as being quite independent. Furthermore, similar to other highly structured crew environments (Dieterly, 1988), practically all training for railroad crews focuses on the technical aspects of specific tasks individual crew members complete, with little or no emphasis on "team" tasks (Morgan et al., 2003). As a result, crews in the railroad industry appear particularly amenable to a team training intervention that is aimed at increasing the accuracy and similarity of team members' perceptions of task interdependence, teamwork behaviors, team performance, and safety.

CRM training is particularly well suited for increasing perceptions of task interdependence in the railroad industry. First, CRM is specifically tailored for crews in high-reliability industries, and the railroad industry is a high-reliability industry that uses crews as its elemental team (Morgan et al., 2003). More importantly, the content and methods used in CRM training are well suited to increase the accuracy and similarity of perceptions of interdependence. For example, CRM training involves content germane to highly interdependent teams including communication processes, team decision-making, briefings, team building, assertiveness, interpersonal relationships, group climate, workload management and distribution, and conflict resolution. Furthermore, that all crew positions attend the same training class and its reliance on case studies or examples of accidents makes CRM particularly relevant for increasing the accuracy and similarity of perceptions of task interdependence in these crews. For example, Ellis, Bell, Ployhart, Hollenbeck, and Ilgen (2005) showed that team training based on case studies and emphasizing task coordination led to higher levels of teamwork and performance. Ellis et al. noted that case studies help trainees understand workload management and distribution as well as issues involved in the appropriate allocation of tasks to team members. Similarly, Salas et al. (1992) stated that "training for coordination should help team members to recognize and identify their task interdependencies and should emphasize the undesirable consequences of failure to coordinate team members efforts correctly" (p. 19).

The case studies discussed during CRM training involve accidents that result from crew members not moving resources (information) through the team for proper and safe task accomplishment, for example, two-way communication failures or improper team decision-making processes. These examples are representative of reciprocal or intensive workflow patterns and would suggest to crew members that it is extremely difficult and unsafe to perform their jobs alone or use lower-order workflow patterns. Based on the preceding argument, CRM training should result in increased accuracy and similarity of crew members' perceptions of task interdependence (team-relatedness and team workflow pattern). However, research suggests that the effectiveness of CRM in changing team knowledge depends on the accuracy of that knowledge prior to training (Brun et al., 2005). As previously noted, specific characteristics of the railroad could negatively affect the accuracy and similarity of a crews' perceptions of task interdependence. However, the pervasiveness of these environmental factors will most likely vary with the type of railroad crew. The two types of crews assessed in the present study were locomotive and maintenance of way (MOW) crews.

Webber and Klimoski (2004) define a crew as "a group of expert specialists each of whom have specific role positions, perform brief events that are closely synchronized with each other, and repeat these events across different environmental conditions" (p. 265). The role specialization and closely synchronized workflow patterns present in crew tasks are often accompanied by strict rules and procedures outlining those workflow patterns as well as specific teamwork behaviors (i.e., communication patterns) needed to complete those tasks successfully (Morgan et al., 2003; Webber & Klimoski, 2004). A locomotive crew is comprised of an engineer and conductor whose task is to safely operate a freight train from one destination to another while complying with all federal and company rules and regulations (Morgan et al., 2003). The locomotive engineer controls the forward and reverse movement of the train while observing and complying with all signal indications and speed restrictions. He or she also ensures trains are loaded in compliance with recognized handling procedures. The conductor is responsible for the overall operation of the train including all train documentation, any restrictions existing on the train, speed requirements, track restrictions, classification of the cars in the train, pre-trip planning, and post-trip summaries. Traditionally, "new" locomotive crews are formed on a just-in-time basis (usually every day) by combining an engineer and a conductor from a railroad's local labor pool.

The general task of a MOW crew is to perform daily assignments of maintenance and repairs of track and rights of way (Morgan et al., 2003). A MOW crew is generally comprised of at least four persons including a foreman, trackman laborer, truck driver, and machine operator. On a larger job, a MOW crew can be augmented by the addition of an assistant foreman, track supervisor, welder and grinder team, speed swing operator, front-end loader operator, dump truck loader, or a flagman "lookout." The foreman supervises the MOW crew and helps with any of the work while track laborers do a variety of jobs related to installing and repairing railroad track. A truck driver operates the work truck that carries the crew to the worksite and the machine operator maneuvers the machinery that moves and lays track or rails to construct, repair, or maintain railroad tracks. A welder and grinder work together to connect rail and a speed swing operator operates a piece of heavy equipment mounted on the ground or track that is used to place new rail on track, and hold it while it is being cut. Finally, a front-end loader operator uses a front-end loader to pick up and move various objects (such as ties, rocks, rail) from one area of the worksite to another. Unlike locomotive crews, MOW crews are not created on a just-in-time basis and traditionally have relatively stable membership.

Although both locomotive and MOW crews would be categorized as “crews” (Webber & Klimoski, 2004), MOW crews tend to have fewer rules and regulations regarding their communication processes as well as less distinction among crew members’ specific tasks (which is most likely the result of working in larger crews and having relatively stable crew membership). Furthermore, unlike locomotive crews, union membership in MOW crews does not cut across jobs. As a result of having more rules, regulations, distinctions between crew member tasks, and divergent within-crew union membership, I suggest that locomotive crew members will have a tendency to perceive their position-specific task as an end in itself, which will cause them to have less accurate and similar perceptions of task interdependence compared to MOW crew members. Thus, although both types of crews are highly interdependent, I suggest that prior to training, there will be higher levels of similarity among MOW crew members’ perceptions of task interdependence compared to locomotive crew members. It is also posited that these perceptions will be more accurate. As a result, the effect of CRM training on crew members’ perceptions of task interdependence will be moderated by crew type such that locomotive crew members will report greater increases in accuracy and similarity than MOW crew members. Consequently, it was hypothesized that:

Hypothesis 1: There will be a significant main effect for training on perceptions of task interdependence. Specifically, when crossed with crew type the (a) accuracy and (b) similarity of participants’ perceptions of team-relatedness and (c) accuracy and (d) similarity of participants’ perceptions of team workflow pattern will be higher at Time 2 (immediate post-training) than at Time 1 (pre-training).

Hypothesis 2: There will be a significant simple effect for crew type on perceptions of task interdependence at Time 1. Specifically, at Time 1 the (a) accuracy, and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be greater for participants who belong to MOW crews compared to locomotive crews.

Hypothesis 3: There will be a significant interaction between training and crew type on perceptions of task interdependence. Specifically, from Time 1 to Time 2, increases in the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern, will be greater for participants who belong to locomotive crews, compared to MOW crews.

Similarly, the extent to which a crew member has taken on the roles of other members of the crew can also affect pre-training perceptions of task interdependence. Interpositional experience should increase one's interpositional knowledge, thereby increasing the accuracy and similarity of that crew member's perception of interdependence. This is consistent with the previously reviewed cross-training literature, which has shown that cross-training increases interpositional knowledge (Cooke et al., 2003; Marks et al., 2002; Salas et al., 1992). Because of the resultant higher levels of interpositional knowledge, crew members who have held other jobs within their crew will most likely have more accurate and similar perceptions of team-relatedness and team workflow pattern prior to training. That is, CRM training should result in an increase in the accuracy and similarity of participants' perceptions of task interdependence for all levels of interpositional experience. However, as a result of posited pre-training differences between participants with higher vs. lower levels, interpositional experience will moderate the effect of CRM training on crew members' perceptions of task interdependence. Consequently, it was hypothesized that:

Hypothesis 4: There will be a significant main effect for training on perceptions of task interdependence. Specifically, when crossed with interpositional experience the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be higher at Time 2 (immediate post-training) than at Time 1 (pre-training).

Hypothesis 5: There will be a significant simple effect for interpositional experience on perceptions of task interdependence at Time 1. Specifically, at Time 1 the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern, will be greater for participants who have higher levels of interpositional experience, compared to participants with lower levels of interpositional experience.

Hypothesis 6: There will be a significant interaction between training and interpositional experience on perceptions of task interdependence. Specifically, from Time 1 to Time 2, increases in the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be greater for participants who have lower levels of interpositional experience, compared to participants with higher levels of interpositional experience.

Effect of CRM on traditional training criteria. Training evaluation is "the systematic collection of descriptive and judgmental information necessary to make effective training decisions related to the selection, adoption, value, and modification of various instructional activities" (Goldstein, 1991, p. 557). Training evaluation is considered essential to any training initiative and the most common framework for characterizing the different types of evaluation data is Kirkpatrick's (1976) typology, which consists of reaction, learning, behavioral, and results criteria. The concept of Kirkpatrick's (1976) learning level has been expanded to include cognitive learning outcomes, skill-based outcomes, and affective outcomes (Kraiger et al., 1993). Meta-analysis showing variable effect sizes for different types of criteria demonstrates the criticality of understanding the fundamental differences in evaluation criteria (Arthur et al., 2003).

As a result of CRM's implementation in other high-reliability industries outside commercial aviation, many authors have suggested that CRM be evaluated in those industries such that its inter-industry effectiveness can be documented (Salas, Rhodenizer, & Bowers, 2000; Sexton, Thomas, & Helmreich, 2000). Furthermore, it has been suggested that non-aviation research employ the same evaluation criteria and measures as commercial aviation research, so that comparisons across industries can be made (see Sexton et al., 2000). CRM awareness training has traditionally been evaluated in the commercial airline industry using reaction, cognitive learning (knowledge), and affective learning (attitudes) outcomes (Salas et al., 2001). Similar to findings in commercial aviation (Salas et al., 2001), research has shown that there are positive reactions, declarative knowledge acquisition, and changes in attitudes toward teamwork in the off-shore oil drilling (O'Connor & Flin, 2003), aircraft maintenance (Taylor & Patankar, 2000), military aviation (Salas et al., 1999; Stout et al., 1997), nuclear power (Harrington & Kello, 1993), and healthcare industries (Howard et al., 1992). However, although it has been implemented in the railroad industry for at least eight years (on a limited basis), there are no evaluations of its effectiveness in this industry (Morgan et al., 2003). Thus the second objective of this paper was to examine the effect of CRM training in the railroad industry, using these traditional measures of CRM training effectiveness.

Reaction criteria represent trainees' feelings toward the training program itself and can include affective reactions to the course content or instructors, as well as judgments as to the utility of the training program for changing on-the-job behavior. Research suggests that participants in the commercial airline industry like CRM training, and believe it to be worthwhile, useful, and applicable in their respective industries (Salas et al., 2001). Consequently, it was hypothesized that:

Hypothesis 7: Participants will have positive reactions to CRM training.

Researchers have recognized two distinct types of team knowledge, labeled "taskwork" and "teamwork," that must be both accurate and shared by team members for effective team performance (Cannon-Bowers & Salas, 2001). Klimoski and Mohammed (1994) state that "models of task work refer to an understanding of the activities and action sequences that both the individual and the team collectively must carry out to perform the team task" (p. 416). Using this conceptualization, crew members' perceptions of task interdependence is a type of "taskwork" knowledge (Arthur et al., 2005; Cannon-Bowers & Salas, 2001). Comparatively, "teamwork" knowledge contains information about performing teamwork behaviors (e.g., communication, performance monitoring, back-up behaviors) which are almost always necessary to coordinate and effectively manage the flow of resources and information through the team (Rentsch & Hall, 1994). Specifically, theories of team performance suggest that team members must have accurate and similar knowledge regarding teamwork, including not only what it is and how it is performed, but also its importance for team functioning and performance (Cannon-Bowers & Salas, 2001).

Interestingly, the effect of CRM awareness training on crew members' understanding of the criticality of specific teamwork behaviors and the consequence of those behaviors on task performance (i.e., "teamwork" knowledge) is the most commonly studied effect in the CRM literature. Although some specific teamwork behaviors may be important for specific teams (McIntyre & Salas, 1995), the literature suggests that in most team tasks, behaviors related to adapting, communicating, coordinating, decision-making, interpersonal relations, leadership, shared situational awareness, and assertiveness are critical (Bowers et al., 1993; Burke et al., 2003).

Conceptualized as a type of "teamwork" knowledge, crew members' perceptions of the criticality of these behaviors is most commonly measured using the "Cockpit Management Attitudes Questionnaire" (CMAQ; Helmreich & Wilhelm, 1991; Helmreich, Wilhelm, & Gregorich, 1988).

Although purported in the CRM literature as a measure of "attitudes" toward teamwork, the CMAQ is clearly measuring team members' knowledge or perceptions of what teamwork behaviors are necessary, how they operate, and their importance in task completion. Evidence suggests that CRM training positively influences cockpit crew members' perceptions of teamwork (Salas et al., 2001). Specifically, CRM training improves perceptions regarding the importance of communication and coordination, command responsibility, and recognizing the effects of stressors (Gregorich, Helmreich, & Wilhelm, 1990; Helmreich & Wilhelm, 1991). Similarly, CRM has been shown to improve crew members' perceptions of teamwork in other high-reliability industries (Flin & O'Connor, 2001; Howard et al., 1992). Because of the similarities of the teams, tasks, and awareness-training environment in the railroad industry and other high-reliability industries (Roop et al., 2007), CRM training should similarly increase the accuracy and similarity of railroad crew members' perceptions of teamwork behaviors.

As previously noted, research suggests that the effectiveness of CRM in changing knowledge is dependent on the accuracy and similarity of that knowledge prior to training (Brun et al., 2005). Crew members who more accurately and similarly perceive the interdependence of their tasks (highly interdependent) should also more accurately and similarly perceive the importance of specific teamwork behaviors in completing those highly interdependent tasks. It was previously noted that compared to locomotive crew members and those with lower levels of interpositional experience, MOW crews and crew members with higher levels of interpositional experience should have more accurate and similar perceptions of interdependence prior to CRM training. Therefore, similar to earlier hypotheses, it is posited that crew type and interpositional experience will moderate the effect of CRM training on the accuracy and similarity of participants' perceptions of teamwork. Consequently, it was hypothesized that:

Hypothesis 8: There will be a significant main effect for training on perceptions of teamwork. Specifically, when crossed with crew type the (a) accuracy and (b) similarity of participants' perceptions of teamwork will be higher at Time 2 (immediate post-training) than at Time 1 (pre-training).

Hypothesis 9: There will be a significant simple effect for crew type on perceptions of teamwork at Time 1. Specifically, the (a) accuracy, and (b) similarity of participants' perceptions of teamwork at Time 1 will be greater for participants who belong to MOW crews compared to locomotive crews.

Hypothesis 10: There will be a significant interaction between training and crew type on perceptions of teamwork. Specifically, increases in the level of (a) accuracy and (b) similarity of participants' perceptions of teamwork from Time 1 to Time 2 will be greater for participants who belong to locomotive crews compared to MOW crews.

Hypothesis 11: There will be a significant main effect for training on perceptions of teamwork. Specifically, when crossed with interpositional experience the (a) accuracy and (b) similarity of participants' perceptions of teamwork will be higher at Time 2 than at Time 1.

Hypothesis 12: There will be a significant simple effect for interpositional experience on perceptions of teamwork at Time 1. Specifically, the (a) accuracy, and (b) similarity of participants' perceptions of teamwork at Time 1 will be greater for participants who have higher levels of interpositional experience, compared to participants with lower levels of interpositional experience.

Hypothesis 13: There will be a significant interaction between training and interpositional experience on perceptions of teamwork. Specifically, increases in the level of (a) accuracy and (b) similarity of participants' perceptions of teamwork from Time 1 to Time 2 will be greater for participants who have lower levels of interpositional experience, compared to participants with higher levels of interpositional experience.

Long-Term Retention

One key aspect of training evaluation is what happens to acquired knowledge or skills after the training is completed and the trainee returns to the real work environment. The extant literature has shown that skills decay over periods of non-use (Arthur, Bennett, Stanush, & McNelly, 1998), and an abundance of literature in the field of cognition and memory suggests knowledge decays when not rehearsed in some way (see Brannon & Koubek, 2001; Ruben & Wenzel, 1996). In an applied setting, knowing the rate of decay of critical knowledge and skills is important in determining the time frame for refresher training (Childs & Spears, 1986; Roop et al., 2007). The three types of cognitive training criteria (Kraiger et al., 1993) used in the present research that are relevant to long-term retention are (a) declarative knowledge regarding the specific CRM course content, (b) the accuracy and similarity of perceptions of task interdependence and (c) the accuracy and similarity of perceptions of teamwork. As previously noted, although there has been quite a bit of research evaluating the retention of declarative knowledge, there is very little research assessing the degree to which the accuracy and similarity of team members' perceptions of task interdependence or perceptions of teamwork are retained over time.

Interestingly, research suggests that there are a variety of factors that can affect the rate of skill decay (Arthur et al., 1998). One factor that seems particularly relevant in this research is the degree to which the work environment supports or reinforces accurate and similar perceptions of task interdependence and teamwork. It is clear from the aviation CRM literature that without organizational reinforcement, commitment, and support, the learning that takes place during CRM training is not retained (Helmreich & Foushee, 1993; Helmreich et al., 1999). As previously suggested, the accuracy and similarity of pre-training perceptions of task interdependence and teamwork are expected to be low because of various factors in the railroad industry including training practices, unionization, and an overall organizational climate which are not supportive of CRM initiatives (Roop et al., 2007). Unfortunately, these factors will continue to be in place post-training. Furthermore, because the movement of freight is often unpredictable, railroads use a labor pool of engineers and conductors to form locomotive crews on a "just-in-time" basis (Morgan et al., 2003). That is, similar to crews in other high-reliability industries, locomotive crew membership is extremely dynamic. Because only a small fraction of engineers and conductors in the host railroad's labor pool will be trained in CRM as part of this research, after training, locomotive crews will undoubtedly be formed by pairing CRM trained and non-CRM trained crew members. These non-trained crew members will lack knowledge of CRM, which will most likely result in a lack of support for trained crew members' CRM-related perceptions, knowledge, and skills. As a result of this lack of environmental support for CRM, it is likely that the accuracy and similarity of perceptions of both task interdependence and teamwork will decay over time. Therefore, it seems likely that without rehearsal or regular use of CRM knowledge, it will also decay. Consequently, it is hypothesized that:

Hypothesis 14: Participants' knowledge of CRM concepts will be higher at Time 2 (immediate post-training) than at Time 3 (2-years post-training).

Similar to earlier hypotheses, it is posited that crew type and interpositional experience will moderate the extent to which the accuracy and similarity of perceptions of task interdependence and teamwork are retained after CRM training. That is, in the same manner that specific characteristics of the railroad could negatively affect the accuracy and similarity of crew members' perceptions of task interdependence and teamwork, the characteristics could affect long-term retention. MOW crews should experience more environmental support for perceiving the task as highly interdependent and requiring specific teamwork skills. For instance, unlike locomotive crews, which are created on a "just-in-time" basis, MOW crews have a relatively fixed membership and are trained as intact crews. This will most likely result in higher levels of within-crew communication and feedback regarding CRM-related perceptions, knowledge, and skills in MOW crews compared to locomotive crews. Similarly, having high levels of interpositional experience should act as an inoculation against negative environmental characteristics like a lack of supervisory support. So, on the basis of the preceding arguments, it was hypothesized that:

Hypothesis 15: There will be a significant main effect for time on perceptions of task interdependence. Specifically, when crossed with crew type the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be higher at Time 2 (immediate post-training) than at Time 3 (2-years post-training).

Hypothesis 16: There will be a significant interaction between time and crew type on perceptions of task interdependence. Specifically, from Time 2 to Time 3, decreases in the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be greater for participants who belong to locomotive crews compared to MOW crews.

Hypothesis 17: There will be a significant simple effect for crew type on perceptions of task interdependence at Time 3. Specifically, at Time 3 the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be greater for participants who belong to MOW crews compared to locomotive crews.

Hypothesis 18: There will be a significant main effect for time on perceptions of task interdependence. Specifically, when crossed with interpositional experience the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be higher at Time 2 than at Time 3.

Hypothesis 19: There will be a significant interaction between time and interpositional experience on perceptions of task interdependence. Specifically, from Time 2 to Time 3, decreases in the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be greater for participants who have lower levels of interpositional experience, compared to participants with higher levels of interpositional experience.

Hypothesis 20: There will be a significant simple effect for interpositional experience on perceptions of task interdependence at Time 3. Specifically, at Time 3 the (a) accuracy and (b) similarity of participants' perceptions of team-relatedness and (c) accuracy and (d) similarity of participants' perceptions of team workflow pattern will be greater for participants who have higher levels of interpositional experience, compared to participants with lower levels of interpositional experience.

Hypothesis 21: There will be a significant main effect for time on perceptions of teamwork. Specifically, when crossed with crew type the (a) accuracy and (b) similarity of participants' perceptions of teamwork will be higher at Time 2 than at Time 3.

Hypothesis 22: There will be a significant interaction between time and crew type on perceptions of teamwork. Specifically, decreases in the level of (a) accuracy and (b) similarity of participants' perceptions of teamwork from Time 2 to Time 3 will be greater for participants who belong to locomotive crews compared to MOW crews.

Hypothesis 23: There will be a significant simple effect for crew type on perceptions of teamwork at Time 3. Specifically, the (a) accuracy, and (b) similarity of participants' perceptions of teamwork at Time 3 will be greater for participants who belong to MOW crews compared to locomotive crews.

Hypothesis 24: There will be a significant main effect for time on perceptions of teamwork. Specifically, when crossed by interpositional experience the (a) accuracy and (b) similarity of participants' perceptions of teamwork will be higher at Time 2 than at Time 3.

Hypothesis 25: There will be a significant interaction between time and interpositional experience on perceptions of teamwork. Specifically, decreases in the level of (a) accuracy and (b) similarity of participants' perceptions of teamwork from Time 2 to Time 3 will be greater for participants who have lower levels of interpositional experience, compared to participants with higher levels of interpositional experience.

Hypothesis 26: There will be a significant simple effect for interpositional experience on perceptions of teamwork at Time 3. Specifically, the (a) accuracy, and (b) similarity of participants' perceptions of teamwork at Time 3 will be greater for participants who have higher levels of interpositional experience, compared to participants with lower levels of interpositional experience.

METHOD

Participants

The study sample consisted of 160 railroad employees including 85 locomotive and 75 MOW crew members working for a Class 1 U.S. railroad. Locomotive crews included engineers and conductors while MOW crews (from six divisions) included foremen, assistant foremen, trackmen (laborers), machine operators, welders, and speed swing operators. The mean age of the study sample was 45 years (SD = 10.80, range = 18 - 65) and mean tenure with the railroad was 18.90 years (SD = 12.39, range = 3 mo. - 36 yrs). All but two participants were male. Due to substantial sample size differences for specified analyses, power calculations for each analyses can be found in Appendix A.

Training

The Texas Transportation Institute (TTI) developed the railroad CRM training program used in this study. Members of the development team included researchers with backgrounds in the rail industry, CRM training in naval aviation, industrial/organizational psychology, and training development, delivery, and evaluation. The content of the training course was initiated first by the National Transportation Safety Board (NTSB) recommendation following the Butler Indiana collision and derailment (NTSB, 1999a). The NTSB (1999b) recommended that the American Short Line and Regional Railroad Association, “develop, for all train crewmembers, train crew resource management training that addresses, at a minimum: crewmember proficiency, situational awareness, effective communication and teamwork, and strategies for appropriately challenging and questioning authority” (p. 3).

With this recommendation as a basis, TTI used a variety of sources and methods to determine the most relevant and effective content to be covered in the current railroad CRM program. This included reviewing CRM training materials from various non-railroad industries (commercial aviation, military, medical, maritime and nuclear power) as well as the FAA's Crew Resource Management Advisor Circular (FAA, 2004), which recommends content areas for commercial airlines.

Although there are many similarities between the types of tasks and crews that work in other industries (Roop et al., 2007), steps were taken to ensure that the training was specifically targeted to and for railroad crews. These steps included reviewing the specific causes of human error railroad accidents, discussions with FRA safety officials and railroad employees, and the examination of previous railroad CRM programs. The current railroad CRM program centered around five modules which included "technical proficiency," "situational awareness," "communication," "teamwork," and "assertiveness." Furthermore, crew specific training materials were developed for both locomotive and MOW crews.

Training was delivered in a one-day course using a combination of lecture, group discussion, case studies, and video. One of the critical aspects of this training was its use of real accident scenarios and exercises, showing how a breakdown in each key area of CRM (e.g., technical proficiency, situational awareness) has caused a serious accident in the past. In these exercises, participants read a 1-2 page summary of a real railroad human error accident that has been investigated by the NTSB. Next, participants were put into groups of 3-4 and discussed a series of questions posed to them at the end of the scenarios. These questions were related to determining who was part of the team, what the human errors were, as well as steps that could have been taken to prevent the accident. The developers of the CRM program also served as the facilitators.

Measures

Perceptions of task interdependence. Trainees' perceptions of task interdependence were measured using team-relatedness and team workflow pattern scales adapted from Arthur et al. (2005). The team-relatedness measure asked participants to identify the extent to which successful team performance requires them to work with members of the crew in order to optimally perform their overall job. Participants were asked to indicate their answer using a five point Likert scale ranging from 1 ("not required to work with crew members at all for optimal performance") to 5 ("very much required to work with crew members for optimal performance"). The team-relatedness scale is presented in Appendix B.

The team workflow pattern measure defined team workflow pattern as "the paths by which work and/or information flows through the crew in order to optimally perform your job." Next, participants were given descriptions and illustrations of five different workflow patterns. Trainees were then asked to rank order the five patterns (1 = high, 5 = low) in terms of the extent to which they are descriptive of the way that work between crew members flows for the optimal performance of their job as a whole. The team workflow pattern scale is presented in Appendix C.

Railroad management attitudes questionnaire (RMAQ). The 28-item questionnaire was used to assess participants' perceptions of teamwork. The questionnaire was based on the Cockpit Management Attitudes Questionnaire (CMAQ) developed by Helmreich and his colleagues (Helmreich, 1984; Helmreich et al., 1988). The three separate teamwork factors measured by the CMAQ are (1) communication and coordination, (2) command responsibility, and (3) recognition of stressor effects (Gregorich et al., 1990). Because leadership and command responsibility in railroad crews is different from those of cockpit crews, this content was not trained in the current railroad CRM training, and thus items pertaining to "command responsibility" were not included in the RMAQ. Likewise, the wording of other items from the CMAQ was slightly modified to make them appropriate for the railroad industry. Finally, additional items that reflect perceptions of specific teamwork behaviors addressed in the current railroad CRM training program that are not in the original CMAQ were added. The RMAQ is presented in Appendix D.

Trainee reactions. A four-item measure using a 7-point Likert-type scale (1 = strongly disagree and 7 = strongly agree) was used to assess trainee reactions to the CRM training. Items assessed the extent to which trainees agreed with the following statements: (a) I found this training to be enjoyable, (b) the training was job relevant, and (c) the training had practical value. Furthermore, the questionnaire asked trainees, "to what degree will this training influence your ability later to perform your job." A 7-point Likert-type scale (1 = no influence and 7 = strongly influence) was also used for this question.

Knowledge test. A 24-item knowledge test was developed and administered to participants to assess their mastery of training concepts that were taught during the CRM course. Nine questions were in multiple-choice format. For example, participants were asked to identify which railroad craft has the highest rate of fatalities a year. The knowledge test also included 10 true/false items. An example of this type of item was, "New equipment can change a crew members' technical proficiency." Lastly, trainees were asked five open-ended questions — for example, "What are two specific characteristics of railroading that could potentially lead to fatigue?" Participants' scores on the knowledge test equaled the number of correct responses. Open ended questions were scored 1 point per correct item.

Interpositional experience. Participants were asked to indicate if they have previously held any other jobs on a railroad. Participants who had held other crew member positions within their current crew were labeled "higher" interpositional experience, while those who had not held other crew member positions within their current crew were labeled "lower" interpositional experience.

Design and Procedures

Pre and immediate post-training data were collected over a 1-year period during 15 initial training sessions. Class sizes ranged from 3 to 23 participants with an average class size of 12 (SD = 6.73). Prior to training, all participants completed the RMAQ, team-relatedness, and team workflow pattern measures. Participants then completed the day long, 6-7 hour railroad CRM training course. Immediately after training, participants completed the "reactions" to training measure, RMAQ, team-relatedness and team workflow pattern measures, and knowledge test. In addition, participants provided demographic information such as job title, age, other jobs they have held on a railroad, and years of service. Participants also indicated if they had previous CRM training or training in any of the CRM content areas (communication, situational awareness, teamwork, assertiveness). Follow-up surveys were sent out 2 years after training. Follow-up surveys asked participants to again complete the RMAQ, team-relatedness and team workflow pattern measures, and knowledge test.

Data Analysis

Operationalization of accuracy. This study assessed accuracy at the individual level. Thus, each participant had an accuracy score, which represented the similarity between that participant and an expert referent. To assess the accuracy of participants' perceptions, I used the similarity index *D* (Cronbach & Gleser, 1953) which represents the Euclidean distance between a participant's profile score (i.e., score on a measure) and that of the expert referent.[1] The formula for computing *D* is presented in Equation 1:

$$D = \sqrt{\sum \left(X_i - Y_i \right)^2} \qquad (1)$$

where:

X	=	an individual's score on a particular element
Y	=	a referent score on the same element

Using the equation above, the accuracy of a participant's perception of his or her team workflow pattern was represented by the square root of the sum of the squared differences between that participant's ranking of a particular workflow pattern and that of the expert referent's ranking of that pattern for each of the five team workflow patterns. The accuracy of a participant's team-relatedness score, represented by *D*, is the square root of the sum of the squared difference between that participant's rating of team-relatedness and that of the expert referent. Because team-relatedness is a 1-item scale, *D* for this scale is simply the absolute value of the difference between the participant's and expert's rating. Representing the accuracy of a participant's perceptions of teamwork, *D* is the square root of the sum of the squared differences between a participant's rating on an item and that of the expert referent rating for that same item, for each of the 28 items in the RMAQ. Similar to other studies, subject matter experts were used to derive the expert referent (Goldsmith, Johnson, & Acton, 1991; Kraiger, Salas, & Cannon-Bowers, 1995).

Operationalization of similarity. Similarity was also assessed at the individual level. As with accuracy, *D* was used as an index of similarity regarding participants' perceptions of team workflow pattern, team-relatedness, and teamwork. However, unlike the index representing accuracy, this *D* represents the similarity between a participant and that participant's referent group. Because there are no "intact" crews in either the transportation or engineering branches of the railroad, a participants' referent group consisted of other participants with whom they would have worked with for *at least* one shift during the 12 months following training. Because participants were from the same host railroad, the referent group for any specific engineer was all conductors, while a specific conductor's referent group encompassed all engineers. As stated earlier, MOW participants from six different divisions participated in this study, and unlike locomotive participants, MOW participants work only within their particular division. Thus, a specific MOW participant's referent group encompassed not all other MOW participants, but only those in his or her particular division.

Creation of Expert Referent Scores

Three experts in railroad safety and CRM were recruited to provide the expert referent scores used to compute participants' accuracy scores (i.e., *D*). Each expert completed the team-relatedness and team workflow pattern measures and the RMAQ from the perspective of a locomotive crew member and once again from the perspective of a MOW crew member. There was complete agreement between all three experts for the team-relatedness and team workflow pattern scales for both the locomotive and MOW versions. Experts agreed that for both locomotive and MOW crew members, the accurate level of team-relatedness was "very much required to work with other members of their crew for optimal performance" (represented by a rating of 5 on a 5-point Likert scale). Additionally, experts had complete agreement on the rankings of the various team workflow patterns. Experts indicated that an intensive workflow pattern best described the work activities of the crews, followed by reciprocal, sequential, and additive. Finally, they agreed that "not a team job" was the least descriptive of crew members' work activities.

Analysis of the 28-item RMAQ revealed 3 items on which one expert's response was at the opposite end of the Likert scale than the other two experts. These items were removed from the scale and the subsequent expert referent scores were computed on each of the remaining 25 items. Assessing absolute agreement, the intraclass correlation coefficient (Shrout & Fleiss, 1979) for both the locomotive RMAQ and the MOW RMAQ was .99, indicating almost perfect agreement. Thus an expert referent score for each item on the RMAQ was formed by averaging experts' ratings for that specific item.

Team Workflow Pattern: Conversion From a Normative to an Ipsative Scale

As detailed earlier and shown in Appendix C, instructions to the team workflow pattern scale asked participants to rank order the workflow patterns in terms of the extent to which they were descriptive of the work activities in their job. Rank orders of this type are ipsative, meaning they are multiple item measures where data are collected (or are modified) in such a way that all participant totals across the items are equal. Unfortunately, some participants misunderstood the instructions and rated each workflow pattern individually on a scale of 1 to 5, resulting in a normative, not an ipsative rank-ordered scale. Of the 75 participants who completed the team workflow pattern measure at Time 1, 45 (60%) completed it correctly (consistent with the instructions). Likewise, 46 (56%) of the 82 participants who completed the team workflow pattern scale at Time 2, completed it correctly. In order to utilize the data from participants who completed the measure incorrectly, their normative ratings were converted to an ipsative scale. Examples of this modification can be seen in Appendix E. This allowed the use of data for all participants who completed the team workflow pattern measure. Accordingly, all descriptive statistics presented and analyses performed utilize these corrected measures. Parallel analyses were also performed using only those participants who filled out the team workflow pattern scale correctly. When divergent results were obtained using only this subset of participants, they are indicated with an endnote.

Unequal (but proportional) Cell Sizes

Due to the nature of collecting data in the field and constraints and circumstances associated with the client organization, it was not possible to present all measures or collect all information from every participant. Furthermore, the examination of moderator variables based on pre-existing group membership (i.e., crew type, interpositional experience) coupled with the experimenter's lack of control over the makeup of training sessions resulted in unequal cell or group sizes. Although cell sizes did not differ substantially for most analyses, large differences in sample sizes were present between crew types for dependent variables pertaining to perceptions of task interdependence (e.g., DV = accuracy of perceptions of team-relatedness; locomotive [n = 8], MOW [n = 55]). Under most circumstances, when equal cell sizes had been planned yet not achieved due to unsystematic events, unweighted means analysis (i.e., cell means have equal weight) is appropriate (Bonett, 1982; Green, Heiberger & Laster, 1976). Although a 3-way repeated measures design (Time × Crew Type × Interpositional Experience) would be appropriate for the current study, the presence of unequal sample sizes within both moderating variables, coupled with the very small sample size in the locomotive condition, resulted in major aberrations when utilizing unweighted means in such a design. Therefore, hypotheses were tested in the current study by examining each moderating variable separately using 2-way mixed repeated measures analyses (i.e., Time × Crew Type; Time × Interpositional Experience).

RESULTS

Correlations Among the Dependent Measures

Table 1 shows the correlations among all variables including the dependent measures at both Time 1 and 2. Although there were no a priori hypotheses regarding the relationships between dependent measures, several interesting results are apparent in Table 1. Results show moderately high positive correlations between the Time 1 and Time 2 measures for each dependent variable (ranging from .33 – .79; $p < .01$ for all correlations). Table 1 also reveals moderately high positive correlations between the accuracy and similarity indices for each dependent variable (ranging from .47 – .82; $p < .01$ for all correlations). Most interesting, however, is the pattern of correlations between perceptions of team-relatedness, team workflow pattern, and teamwork. The correlations between indices of participants' perceptions of team-relatedness and team workflow pattern were positive, yet uniformly low (ranging from .04 – .25; $p < .05$ for correlation between the accuracy of team-relatedness and team workflow pattern at Time 2 only). Furthermore, although the correlations between perceptions of team-relatedness and team workflow pattern were positive and low, overall, they were higher than the positive correlations between either perceptions of team-relatedness or team workflow pattern, and perceptions of teamwork (ranging from .03 – .13; $p > .05$ for all). In fact, there were negative correlations between the similarity of participants' perceptions of team workflow pattern and teamwork at Time 1 ($r = -.10$; $p > .05$), and Time 2 ($r = -.24$; $p < .05$). Furthermore, there was a negative correlation between the similarity of participants' perceptions of team-relatedness and teamwork at Time 1 ($r = -.05$; $p > .05$).

Table 1

Descriptive Statistics and Correlations Among Variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
IV's														
1. Crew Type[A]	—													
2. IPE[B]	.13	—												
Time 1														
3. Acc TR	.07	.09	—											
4. Sim TR	-.15	.09	.47**	—										
5. Acc TWFP	.00	.01	.22	.19	—									
6. Sim TWFP	-.06	-.06	.27*	.20	.82**	—								
7. Acc RMAQ	-.17*	.05	.13	.10	.05	.04	—							
8. Sim RMAQ	.04	.00	.05	.04	-.13	-.10	.58**	—						
Time 2														
9. Acc TR	.22	.16	.79**	.32*	.08	.13	.05	.09	—					
10. Sim TR	.19	.27*	.35**	.60**	.07	-.01	.15	-.08	.54**	—				
11. Acc TWFP	-.16	-.12	.17	.22	.60**	.47**	-.15	-.20	.25*	.19	—			
12. Sim TWFP	.00	-.09	-.07	.02	.31*	.33**	-.20	-.29*	.02	.04	.53**	—		
13. Acc RMAQ	.05	.09	.12	.11	.14	.12	.54**	.23*	.11	.11	.03	-.10	—	
14. Sim RMAQ	.07	.03	.05	-.03	-.11	-.08	.23*	.45**	.12	-.05	-.06	-.24*	.62**	—
Mean	—	—	1.10	1.31	2.70	2.78	11.34	8.47	0.81	0.89	3.23	2.80	10.11	7.99
SD	—	—	1.35	0.87	1.84	0.87	2.63	2.05	1.26	0.79	2.20	0.84	2.99	2.41

Note. [A]Crew Type is coded 1 = locomotive, 2 = MOW. [B]IPE is coded 0 = lower, 1 = higher. IPE = interpositional experience; Acc = accuracy; Sim = similarity; TR = perceptions of team-relatedness; TWFP = perceptions of team workflow pattern; RMAQ = perceptions of teamwork.

$^*p < .05$, $^{**}p < .01$.

These correlations suggest that as measures of participants' perceptions of taskwork (Arthur et al., 2005; Cannon-Bowers & Salas, 2001), team-relatedness and team workflow pattern are distinct yet related aspects of task interdependence, and show a uniformly low (and sometimes negative) relationship with participants' perceptions of teamwork.

Hypotheses Pertaining to Perceptions of Task Interdependence (Hypotheses 1 - 6)

A summary of the results for the research hypotheses pertaining to the effect of training on participants' perceptions of task interdependence are presented in Table 2.

Effect of training and crew type on perceptions of task interdependence (Hypotheses 1 - 3). As stated previously, equal cell sizes had been planned yet not achieved due to unsystematic events, and in these circumstances, research suggests using unweighted means analysis which weights the group mean equally regardless of the sample size obtained (Bonett, 1982). However, the very small sample of locomotive participants with data regarding perceptions of task interdependence (avg. $n = 8$) is a cause for concern. This is because analyses based on these data are susceptible to the standard problems with small sample sizes (i.e., variance of the sampling distribution and low power) which are statistically masked by the larger sample in the overall unweighted means analyses. Although such a small sample size might cause a reader hesitation in terms of statistical validity, for the sake of completeness they are presented in the standard statistical analyses below. That is, the descriptive statistics, main effects, and interactions regarding perceptions of task interdependence that also include the independent variable Crew Type (i.e., Hypotheses 1 - 3) should be viewed as exploratory in nature. Later discussions regarding these findings should also be viewed as exploratory and future research should be conducted to determine if results are replicated using a larger locomotive sample.

Table 2

Hypotheses Pertaining to the Effect of Training on Perceptions of Task Interdependence

Hypothesis and dependent variable	Level of independent variable and proposed effect on	Support
Hypothesis 1 - main effect for Training (when crossed with Crew Type)		
ACC Team-Relatedness	Time 1 < Time 2	No[a]
ACC Team Workflow Pattern	Time 1 < Time 2	No[a]
SIM Team-Relatedness	Time 1 < Time 2	Yes[a]
SIM Team Workflow Pattern	Time 1 < Time 2	No[a]
Hypothesis 2 - simple effect for Crew Type (at Time 1)		
ACC Team-Relatedness	MOW > LOC	No[a]
ACC Team Workflow Pattern	MOW > LOC	No[a]
SIM Team-Relatedness	MOW > LOC	No[a]
SIM Team Workflow Pattern	MOW > LOC	No[a]
Hypothesis 3 – interaction (Training × Crew Type)		
Increase in ACC Team-Relatedness from Time 1 to Time 2	LOC > MOW	No[a]
Increase in ACC Team Workflow Pattern from Time 1 to Time 2	LOC > MOW	No[a]
Increase in SIM Team-Relatedness from Time 1 to Time 2	LOC > MOW	Yes[a]
Increase in SIM Team Workflow Pattern from Time 1 to Time 2	LOC > MOW	No[a]
Hypothesis 4 - main effect for Training (when crossed with IPE)		
ACC Team-Relatedness	Time 1 < Time 2	Yes
ACC Team Workflow Pattern	Time 1 < Time 2	No
SIM Team-Relatedness	Time 1 < Time 2	No
SIM Team Workflow Pattern	Time 1 < Time 2	No
Hypothesis 5 - simple effect for IPE (at Time 1)		
ACC Team-Relatedness	Higher IPE > Lower IPE	No
ACC Team Workflow Pattern	Higher IPE > Lower IPE	No
SIM Team-Relatedness	Higher IPE > Lower IPE	No
SIM Team Workflow Pattern	Higher IPE > Lower IPE	No
Hypothesis 6 – interaction (Training × IPE)		
Increase in ACC Team-Relatedness from Time 1 to Time 2	Lower IPE > Higher IPE	No
Increase in ACC Team Workflow Pattern from Time 1 to Time 2	Lower IPE > Higher IPE	No
Increase in SIM Team-Relatedness from Time 1 to Time 2	Lower IPE > Higher IPE	Yes
Increase in SIM Team Workflow Pattern from Time 1 to Time 2	Lower IPE > Higher IPE	No

Note. ACC = accuracy; SIM = similarity; LOC = locomotive crew; MOW = maintenance of way crew; IPE = interpositional experience. [a]Test (and subsequent support) is based on a small locomotive sample size (avg. n = 8) in combination with an unweighted means analysis.

The first hypothesis posited that when crossed with crew type, CRM training would increase the accuracy and similarity of participants' perceptions of both team-relatedness and team workflow pattern. Consistent with Hypothesis 1, the results presented in Table 3 indicate that, when crossed with crew type, the accuracy of participants' perceptions of team-relatedness was higher after training than before training. However, the within-subjects main effect for training did not reach conventional levels of statistical significance, $F(1, 61) = 3.15$, $p > .05$, $\eta^2 = .05$.[2] However, a mixed repeated measures analysis of variance (ANOVA) did reveal a significant main effect for training on the similarity of participants' perceptions of team-relatedness, $F(1, 60) = 9.27$, $p < .01$, $\eta^2 = .13$.[3] Consistent with Hypothesis 1, the similarity of participants' perceptions of team-relatedness was higher after training than before training. Finally, a mixed repeated measures ANOVA revealed no significant difference between Time 1 and Time 2 in terms of the accuracy, $F(1, 62) = 2.55$, $p > .05$, $\eta^2 = .04$, or the similarity $F(1, 61) = 0.01$, $p > .05$, $\eta^2 = .00$, of participants' perceptions of team workflow pattern.[4] Hence, although there was no support for Hypotheses 1a, 1c, or 1d, Hypothesis 1b was supported.

Table 3
Accuracy and Similarity of Perceptions of Task Interdependence (i.e., D) by Crew Type Across Time 1 and 2

		Time 1		Time 2		*d*
	N	Mean	*SD*	Mean	*SD*	
ACCURACY						
Team-relatedness						
Overall	63	1.10	1.35	0.81	1.26	.22
LOC	8	0.88	1.64	0.50	0.93	.29
MOW	55	1.33	1.31	1.13	1.28	.15
Team workflow pattern						
Overall	64	2.70	1.84	3.23	2.20	-.26
LOC	9	2.68	1.93	3.70	2.24	-.49
MOW	55	2.72	1.86	2.75	2.18	-.01
SIMILARITY						
Team-relatedness						
Overall	62	1.31	0.87	0.89	0.79	.51*
LOC	7	1.49	1.34	0.67	0.67	.77*
MOW	55	1.14	0.79	1.10	0.79	.05
Team workflow pattern						
Overall	63	2.78	0.99	2.80	0.84	-.02
LOC	8	2.88	1.31	2.75	1.01	.11
MOW	55	2.67	0.95	2.84	0.83	-.19

Note. Lower means indicate increased accuracy (or similarity). Positive *d* indicates increased accuracy or similarity (consistent with hypothesis). LOC = locomotive crew members; MOW = maintenance of way crew members. Overall means are unweighted. Follow-up simple effect analyses were only performed when a significant interaction was found (i.e., only for similarity of team-relatedness × crew-type).
*$p < .01$ (two-tailed).

Descriptive statistics pertaining to Hypotheses 2 and 3 are also presented in Table 3. Hypothesis 2 posited that at Time 1 (prior to training), participants who belong to MOW crews would have greater (a) accuracy and (b) similarity of perceptions of task interdependence compared to locomotive crew members. This hypothesis can be tested by running a *t*-test between locomotive and MOW crew members at Time 1, or by running a between-groups simple effect analysis at Time 1 within a Time 1 – Time 2 repeated measures design. The difference is the between-groups simple effect analysis uses only data from participants who completed the relevant measure at both Time 1 and 2. For subsequent hypotheses, I will describe results from the between-groups simple effect analyses and will indicate with an endnote divergent one-way analysis results. Thus, regarding Hypothesis 2, between-group simple effect analyses revealed that prior to training, there was no significant difference between locomotive and MOW crew members in either the accuracy of perceptions of team-relatedness, $F(1, 61) = 0.79$, $p > .05$, $\eta^2 = .01$, or team workflow pattern, $F(1, 62) = 0.00$, $p > .05$, $\eta^2 = .00$. Likewise, between-group simple effect analyses revealed that prior to training, there was no significant difference between MOW and locomotive crew members in either their within-crew similarity of perceptions of team-relatedness, $F(1, 60) = 0.99$, $p > .05$, $\eta^2 = .02$, or team workflow pattern, $F(1, 61) = 0.31$, $p > .05$, $\eta^2 = .00$. Hence, Hypothesis 2 was not supported.

Hypothesis 3 posited that increases in the levels of (a) accuracy and (b) similarity of participants' perceptions of task interdependence from Time 1 to Time 2 will be greater for participants who belong to locomotive crews compared to MOW crews. Mixed repeated measures ANOVAs indicated that crew type did not moderate the effect of training on the accuracy of crew members' perceptions of team-relatedness, $F(1, 61) = 0.29$, $p > .05$, η^2=.01, or team workflow pattern $F(1, 62) = 2.27$, $p > .05$, η^2=.04. Regarding changes in similarity however, a mixed repeated measures ANOVA revealed that the effect of training on the similarity of crew members' perceptions of team-relatedness was moderated by crew type, $F(1, 60) = 7.76$, $p < .01$, η^2=.11. This interaction is illustrated in Figure 1. A follow-up within-group simple effect analysis revealed a significant increase in the similarity of perceptions of team-relatedness for locomotive crew members, $F(1, 60) = 9.58, p < .01, \eta^2 = .14$, but not for MOW crew members, $F(1, 60) = .15, p > .05, \eta^2 = .00$. Lastly, a mixed repeated measures ANOVA indicated that the effect of training on the similarity of crew members' perceptions of team workflow pattern was not moderated by crew type, $F(1, 61) = .53, p > .05, \eta^2 = .01$.[5] Thus, although there was no support for Hypotheses 3a, 3c, or 3d, Hypothesis 3b was supported.

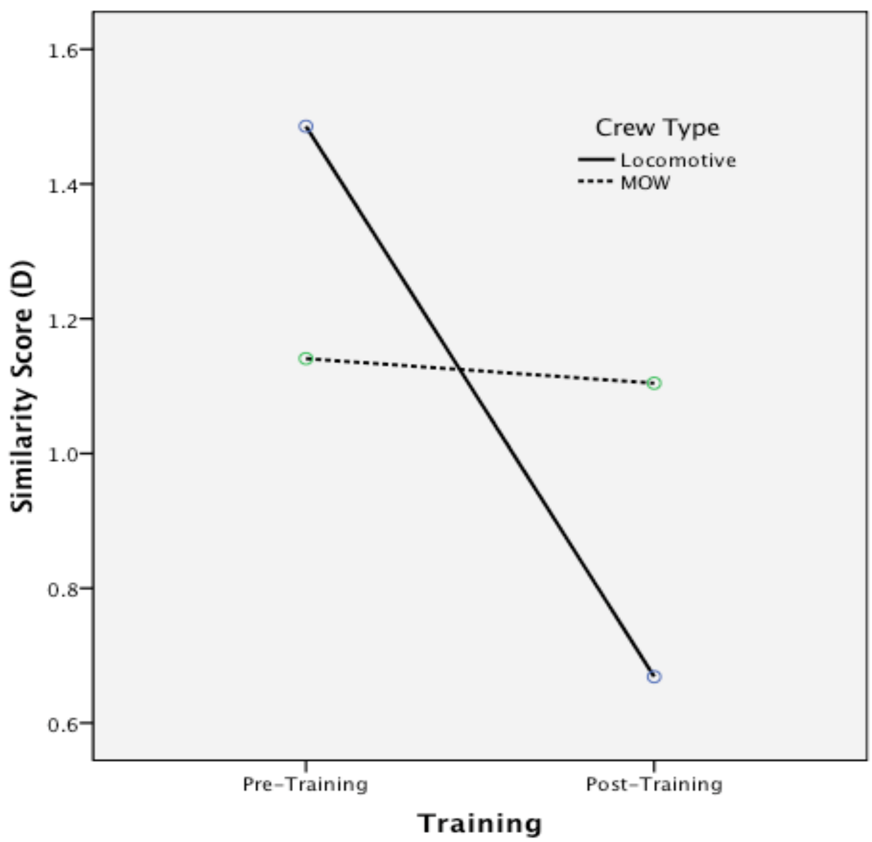


Figure 1. The effects of training and crew type on the similarity of crew members' perceptions of team relatedness. Lower score indicates increased similarity.

Effect of training and interpositional experience on perceptions of task interdependence (Hypotheses 4 - 6). Results of Hypotheses 4, 5, and 6 are presented in Table 2 and descriptive statistics pertaining to these hypotheses are presented in Table 4. Hypothesis 4 posited that when crossed with interpositional experience, CRM training would increase the accuracy and similarity of participants' perceptions of both team-relatedness and team workflow pattern. Consistent with Hypothesis 4, a mixed repeated measures analysis (across levels of interpositional experience) found a significant main effect for training on the accuracy of participants' perceptions of team-relatedness $F(1, 61) = 4.63$, $p < .05$, $\eta^2 = .07$. As predicted, perceptions were more accurate after training than before. However, a mixed repeated measures ANOVA did not find a significant main effect for training on the similarity of participants' perceptions of team-relatedness, $F(1, 60) = 2.41$, $p > .05$, $\eta^2 = .04$. Finally, a mixed repeated measures ANOVA revealed no significant difference between Time 1 and Time 2 regarding the accuracy, $F(1, 62) = 0.91$, $p > .05$, $\eta^2 = .01$, or the similarity $F(1, 61) = 0.95$, $p > .05$, $\eta^2 = .02$, of participants' perceptions of team workflow pattern. Thus, Hypothesis 4a was supported, although there was no support for Hypotheses 4b, 4c, or 4d.

Hypothesis 5 posited that the (a) accuracy and (b) similarity of participants' perceptions of task interdependence at Time 1 will be greater for participants who have higher levels of interpositional experience, compared to participants with lower levels of interpositional experience. Between-group simple effect analyses revealed that prior to training, there was no significant difference between those with lower and higher levels interpositional experience in terms of the accuracy of perceptions of team-relatedness,

$F(1, 61) = 0.12, p > .05, \eta^2 = .00$, or team workflow pattern, $F(1, 62) = 0.13, p > .05, \eta^2 = .00$. Likewise, between-group simple effect analyses revealed that prior to training, there was no significant difference between those with lower and higher levels of interpositional experience in terms of the within-crew similarity of perceptions of team-relatedness, $F(1, 60) = 0.22, p > .05, \eta^2 = .00$, or team workflow pattern, $F(1, 61) = 0.05, p > .05, \eta^2 = .00$. Thus, Hypothesis 5 was not supported.

Table 4
Accuracy and Similarity of Perceptions of Task Interdependence (i.e., D) by Interpositional Experience Across Time 1 and 2

		Time 1		Time 2		*d*
	N	Mean	*SD*	Mean	*SD*	
ACCURACY						
Team-relatedness						
Overall	63	1.27	1.34	1.03	1.24	.19*
Low IPE	29	1.21	1.29	0.86	0.95	.31
High IPE	34	1.32	1.41	1.21	1.45	.08
Team workflow Pattern						
Overall	64	2.71	1.84	2.92	2.18	-.10
Low IPE	29	2.62	1.74	3.33	2.18	-.36
High IPE	35	2.79	1.97	2.51	2.16	.14
SIMILARITY						
Team-relatedness						
Overall	62	1.18	0.86	1.03	0.77	.18
Low IPE	28	1.12	0.92	0.80	0.60	.41*
High IPE	34	1.23	0.83	1.27	0.86	-.05
Team workflow pattern						
Overall	63	2.71	0.98	2.84	0.84	-.14
Low IPE	28	2.73	1.05	2.92	0.77	-.21
High IPE	35	2.68	0.96	2.76	0.90	-.09

Note. Lower means indicate increased accuracy (or similarity). Positive *d* indicates increased accuracy or similarity (consistent with hypothesis). IPE = interpositional experience. Overall means are unweighted. Follow-up simple effect analyses were only performed when a significant interaction was found (i.e., only for similarity of team-relatedness × IPE).
*$p < .05$ (two-tailed).

Hypothesis 6 posited that increases in the levels of (a) accuracy and (b) similarity of participants' perceptions of task interdependence from Time 1 to Time 2 will be greater for participants who have lower levels of interpositional experience, compared to participants with higher levels of interpositional experience. A mixed repeated measures ANOVA indicated that the effect of CRM on the accuracy of crew members' perceptions of team-relatedness was not moderated by interpositional experience, $F(1, 61) = 1.12$, $p > .05$, $\eta^2 = .02$. However, a subsequent analysis indicated that the effect of training on the accuracy of crew members' perceptions of team workflow pattern was moderated by interpositional experience, $F(1, 62) = 4.81$, $p < .05$. $\eta^2 = .07$.[6] Follow-up within-group simple effect analyses revealed significant differences in mean change scores over time for participants with lower levels of interpositional experience, $F(1, 62) = 4.52$, $p < .05$, $\eta^2 = .07$, but not for those with higher levels of interpositional experience, $F(1, 62) = 0.85$, $p > .05$, $\eta^2 = .01$. However, as indicated by the means presented in Table 4 and the interaction illustrated in Figure 2, there was a decrease in the accuracy of perceptions of team workflow pattern for those with lower levels of interpositional experience. This is the opposite of what was predicted.

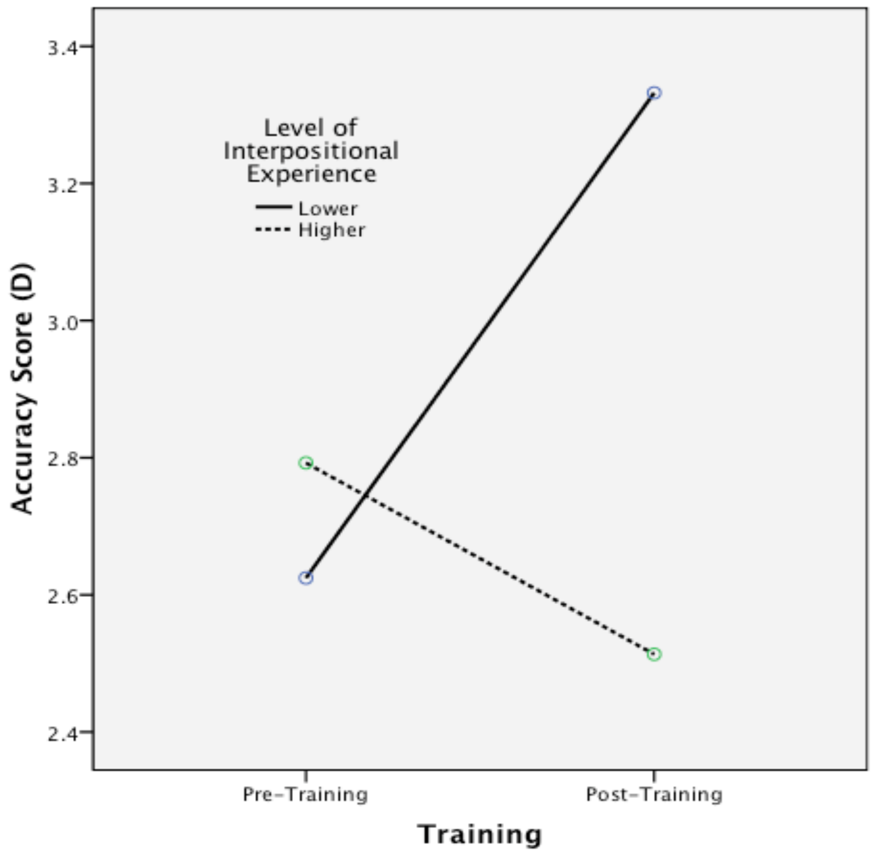


Figure 2. The effects of training and interpositional experience on the accuracy of crew members' perceptions of team workflow pattern. Lower score indicates increased accuracy.

Regarding changes in similarity, a mixed repeated measures ANOVA indicated that the effect of training on the similarity of crew members' perceptions of team-relatedness was moderated by interpositional experience, $F(1, 60) = 3.96$, $p < .05$, $\eta^2 = .06$. This interaction is illustrated in Figure 3. Follow-up within-group simple effect analyses revealed a significant increase in the similarity of perceptions of team-relatedness for those with lower levels of interpositional experience $F(1, 60) = 5.72$, $p < .05$, $\eta^2 = .09$, but not for those participants with higher levels of interpositional experience, $F(1, 60) = .11, p > .05, \eta^2 = .00$. Lastly, a mixed repeated measures ANOVA indicated that the effect of CRM on the similarity of crew members' perceptions of team workflow pattern was not moderated by interpositional experience, $F(1, 61) = 0.15, p > .05, \eta^2 = .00$. Hence, Hypothesis 6b was supported, although there was no support for Hypotheses 6a, 6c, or 6d.

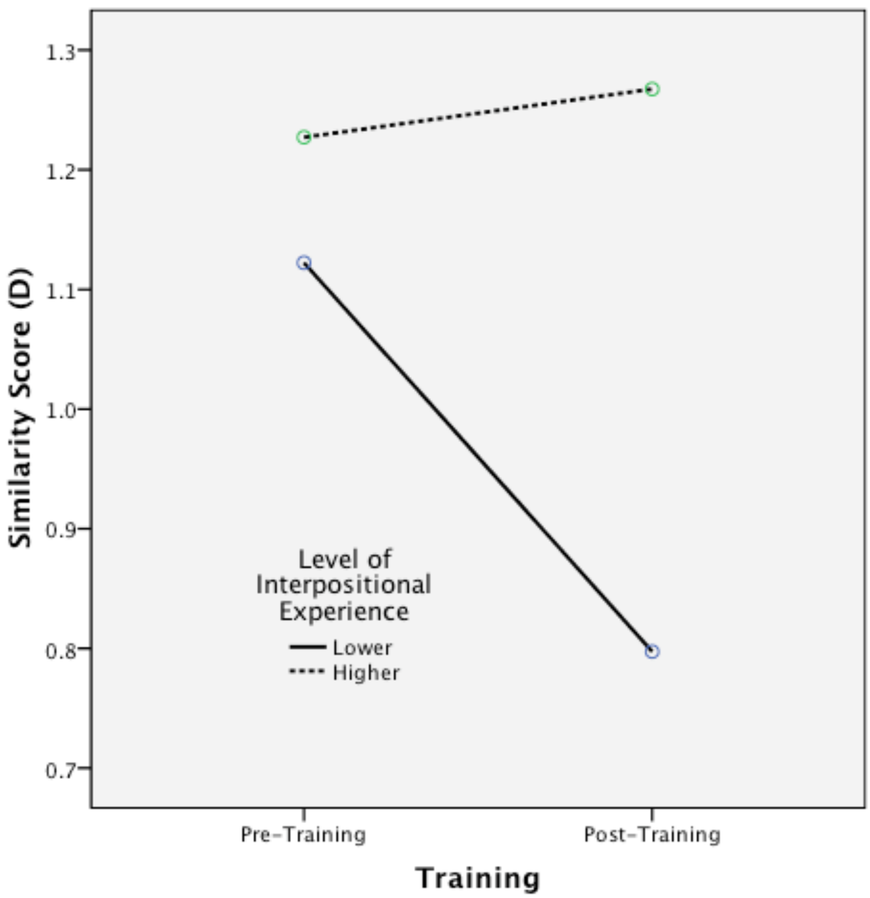


Figure 3. The effects of training and interpositional experience on the similarity of crew members' perceptions of team-relatedness. Lower score indicates increased similarity.

Trainees' Reactions to CRM Training (Hypothesis 7)

Hypothesis 7 suggested that participants would have positive reactions to CRM training. Participants indicated that they enjoyed the training, with a mean rating of 5.08 (*SD* = 1.35) out of 7. More importantly, participants indicated that they found the training to be job relevant (*M* = 5.61, *SD* = 1.43) and practical (*M* = 5.59, *SD* = 1.32). Lastly, participants indicated that CRM training "would influence" their ability to later perform their job (*M* = 5.19, *SD* = 1.27). One-sample *t*-tests on each reaction variable revealed a statistically reliable difference between each mean rating and a neutral affection rating of 4. Hence, Hypothesis 7 was supported. Follow-up analyses suggest neither crew type, interpositional experience, or crew position moderated participants' reactions to CRM training.

Hypotheses Pertaining to Perceptions of Teamwork (Hypotheses 8 - 13)

A summary of the results for the research hypotheses pertaining to the effect of CRM training on participants' perceptions of teamwork are presented in Table 5.

Table 5
Hypotheses Pertaining to the Effect of Training on Perceptions of Teamwork

Hypothesis and dependent variable	Level of independent variable and proposed effect on	Support
Hypothesis 8 - main effect for Training (when crossed with Crew Type)		
ACC Teamwork	Time 1 < Time 2	Yes
SIM Teamwork	Time 1 < Time 2	Yes
Hypothesis 9 - simple effect for Crew Type (at Time 1)		
ACC Teamwork	MOW > LOC	No
SIM Teamwork	MOW > LOC	No
Hypothesis 10 – interaction (Training × Crew Type)		
Increase in ACC of Teamwork from Time 1 to Time 2	LOC > MOW	No
Increase in SIM of Teamwork from Time 1 to Time 2	LOC > MOW	No
Hypothesis 11 - main effect for Training (when crossed with IPE)		
ACC Teamwork	Time 1 < Time 2	Yes
SIM Teamwork	Time 1 < Time 2	Yes
Hypothesis 12 - simple effect for IPE (at Time 1)		
ACC Teamwork	Higher IPE > Lower IPE	No
SIM Teamwork	Higher IPE > Lower IPE	No
Hypothesis 13 – interaction (Training × IPE)		
Increase in ACC Teamwork from Time 1 to Time 2	Lower IPE > Higher IPE	No
Increase in SIM Teamwork from Time 1 to Time 2	Lower IPE > Higher IPE	No

Note. ACC = accuracy; SIM = similarity; LOC = locomotive crew; MOW = maintenance of way crew; IPE = interpositional experience.

Effect of training and crew type on perceptions of teamwork (Hypotheses 8 - 10). Table 6 displays the means pertaining to Hypotheses 8-10. Hypothesis 8 posited that when crossed with crew type, CRM training would increase the accuracy and similarity of participants' perceptions of teamwork. Mixed repeated measures ANOVAs revealed significant main effects for training on both the accuracy of, $F(1, 138) = 28.50$, $p < .001$, $\eta^2 = .17$, and the similarity of, $F(1, 106) = 4.35$, $p < .05$, $\eta^2 = .04$, participants' perceptions of teamwork. Consistent with Hypothesis 8, both the accuracy and similarity of participants' perceptions of teamwork were higher after training than before training. Hypothesis 8 was supported.

Table 6
Accuracy and Similarity of Perceptions of Teamwork (i.e., D) by Crew Type Across Time 1 and 2

		Time 1		Time 2		*d*
	N	Mean	*SD*	Mean	*SD*	
Accuracy of Teamwork						
Overall	140	11.34	2.63	10.11	2.99	.44***
LOC	78	11.70	2.81	10.06	3.10	.55
MOW	62	10.98	2.37	10.17	2.89	.31
Similarity of Teamwork						
Overall	108	8.47	2.05	7.99	2.41	.21*
LOC	48	8.35	2.27	7.89	2.50	.19
MOW	60	8.59	1.89	8.10	2.37	.23

Note. Lower means indicate increased accuracy or similarity. Positive *d* indicates increased accuracy or similarity (consistent with hypothesis). LOC = locomotive; MOW = maintenance of way. Overall means are unweighted. No follow-up simple effect analyses were performed because no significant interactions were found.
*p < .05 (two-tailed), ***p < .001 (two-tailed).

Hypothesis 9 posited that prior to training (Time 1), participants who belong to MOW crews would have greater (a) accuracy and (b) similarity of perceptions of teamwork compared to locomotive crew members. Between-group simple effect analyses revealed that prior to training, there were no significant differences between locomotive and MOW crews regarding the accuracy, $F(1, 138) = 17.97, p > .05, \eta^2 = .02$,[7] or similarity, $F(1, 106) = 0.36, p > .05, \eta^2 = .00$, of participants' perceptions of teamwork. Hypothesis 9 was not supported.

Hypothesis 10 posited that increases in the levels of (a) accuracy and (b) similarity of participants' perceptions of teamwork from Time 1 to Time 2 will be greater for participants who belong to locomotive crews, compared to MOW crews. Mixed repeated measures ANOVAs revealed no significant interactions between training and crew type regarding either the accuracy, $F(1, 138) = 3.27$, $p > .05$, $\eta^2 = .02$, or similarity, $F(1, 106) = 0.01$, $p > .05$, $\eta^2 = .00$, of crew members' perceptions of teamwork. Hence, Hypothesis 10 was not supported.

Effect of training and interpositional experience on perceptions of teamwork (Hypotheses 11 - 13). Results of Hypotheses 11, 12, and 13 are presented in Table 5 and descriptive statistics pertaining to these hypotheses are presented in Table 7. Hypothesis 11 posited that when crossed with interpositional experience, CRM training would increase the accuracy and similarity of participants' perceptions of teamwork. Mixed repeated measures ANOVAs revealed significant main effects for training on both the accuracy, $F(1, 69) = 30.91$, $p < .001$, $\eta^2 = .14$, and the similarity, $F(1, 68) = 5.95$, $p < .05$, $\eta^2 = .08$, of participants' perceptions of teamwork. Consistent with Hypothesis 11, both the accuracy and similarity of participants' perceptions of teamwork were higher after training than before training. Hypothesis 11 was supported.

Hypothesis 12 posited that prior to training (Time 1), participants who have higher levels of interpositional experience would have greater (a) accuracy and (b) similarity of perceptions of teamwork compared to participants with lower levels of interpositional experience. Between-group simple effect analyses revealed that prior to training, there was no significant difference between participants with lower and participants with higher levels of interpositional experience regarding the accuracy, F(1,

69) = 0.17, $p > .05$, $\eta^2 = .00$, or similarity, $F(1, 68) = 0.00$, $p > .05$, $\eta^2 = .00$, of participants' perceptions of teamwork. Hypothesis 12 was not supported.

Table 7
Accuracy and Similarity of Perceptions of Teamwork (i.e., D) by Interpositional Experience Across Time 1 and 2

		Time 1		Time 2		*d*
	N	Mean	*SD*	Mean	*SD*	
Accuracy of Teamwork						
Overall	71	11.17	2.38	10.24	2.91	$.35^{***}$
Low IPE	33	11.05	2.31	9.75	2.38	.55
High IPE	38	11.29	2.49	10.72	3.30	.19
Similarity of Teamwork						
Overall	70	8.59	1.97	8.05	2.32	$.25^{*}$
Low IPE	32	8.58	1.90	7.77	2.08	.41
High IPE	38	8.60	2.08	8.32	2.53	.12

Note. Decrease in means indicate increased accuracy or similarity. Positive *d* indicates increased accuracy or similarity (consistent with hypothesis). IPE = interpositional experience. No follow-up simple effect analyses were performed because no significant interactions were found.
$^{*}p < .05$ (two-tailed), $^{***}p < .001$ (two-tailed).

Hypothesis 13 posited that increases in the levels of (a) accuracy and (b) similarity of participants' perceptions of teamwork from Time 1 to Time 2 will be greater for participants who have lower levels of interpositional experience compared to participants with higher levels of interpositional experience. Mixed repeated measures ANOVAs indicated that interpositional experience did not moderate the effect of CRM on the accuracy, $F(1, 69) = 1.74$, $p > .05$, $\eta^2 = .03$, or the similarity, $F(1, 68) = 1.37$, $p > .05$, $\eta^2 = .02$, of crew members' perceptions of teamwork. Hypothesis 13 was not supported.

Hypotheses Pertaining to 2-year Follow-up (Hypotheses 14 - 26)

Preliminary analysis due to low response rate. As described in the procedures section, a follow-up package consisting of the team-relatedness, team workflow pattern, and RMAQ measures, and the knowledge test was sent to participants two years after their participation in CRM training. Unfortunately, only 15 of 160 surveys were returned, resulting in a 9% return rate. Correspondence with participants and management at the host railroad suggested that specific issues between railroad labor and management were the most likely causes for the low level of responsiveness. Discussions with railroad personnel suggested that continued effort to collect follow-up data (i.e., a second round of mailings) would not yield additional responses. Although the small sample size and low power precluded the running of quantitative statistical analysis on the follow-up measures, for the sake of completeness, the means, standard deviations, and effect sizes are presented in Appendix F.

DISCUSSION AND CONCLUSIONS

Research has reported positive relationships between the accuracy and similarity of team members' knowledge or perceptions regarding the interdependence of their task and team performance (Arthur et al., 2005; Cooke et al., 2003; Edwards et al., 2006; Marks et al., 2000; Marks et al., 2002; Mathieu et al., 2000; Mathieu et al., 2005; Smith-Jentsch et al., 2005). Consequently, there have been calls for investigations of how to improve the accuracy and similarity of team members' perceptions or knowledge of the task (Arthur et al., 2005; Brun et al., 2005; Cooke et al., 2003; Edwards et al., 2006; Webber et al., 2000). However, there are only a handful of studies assessing the impact of training on criteria pertaining to team members' perceptions of task interdependence (i.e., team-interaction models; see Brun et al., 2005; Cooke et al., 2003; Marks et al. 2002; Marks et al., 2000), and none on the impact of training interventions on perceptions of task interdependence directly.

This study begins to fill in that gap by examining CRM training's impact on the accuracy and similarity of railroad crew members' perceptions regarding their team tasks. The overall results suggest that CRM training is effective in increasing the accuracy and similarity of crew members' perceptions of team-relatedness (amount of task interdependence) and perceptions regarding the importance of teamwork. However, the effectiveness is often dependent on the metric used (i.e., accuracy vs. similarity), and the specific characteristics of the crew members (i.e., locomotive vs. MOW, higher vs. lower levels of interpositional experience). Furthermore, the results suggest that training did not increase the accuracy or similarity of crew members' perceptions of team workflow pattern (form of task interdependence). Lastly, a low return rate (9%) for the 2-year follow-up survey resulted in a very small sample size (*n* = 15) and low power which precluded any quantitative statistical analyses to test the specified hypotheses. However, summary descriptive statistics for the limited sample of participants are presented in Appendix F.

Crew Members' Pre-Training Perceptions

Theorists suggest and research has shown that team members can misperceive the structural interdependence of their task and consequently have different perceptions than their teammates (Arthur et al., 2005; Comeau & Griffith, 2005; Tjosvold, 1986; Wageman, 1995). Likewise, I posited that various factors specific to the railroad environment, including strict rules and regulations regarding communication as well as disparate union representation within a crew, could lead crew members to perceive their jobs as less team-related and involving a less intensive workflow pattern than is actually the case. However, results of the current study suggest that the accuracy and similarity of participants' pre-training perceptions of the task and team were relatively accurate (see Tables 3, 4, 6, and 7). This suggests that strict rules and regulations, divergent within-crew union membership, and the lack of previous team training, does not substantially affect the accuracy of crew members' perceptions of task interdependence and teamwork. These findings are consistent with theories of cooperation (Deutsch, 1949) and team mental models (Cannon-Bowers et al., 1993) that suggest the structural interdependence of the task has a direct influence on team members' perceptions of task interdependence. Previous research also supports the positive relationship between actual and perceived task interdependence (Arthur et al., 2005; Wageman, 1995).

Additionally, research has outlined qualitative differences between locomotive and MOW crews (Morgan et al., 2003) and many of the a priori hypotheses regarding the moderation of crew type and interpositional experience on crew member perceptions (Hypotheses 3, 6, 10, and 13) were based on posited pre-training difference between these groups' perceptions (Hypotheses 2, 5, 9, and 12). Specifically, MOW crews are larger, have fewer rules and regulations regarding their communication processes, and have less distinction among crew members' specific tasks, and it was posited that as a result MOW crews would have more accurate and similar perceptions of task interdependence and teamwork before training then locomotive crews. Likewise, it was suggested that higher levels of interpositional experience resulting from taking on the role of another crew position would increase the accuracy and similarity of that crew member's pre-training perceptions. However, the present study found that prior to training, there were no significant differences between locomotive and MOW crew members or those with higher or lower levels of interpositional experience, in reference to their perceptions of team-relatedness, team workflow pattern, or teamwork. This suggests that these group differences do not substantially affect crew members' perceptions of their task or team.

There are two important caveats. First, although testing the hypotheses described above using repeated measures analyses (with the similarity metric *D)* failed to find significant differences between locomotive and MOW crew members in terms of the accuracy or similarity of any pre-training perceptions, using the alternative similarity metric $|D|$ (Dougherty & Pritchard, 1985), or running a one-way analysis on *D*, significant differences were found regarding the accuracy of perceptions of teamwork (see Endnote 6). Results of these analyses suggest that, as hypothesized, locomotive crew members had significantly less accurate pre-training perceptions of teamwork compared to MOW crew members. That is, prior to training, MOW crew members more accurately understood that behaviors such as communicating, adapting, and being assertive are critical for crew safety and performance. Second, although not significant (and in the opposite direction as predicted), results in Table 3 show that locomotive crew members had more accurate perceptions of team-relatedness than MOW crew members (d =.30).

This brings up an important question. Why is it that prior to training, locomotive crew members have more accurate perceptions regarding the extent to which they can successfully perform their job alone without other members of their crew (i.e., team-relatedness), yet MOW crew members have more accurate perceptions of the *teamwork behaviors* needed when interacting with other crew members? One plausible explanation for why locomotive crew members have more accurate perceptions of team-relatedness is because there is greater distinction between locomotive crew members' specific tasks and they have more rules and regulations regarding those distinctions than MOW crew members. If these greater distinctions between locomotive crew members' specific tasks do not lead them to perceive their position specific job duties as ends in themselves (as I posited), then these distinctions would most likely make locomotive crew members perceive that they can not complete the "overall" locomotive crew task alone under any circumstances (i.e., accurate perception of team-relatedness). However, the specific crew member behaviors during within-crew interactions (i.e., coordination, communication) are more *highly structured* within locomotive crews' tasks compared to MOW crews' tasks. For example, the communication of a track warrant between a dispatcher, conductor, and engineer, necessitates particular words, phrasing, and pronunciation.[8] These structured interaction patterns in the locomotive cab could result in locomotive crew members not perceiving (i.e., less accurate) the criticality of the less structured teamwork behaviors being measured by the RMAQ (e.g., helping behavior, informal communication). Because there are no analogous rules and regulations regarding specific interaction patterns during MOW tasks (Morgan et al., 2003), MOW crew members should have a better (i.e., more accurate) understanding of the importance of the *less structured* teamwork behaviors being measured by the RMAQ. This would account for the present study's results showing that locomotive crew members have more accurate perceptions of team-relatedness (not significant), while MOW crew members have more accurate perceptions of teamwork (alternative analyses). However, future research with a larger sample and more power is needed to investigate the extent to which this is the case.

An alternative explanation for why locomotive crew members' have more accurate pre-training perceptions of task interdependence than hypothesized, lies in the specific regulatory climate present during training. Specifically, during training, locomotive crew members communicated their anxiety about their suspicions that railroads are moving toward one-man locomotive operations. The host railroad's single-person crew pilot programs (United Transportation Union, 2006), studies of positive train control (FRA, 2007a), and the use of remote-control yard operations (FRA, 2007b) suggest that locomotive crew members' fears are not entirely unfounded. In fact, this has been a contentious issue in labor-management relations at many railroads. It is possible that these fears could have unduly influenced locomotive crew members' responses to the team-relatedness measure (i.e., "the extent to which they can not successfully perform the job alone without other members of their team"). Knowing the study results would be seen by the FRA, participants might have been tempted to inflate their pre-training responses to the measure suggesting that "they can not perform the job alone under any circumstances," when without such a fear, a less team-related response might have been given. Although there is no way to determine if this influenced their responses, if true, it would most likely be limited to the team-relatedness scale.

Training's Effect on Perceptions of Team-relatedness

Overall, this study found that training can increase the accuracy of crew members' perceptions of team-relatedness. As seen in Tables 3 and 4, effect sizes (*d*) ranged from .08 to .31 (i.e., accuracy increased for all groups) with larger effect sizes for locomotive crew members and those with lower levels of interpositional experience (no significant interactions). Furthermore, effect sizes for main effects ranged from .19 to .22 – effects that would generally be considered small (Cohen, 1988). Specifically, training led participants to perceive their tasks as being more difficult to perform alone and more similar to the expert referent "it can not be performed alone under any circumstances." Recent research has also found that other types of team training can indeed influence the accuracy of "team-interaction" mental models (Cooke et al., 2003; Marks et al., 2002; Marks et al., 2000). This is consistent with theories of team mental models suggesting that team training, which emphasizes the requirements of coordination and communication, should help team members recognize and identify the structural interdependence of their tasks (Rouse et al., 1992; Salas et al., 1992; Tannenbaum et al., 1992). These theories posit that as a result of this increased accuracy, crew members' approach to their tasks will be more compatible with the task demands and requirements needed for successful performance. Although there were no performance measures in the current study, research has shown that increases in the accuracy of team members' team interaction mental models are associated with improved performance (Edwards et al., 2006).

It is important to note that although, overall, CRM increased the *accuracy* of crew members' perceptions of team-relatedness for all participants, the present study suggests that training *only* increased the within-crew *similarity* of perceptions of team-relatedness for locomotive crew members and participants with lower levels of interpositional experience.[9] In order to further understand the nature of this finding within locomotive crews, follow-up analysis by position within the crew (engineer vs. conductor) was also performed. Although based on a small sample, position level data revealed that at both Time 1 and 2, every conductor (n = 4) rated his overall task as a 5 (i.e., it can not be performed alone under any circumstances) on the team-relatedness scale, reflecting complete accuracy at both times and no change as a result of training. It appears that prior to training, conductors understand they can not complete their jobs without engineers and training most likely reinforces that accurate perception. It is important to note that although conductors' duties are critical before, during, and after a train movement, in terms of "safely" moving a train, he or she does not have the proper authority (or perhaps the skill and knowledge) to put the train in motion. In this instance it appears the structural interdependence of the task is having a direct effect on conductors' perceptions of team-relatedness.

In contrast, the mean of engineers' ($n = 3$) pre-training ratings regarding the team-relatedness of their overall job was 3.66 ($SD = 2.31$) and 4.33 ($SD = 1.16$) immediate post-training, suggesting that training led engineers to perceive their job as being more difficult to complete alone (i.e., more accurate). Anecdotal evidence suggests training resulted in gamma change in that it led engineers to redefine or reconceptualize their jobs (Golembiewsi, Billingsley, & Yeager, 1976). Specifically, because an engineer can move the train without the conductor (albeit against current safety regulations) his pre-training perceptions of team-relatedness are low (i.e., inaccurate). However, conversations with engineers after training suggested that CRM changed their definition of their overall task from simply "moving the train" to "moving the train safely" and that increases in the accuracy of engineers' perceptions of team-relatedness are the result of recognizing that they would have a difficult time accomplishing the latter without a conductor. However, additional research investigating the extent to which CRM training results in crew members "redefining" their jobs is warranted. In summary, increases in the accuracy of engineers' perceptions from Time 1 to Time 2, coupled with the range-restricted conductors' perceptions, explains the current findings regarding increased accuracy and similarity of locomotive crew members' perceptions of team-relatedness.

Similarly, CRM training significantly increased the similarity of perceptions of team-relatedness for participants with lower levels but not for those with higher levels of interpositional experience (see Table 4 and Figure 1). A reasonable explanation as to why participants with higher levels of interpositional experience are less responsive to training is simply because as a result of that experience, they should have less difficulty completing the job alone than those with lower levels of interpositional experience. That is, crew members with higher levels of interpositional experience would rely on their experience to complete their fellow crew members' duties.

Additionally, the findings regarding the differential effects of CRM on crew members with higher vs. lower levels of interpositional experience could be the result of the unequal distribution of interpositional experience across crew type. Specifically, 6 of the 7 locomotive crew members who had data on interpositional experience, had not held other jobs within their crew (i.e., lower levels of interpositional experience). As a result, the previously discussed large increases in the similarity of team-relatedness for locomotive crew members is disproportionately present in the lower level of interpositional experience condition. Because of the unequal group sizes and inability to run analyses using all variables simultaneously, it is difficult to determine the degree to which crew type is confounding the results regarding interpositional experience. Future research should be conducted with greater control over training make up such that equal cell sizes can be achieved or levels of specific independent variables can be proportionately distributed across levels of other independent variables. Only under these circumstances would it be possible to test for the effects of an independent variable on a dependent variable while statistically controlling for variations in other independent variables.

Training's Effect on Perceptions of Team Workflow Pattern

Interestingly, the results suggest that CRM training had no positive effect on the accuracy or similarity of participants' perceptions of team workflow pattern. In fact CRM training resulted in a significant decrease in the accuracy of perceptions of team workflow pattern for crew members who had lower levels of interpositional experience (see Figure 2). This is most likely the result of the training itself. Theory and research suggests that the effect of training on team member knowledge and perceptions depends on the type of training, the type of team, and the specific knowledge being trained (Arthur et al., 2003; Marks et al., 2002; Rouse et al., 1992). Mismatches between the type of training and knowledge content have been put forth as reasons for disappointing results in the past (see Brun et al., 2005). In the present study, although researchers have broadly suggested that CRM might be useful for creating and reinforcing team mental models (Rouse et al., 1992), a closer examination of CRM suggests that it might not be ideal for influencing perceptions of team workflow pattern. That is, CRM content does not detail specific workflow patterns to be used by crews, which would seem necessary to increase the accuracy and similarity of a crew members' perceptions regarding those patterns. As stated previously, CRM content is aimed more at teaching the importance of working together as a team, using two-way communication, and outlining specific instructions on how to communicate one-on-one with another crew member. This content seems more relevant to perceptions of team-relatedness and is most likely the reason for the positive effects discussed in the previous section. The disparate results regarding CRM's effect on team-relatedness and team workflow pattern, as well as the low correlations between indices of team-relatedness and team workflow pattern shown in Table 1, is consistent with theory and research suggesting that they are related, yet distinct aspects of task interdependence (Arthur et al., 2005; Wageman et al., 2005). Fortunately, increases in the accuracy of crew members' perceptions regarding their task's team-relatedness should lead to improved coordination and performance within any of the workflow patterns which involve interaction (i.e., sequential, reciprocal, or intensive).

Alternatively, it could be the case that the reason CRM did not positively influence crew members' perceptions of team workflow pattern is the team workflow pattern measure itself might not have been appropriate in the current study. For example, the previously noted misunderstanding of instruction by some participants suggests crew members might have found it difficult to rank-order workflow patterns in terms of the extent to which they are descriptive of the activities in their overall job. Similar to other jobs, railroad jobs encompasses many, sometimes divergent tasks or activities (Morgan et al., 2003; U.S. Department of Labor, 2006), each of which may involve different workflow patterns. For example, a conductor might work autonomously when reviewing a wheel report, yet work in an intensive workflow pattern with the engineer when addressing a critical situation involving train movement (e.g., braking problem). Although such critical decision-making tasks require intensive workflow patterns (and were represented in several case-studies during training), they occur less frequently than lower-order workflow patterns (i.e., additive – reciprocal) during the completion of normal work tasks. An alternative measurement method which divides a crew members' overall job into more distinct subtasks might give insight into the range of, and frequency in which, specific workflow patterns occur during a job. For example, in a laboratory study using 3-person crews, Arthur et al. (2005) asked participants to rate the workflow pattern of each of their 34 individual tasks. The authors then created a composite of the task-level ratings to form a index of the task workflow pattern of the job as a whole. Because of the wide array of tasks that many real-world crews (including railroad crews) accomplish, such a composite measure might yield a more valid indicator of non-laboratory crew members' perceptions of their tasks and jobs. Future field research should be conducted comparing an overall job index and a composite index in terms of the extent to which they accurately capture crew members' perceptions of task workflow pattern.

Lastly, one is inclined to wonder if the results were influenced by the fact that some participants' responses to the team workflow pattern scale were rescored due to their misunderstanding the instructions. As detailed earlier and shown in Appendix E, the normative ratings of participants who completed the measure incorrectly were modified to that of an ipsative ranking scale. Follow-up analyses revealed that, compared to participants who completed the scale incorrectly, those who completed it correctly tended to have more accurate pre-training and immediate post-training perceptions of team workflow pattern. In three instances the statistical support of hypotheses changed as a result of using only participants who filled out the team workflow measure correctly. Two of these instances (see Endnote 4 and 5) were the result of statistical abnormalities resulting from an extremely small sample size ($n = 2$) combined with unweighted means analysis. In the third instance (see Endnote 6), although significant results were obtained using all participants (means were in the opposite direction as predicted) using only those participants who filled out the team workflow pattern measure correctly, interpositional experience did not moderated the effect of CRM on the accuracy of crew members' perceptions of team workflow pattern. Thus, these alternative analyses suggest that overall, results of the current study were not influenced by the process of correcting for participants' misunderstanding of the instructions to the team workflow pattern measure.

Training's Effect on Perceptions of Teamwork (RMAQ)

The results of the current study indicate that CRM training also increased the accuracy and similarity of crew members' perceptions of teamwork. Specifically, training led participants to more accurately and similarly perceive their tasks as requiring greater levels of interpersonal awareness, communication and coordination between crew members, and attention to crew stress and fatigue. This is consistent with theories of team training suggesting that training that focuses on the nature of team tasks and provides examples of teamwork failures should help team members to recognize the importance of specific teamwork skills and behaviors (Rouse et al., 1992; Salas et al., 1992; Tannenbaum et al., 1992). Models of team effectiveness posit that as a result of increased accuracy regarding crew members' perceptions of specific teamwork behaviors, crew members will be more likely to perform those behaviors, thus enhancing team performance (Rentsch & Hall, 1994; Rouse et al., 1992; Salas et al., 2004). Similarly, these models posit that as team members' knowledge of relevant teamwork behaviors become more similar, they can better predict the behavior and interactions with fellow team members, resulting in improved team processes and performance (Cannon-Bowers & Salas, 2001; Cannon-Bowers et al., 1993).

These results of the current study are consistent with recent research that has found that team training can indeed influence the accuracy and similarity of teamwork mental models (Smith-Jentsch, Campbell, Milanovich, & Reynolds, 2001). Furthermore, these findings are consistent with research in the aviation (Salas et al., 2001), off shore drilling (Flin & O'Connor, 2001), and healthcare (Howard et al., 1992) industries that use the CMAQ to assess CRM training effectiveness. However, unlike previous CRM research, this paper employed the RMAQ in a mental model framework to assess the accuracy of crew members' responses. This is important because, unlike the airline industry (Helmreich et al., 1986), there is no research associating higher scores on the RMAQ with improved team behaviors or performance. Having experts complete the RMAQ and comparing participants' perceptions of teamwork (RMAQ) to the computed expert referent (i.e., accuracy) allowed for the qualification of participants' RMAQ scores without assuming higher scores are associated with higher levels of performance in the railroad environment.

Limitations and Directions for Future Research

In addition to those already noted, several limitations and directions for future research are examined below. One limitation of the present study was the small number of locomotive crew members who completed the task interdependence measures (see Table 3). This was the direct result of the author's limited control over training and data collection in the field setting. The current study employed repeated measures designs which resulted in proportional cell sizes and allowed for analyses using unweighted means (e.g., equal weight was given to both locomotive and MOW crew means). Although this is appropriate in the circumstances present in the current study (Bonnett, 1982; Green et al., 1976), caution must be taken when interpreting these results. Future research should attempt to replicate these results using a larger sample of locomotive crew members.

Another limitation of the present study was the lack of a measure of behavior or performance. Over a 20 year history in aviation, CRM training has not only been shown to increase crew member knowledge, skills, and attitudes in the cockpit, but also to improve teamwork behavior and performance (Salas et al., 2001; Salas, Wilson, Burke, & Wightman, 2006). In the aviation industry, team behavior and performance can be assessed during line oriented flight training (LOFT) and jump seat observation or line operation safety audits (LOSA; Ikomi, Boehm-Davis, Holt, & Incalcaterra, 1999). Unfortunately, simulators and programs such as LOSA are in their infancy in the railroad industry (Morgan et al., 2003; Roop et al., 2007). Hopefully, as a result of the increased attention to CRM in the railroad industry, future research will utilize more complex training evaluation designs aimed at assessing the impact of CRM on behavioral and performance measures. Unlike the present study, this would allow researchers to assess the relationship between the accuracy and similarity of crew members' perceptions of task interdependence and crew behavior or performance.

Several general limitations to the study design, which can impact the study's internal validity are also important to note. First, in the absence of a control group, it is unknown whether it was CRM training in particular (and not just any training) that led to changes in outcome measures. Furthermore, as a result of the ex post facto nature of the quasi-experimental design (i.e., crew type, interpositional experience are not manipulated variables), there is no way to be certain that any of the differences found were specifically *caused* by being in a particular group. Similarly, there is no way to be certain that any posited differences between the conditions (e.g., more rules and regulations for locomotive crews) were direct causes of differences in outcome measures (Myers & Hansen, 1997). Future research with greater experimental control can manipulate these posited mediating variables to determine the exact nature of the relationship.

Another potential area of future research is investigating what factors or conditions influence how crew members perceive the interdependence of their tasks. Aside from the influence of the task's structural interdependence (Arthur et al., 2005; Wageman, 1995), there is little research assessing the effect of other variables on perceived interdependence. I suggest, as others have (Tjosvold, 1984), that more research should be conducted to determine what factors influence the accuracy and similarity of team members' perceptions regarding the interdependence of their tasks. For example, the physical proximity of team members has been shown to be related to the perceptions of team unity (Urban, Bowers, Cannon-Bowers, & Salas, 1995) and research on its effect on perceptions of task interdependence could contribute to the literature on virtual teams (Pinsonneault & Caya, 2005). Other variables of interest include the organizational climate in which the team exists, reward and promotional practices, the size and complexity of work crews, as well as team members' positions within the team.

Lastly, assessing task workflow pattern at the level of the task (compared to the job) might give a more reliable and clearer picture regarding the range of workflow patterns that occur during the accomplishment of crew members' tasks. However, whether using holistic or decomposed ratings, there is relatively little research regarding the extent to which the current task workflow pattern measure accurately captures the true workflow patterns that exist in "real" world teams. Future research should be conducted to determine the degree to which crew members' perceive the five workflow pattern used in the current measure as accurately and comprehensively capturing the range of patterns that exist in their jobs.

Conclusion

The overall results of the present study suggests that CRM training is effective in increasing the accuracy and similarity of crew members' perceptions of team-relatedness (amount of task interdependence) and perceptions of teamwork. However, the effectiveness is often dependent on the metric used (i.e., accuracy vs. similarity), and the specific characteristics of the crew members (i.e., locomotive vs. MOW, higher vs. lower interpositional experience). Furthermore, the results suggest that training did not increase the accuracy or similarity of crew members' perceptions of team workflow pattern (form of task interdependence). These findings, along with the low positive correlations between the two indices, are consistent with research in team task analysis that suggests that team-relatedness and team workflow pattern represent related yet distinct facets of task interdependence (Arthur et al., 2005). Coupled with data indicating participants found the training to be job relevant and practical, and stated it would influence their ability to later perform their job, these findings suggest that CRM training can likely influence crew member behavior and ultimately safety.

ENDNOTES

[1]Alternative similarity indices include D^2 (Cronbach & Gleser, 1953), $|D|$ (Dougherty & Pritchard, 1985), or for rank ordered data (i.e., team workflow pattern scale), the rank order correlation r_s (Spearman, 1904). However, as a result of squaring differences (without later taking the square root of the sum of squared differences), both D^2 and r_s (a linear combination of D^2) give exponentially greater weight to differences of larger magnitude. Theory and research on team knowledge and team mental models does not suggest that an exponential function more accurately represents the similarity of the perceptions (e.g., team-relatedness, teamwork) assessed in the current study.

[2]A one-way repeated measures ANOVA which uses weighted means (unlike the two-way [Training × Crew Type] repeated measures ANOVA) indicated that CRM training did increase the accuracy of participants' perceptions of team-relatedness, $F(1, 62) = 4.297$, $p < .05$, $\eta^2 = .07$.

[3]A one-way repeated measures ANOVA which uses weighted means (unlike the two-way [Training × Crew Type] repeated measures ANOVA) indicated that CRM training did not affect the similarity of participants' perceptions of team-relatedness, $F(1, 61) = 1.777$, $p > .05$, $\eta^2 = .03$.

[4]Unlike the analysis using all participants' data, a mixed repeated measures analysis of variance using only those participants who filled out the team workflow pattern measure correctly obtained a significant effect of CRM on the similarity of crew members' perceptions of team workflow pattern, $F(1, 31) = 22.77$, $p < .001$, $\eta^2 = .42$. However, the direction of the effect was the opposite of what had been predicted (M_{pre} =

2.51, $SE = .43$; $M_{post} = 4.56$, $SE = .42$. Because this effect is the direct result of sampling error caused by the extremely small sample size in the locomotive crew condition ($n = 2$) combined with unweighted means analysis, it precludes further analysis of this finding.

[5]Unlike the analysis using all participants data, a mixed repeated measures analysis of variance using only those participants who filled out the team workflow pattern measure correctly indicated that the effect of CRM on the similarity of crew members' perceptions of team workflow pattern was moderated by crew type, $F(1, 31) = 14.087$, $p < .001$, $\eta^2 = .31$. Means were in the opposite direction as predicted however ($M_{pre/mow} = 2.43$, $SD = 1.20$; $M_{pre/loc} = 2.58$, $SD = .82$; $M_{post/mow} = 2.87$, $SD = 1.16$; $M_{post/loc} = 6.24$, $SD = .11$). Because this effect is the direct result of sampling error caused by the extremely small sample size in the locomotive crew condition ($n = 2$) combined with unweighted means analysis, it precludes further analysis of this finding.

[6]When using only those participants who filled out the team workflow measure absolutely correctly, there were no significant effects.

[7]Alternatively, when running a one-way analysis at Time 1, a significant difference was found regarding the accuracy of perceptions of teamwork between MOW and locomotive crews prior to training (Time 1), $F(1, 143) = 4.15$, $p < .05$, $\eta^2 = .03$. Means were in the predicted direction ($M_{mow} = 10.81$, $SD = 2.48$ $M_{loc} = 11.71$, $SD = 2.79$). Furthermore, means were in the predicted direction and significant results were found when running a between-group simple effect analysis using the similarity metric $|D|$ (Dougherty & Pritchard, 1985).

[8]A track warrant is permission for a train to occupy a specific section of track and is used to control trains operating on track without a wayside signaling system (i.e., dark territory). Communication of a track warrant involves "written" verbal orders which may be modified or rescinded by communication over a radio with a dispatcher. During communication of a track warrant, a dispatcher gives a train or a maintenance crew verbal authority (a warrant) to occupy a portion of main line track between particular named locations (e.g., mile markers, switches, stations, or other points). In addition, track warrants can specify speed limits, direction, time limits, and whether to clear the main line (e.g., by entering a secondary track such as a siding) or any other section of track (sidings, yards secondary track). There is a complicated and time consuming procedure by which track warrants are issued which involves the train conductor or engineer reading back the warrant to the dispatcher before the warrant goes into effect.

[9]The finding that there is an increase in accuracy but a decrease in similarity seems to contradict current thinking on team mental models. For example, team mental model research suggests that as crew members' mental models increase in accuracy, similarity should also increase (see Edwards et al., 2006). However, this assumes that increases in accuracy are the result of a convergence around a "true" score which is midway between all team members' scores. It is true that in these circumstances increases in accuracy would lead to parallel increase in similarity. However, when a "true" score is at the far end of a scale or metric, a team's mean score on that scale can move closer to that referent (i.e., increase in accuracy), while the variance in team member's scores can remain constant or even increase (i.e., decrease in similarity).

REFERENCES

Anderson, J., & Sandza, R. (1987, August 3). A sting in the Gulf. *Newsweek, 110*, 24-29.

Anderson, S. E., & Williams, L. J. (1996). Interpersonal, job, and individual factors related to helping processes at work. *Journal of Applied Psychology, 81*, 282-296.

Arthur, W. Jr., Bennett, W. Jr., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology, 88*, 234-245.

Arthur, W., Jr., Bennett, W., Jr., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance, 11*, 57-123.

Arthur, W. Jr., Edwards, B. D., Bell, S. T., Villado, A. J., & Bennett, W. J. (2005). Team task analysis: Identifying tasks and jobs that are team based. *Human Factors, 47*, 654-669.

Arthur, W. Jr., Glaze, R. M., Bhupatkar, A., & Villado, A. J. (2008). *Team task analysis: Differentiating between jobs using team-relatedness and team workflow as metrics of team task interdependency.* Poster presented at the 23rd annual meeting of the Society for Industrial and Organizational Psychology, San Francisco, CA.

Arthur, W. Jr., Villado, A. J., & Bennett, W. Jr. (in press). Innovations in team task analysis: Identifying task elements, tasks, and jobs that are team-based. In W. Bennett, Jr. (Ed.), *The future of job analysis*. Mahwah, NJ: LEA.

Bishop, J. W., & Scott, D. K. (2000). An examination of organizational and team commitment in a self-directed team environment. *Journal of Applied Psychology, 85*, 439-450.

Bonett, D. G. (1982). A weighted harmonic means analysis for the proportional unbalanced design. *Educational and Psychological Measurement, 42*, 401-407.

Bowers, C. A., Morgan, B. B., Salas, E., & Prince, C. (1993). Assessment of coordination demand for aircrew coordination training. *Military Psychology, 5,* 95-112.

Brannon, N. G., & Koubek, R. J. (2001). Towards a conceptual model of procedural knowledge degradation. *Theoretical Issues in Ergonomic Science, 2*, 317-335.

Brass, D. J. (1985). Technology and the structuring of jobs: Employee satisfaction, performance and influence. *Organizational Behavior and Human Decision Processes, 35*, 216-240.

Brun, W., Eid, J., Johnsen, B. H., Laberg, J. C., Ekornas, S. B., & Kobbeltvedt, T. (2005). Bridge resource management training: Enhancing shared mental models and task performance? In H. Montgomery, R. Lipshitz & B. Brehmer (Eds.) *How professionals make decisions (*pp. 183-193). Hillsdale, NJ: Lawrence Erlbaum Associates.

Burke, C. S., Wilson, K. A., & Salas, E. (2003). Teamwork at 35,000 feet: Enhancing safety through team training. *Human Factors and Aerospace Safety, 3*, 287-312.

Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel Psychology, 46*, 823-850.

Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior, 22*, 195-202.

Cannon-Bowers, J. A., Salas, E., Blickensderfer, E., & Bowers, C. A. (1998). The impact of cross-training and workload on team function: A replication and extension of initial findings. *Human Factors, 40*, 92-101.

Cannon-Bowers, J. A., Salas, E., & Converse, S. A. (1990). Cognitive psychology and team training: Training shared mental models and complex systems. *Human Factors Society Bulletin, 33*, 1-4.

Cannon-Bowers, J. A., Salas, E., & Converse, S. A. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and group decision making* (pp. 221-246). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining competencies and establishing team training requirements. In R. A. Guzzo, E. Salas, and Assoc. (Eds.), *Team effectiveness and decision making in organizations* (pp. 333-380). San Francisco, CA: Jossey-Bass.

Childs, J. M., & Spears, W. D. (1986). Flight-skill decay and recurrent training. *Perceptual and Motor Skills, 62*, 325-242.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Comeau, D. J., & Griffith, R. L. (2005). Structural interdependence, personality, and organizational citizenship behavior: An examination of person-environment interaction. *Personnel Review, 34*, 310-330.

Cooke, N. J., Kiekel, P. A., & Helm, E. E. (2001). Measuring team knowledge during skill acquisition of a complex task. *International Journal of Cognitive Ergonomics, 5*, 297-315.

Cooke, N. J., Kiekel, P. A., Salas, E., Stout, R. E., Bowers, C., & Cannon-Bowers, J. (2003). Measuring team knowledge: A window to the cognitive underpinnings of team performance. *Group Dynamics: Theory, Research, and Practice, 7,* 179-199.

Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. E. (2000). Measuring team knowledge. *Human Factors, 42*, 151-173.

Cooper, J. B., Newbower, R. S., & Kitz, R. J. (1984). An analysis of major errors and equipment failures in anesthesia management: Considerations for prevention and detection. *Anesthesiology, 60*, 34-42.

Cronbach, L. J., & Gleeser, G. C. (1953). Assessing similarity between profiles. *The Psychological Bulletin, 50*, 456-473.

Cullen, D. (1990). *The public inquiry into the Piper Alpha Disaster* (CM 1310). London: HMSO.

Day, E. A., Fein, E., & Arthur, W., Jr. (2000). *Relating member ability and personality to dyadic crew performance.* Unpublished manuscript, Texas A&M University, College Station.

Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations, 2*, 129-152.

Deutsch, M. (1962). Cooperation and trust: Some theoretical notes. In M. R. Jones (Ed.), *Nebraska symposium on motivation* (pp. 275-319). Lincoln, NE: University of Nebraska Press.

Devine, D. J., Clayton, L. D., Philips, J. L., Dunford, B. B., & Melner, S. B. (1999). Teams in organizations: Prevalence, characteristics and effectiveness. *Small Group Research, 30*, 678-711.

Dieterly, D. L. (1988). Team performance requirements. In S. Gael (Ed.), *The job analysis handbook for business, industry, and government* (pp. 766-777). New York: Wiley.

Doerr, K. H., Mitchell, T. R., Schriesheim, C. A., Freed, T., & Zhou, X. (2002). Heterogeneity and variability in the context of flow lines. *Academy of Management Review, 27*, 594-607.

Dougherty, T. W., & Pritchard, R. D. (1985). The measurement of role variables: Exploratory examination of a new approach. *Organizational Behavior and Human Decision Processes, 35*, 141-155.

Edwards, B. D., Day, E. A., Arthur, W. J., & Bell, S. T. (2006). Relationship among team ability composition, team mental models, and team performance. *Journal of Applied Psychology, 91*, 727-736.

Ellis, A. P. J., Bell, B. S., Ployhart, R. E., Hollenbeck, J. R., & Ilgen, D. R. (2005). An evaluation of generic teamwork skills training with action teams: Effects on cognitive and skill-based outcomes. *Personnel Psychology, 58*, 641-672.

Federal Aviation Administration (FAA). (2004). *Advisory Circular: Crew Resource Management Training* (120-51E). Washington, DC: U.S. Government Printing Office.

Federal Railroad Administration (FRA). (1999). *Railroad Safety Statistics Annual Report*. Washington, DC: United States Department of Transportation, Office of Public Affairs.

Federal Railroad Administration (FRA). (2004). *Train Accident Data (by cause)*. Report generated October 15, 2005, from the Office of Safety Analysis: http://safetydata.fra.dot.gov/OfficeofSafety/Query/default.asp?page=inccaus.asp

Federal Railroad Administration (FRA). (2007a). *Human factors considerations in the evaluation of processor-based signal and train control systems* (DOT/FRA/ORD-07/07). Retrieved April, 25th, 2008 from http://www.fra.dot.gov/us/content/288

Federal Railroad Administration (FRA). (2007b). *Development of a preliminary set of human factors functional design guidelines for remote control locomotive systems* (DOT/FRA/ORD-07/19). Retrieved April, 25th, 2008 from http://www.fra.dot.gov/us/content/288

Fleishman, E. A., & Zaccaro, S. J. (1992). Toward a taxonomy of team performance functions. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 31-56). Norwood, NJ: Ablex Publishing.

Flin, R., & O'Connor, P. (2001). Applying crew resource management on offshore oil platforms. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations: Applications of resource management training* (pp. 217-233). Mahwah, NJ: Lawrence Erlbaum.

Foushee, J. C. (1984). Dyads and triads at 35,000 feet: Factors affecting group process and aircrew performance. *American Psychologist, 39*, 885-893.

Galbraith, J. (1977). *Organizational design*. Reading, MA: Addison-Wesley.

Gladstein, D. L. (1984). Groups in context: A model of task group effectiveness. *Administrative Science Quarterly, 29*, 499-517.

Golembiewski, R. T., Billingsley, K., Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *Journal of Applied Behavioral Science, 12,* 133-157.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 83*, 88-96.

Goldstein, I. L. (1991). Training in work organizations. In M.D. Dunnette & L.M. Hough (Eds.) *Handbook of industrial and organizational psychology, Vol. 2* (2nd ed.), p. 507-619. Palo Alto, CA: Consulting Psychologists Press, Inc.

Green, P. A., Heiberger, R. M., & Laster, L. L. (1976). Ambiguous computing of the ANOVA for proportional data. *Proceedings of the American Statistical Association, Statistical Computing Section*, 165-170.

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

Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology, 47*, 307-338.

Guzzo, R. A., Jette, R. D., & Katzell, R. A. (1985). The effects of psychologically based intervention programs on worker productivity: A meta-analysis. *Personnel Psychology, 38*, 275-291.

Haines, V. Y., III, Merrheim, G. G., & Roy, M. (2001). Understanding reactions to safety incentives. *Journal of Safety Research, 32*, 17-30.

Harrington, D., & Kello, J. (1993). *Systematic evaluation of nuclear operator team skills training*. Paper presented at the meeting of the American Nuclear Society, San Francisco, CA.

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

Helmreich, R. L., & Foushee, H. C. (Eds.). (1993). *Why crew resource management? Empirical and theoretical bases of human factors training in aviation.* Academic Press: San Diego, CA.

Helmreich, R. L., Foushee, H. C., Benson, R., & Russini, W. (1986). Cockpit resource management: Exploring the attitude-performance linkage. *Aviation, Space, and Environmental Medicine, 57*, 1198-1200.

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

Helmreich, R. L., & Wilhelm, J. A. (1987). *Human performance in aerospace environments: The search for psychological determinants.* Unpublished manuscript, University of Texas, Austin.

Helmreich, R. L., & Wilhelm, J. A. (1991). Outcomes of crew resource management training. *International Journal of Aviation Psychology, 1*, 287-300.

Helmreich, R. L., Wilhelm, J. A., & Gregorich, S. E. (1988). *Revised version of the cockpit management attitudes questionnaire and CRM seminar evaluation form* (NASA/UT Tech. Man. No. 88-3). Austin: The University of Texas at Austin.

Holyoak, K. J. (1984). Mental models in problem solving. In J. R. Anderson & S. M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower* (pp. 193-218). San Francisco: Freeman.

Howard, S. K., Gaba, D. M., Fish, K. J., Yang, G. S., & Sarnquist, F. H. (1992). Anesthesia crisis resource management training: Teaching anesthesiologists to handle critical incidents. *Aviation, Space, and Environmental Medicine, 63*, 763-770.

Ikomi, P. A., Boehm-Davis, D. A., Holt, R. W., & Incalcaterra, K. A. (1999). Jump seat observations of advanced crew resource management (ACRM) effectiveness. *Proceedings of the Tenth International Symposium on Aviation Psychology* (pp. 292-297). Columbus, OH: The Ohio State University Press.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina, MN: Interaction Book Company.

Kiggundu, M. N. (1981). Task interdependence and the theory of job design. *Academy of*

Management Review, 6, 499-508.

Kiggundu, M. N. (1983). Task interdependence and job design: Test of a theory. *Organizational Behavior and Human Performance, 31*, 145-172.

Kirkpatrick, D. L. (1976). Evaluation of training. In R.L. Craig (Ed.), *Training and development handbook: A guide to human resources development* (pp. 18.1-18.27). New York: McGraw-Hill.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management, 20*, 403-437.

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

Kraiger, K., Salas, E., & Cannon-Bowers, J. A. (1995). Measuring knowledge organization as a method for assessing learning during training. *Human Factors, 37*, 804-816.

Marks, M. A., Sabella, M. J., Burke, C. S., & Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *Journal of Applied Psychology, 87*, 3-13.

Marks, M. A., Zaccaro, S. J., & Mathieu, J. E. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology, 85*, 971-986.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Cannon-Bowers, J. A., & Salas, E. (2005). Scaling the quality of teammates' mental models: Equifinality and normative comparisons. *Journal of Organizational Behavior, 26*, 37-56.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology, 85*, 273-283.

McIntyre, R. M., & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. A. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 9-45). San Francisco: Jossey-Bass.

Mearns, K., Flin, R., Fleming, M., & Gordon, R. (1997). *Human and organizational factors in offshore safety*. Suffolk, U.K.: HSE Books.

Medsker, G. J., & Campion, M. A. (1997). Job and team design. In G. Salvendy (Ed.) *Handbook of human factors and ergonomics* (pp. 450-489). New York: Wiley and Sons.

Morgan, C. A, Olson, L. E., Kyte, T. B., & Roop, S. S. (2003). *Railroad crew resource management (CRM): Survey of teams in the railroad operating environment and identification of available CRM training methods* (Technical Report DFTR-091-00030/474450-00001). College Station, TX: Texas Transportation Institute.

Myers, A., & Hansen, C. H. (1997). *Experimental psychology*. Pacific Grove, CA: Brooks/Cole Publishing.

National Transportation Safety Board (NTSB). (1999a). *Collision of Norfolk Southern Corporation Train 255L5 With Consolidated Rail Corporation Train TV 220 in Butler, Indiana, on March 25, 1998* (NTSB/RAR-99/02). Washington, DC: U.S. Government Printing Office.

National Transportation Safety Board (NTSB). (1999b). *Safety Recommendation (NTSB/R/99/26)*. Washington, DC: U.S. Government Printing Office.

Neuman, G. A., Edwards, J. E., & Raju, N. S. (1989). Organizational development interventions: A meta-analysis of their effects on satisfaction and other attitudes. *Personnel Psychology, 42*, 461-489.

O'Connor, P. & Flin, R. (2003). Crew resource management training for offshore teams. *Safety Science, 41*, 591-609.

Paulus, P. B. (2000). Groups, teams, and creativity: The creative potential of idea-generating groups. *Applied Psychology: An International Review, 49*, 237-262.

Pearce, J. L., & Gregersen, J. B. (1991). Task interdependence and extarole behavior: A test of the mediating effects of felt responsibility. *Journal of Applied Psychology, 76*, 838-844.

Pinsonneault, A., & Caya, O. (2005). "Virtual teams: What we know, what we don't know." *International Journal of e-Collaboration, 1*, 1-16.

Prince, A., Brannick, M. T., Prince, C., & Salas, E. (1992). *Team process measurement and implications for training*. Paper presented at the Human Factors Society, Santa Monica, CA.

Rentsch, J. R., & Hall, R. J. (1994). Members of great teams think alike: a model of team effectiveness and schema similarity among team members. In M. M. Beyerlein & D. A. Johnson (Eds.), *Advances in interdisciplinary studies of work teams, volume 1* (pp. 223-262). Greenwich, CT: JAI.

Rentsch, J. R., & Woehr, D. (2004). Quantifying congruence in cognition: Social relations modeling and team member schema similarity. In E. Salas & S.M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 11-31) Washington, DC: American Psychological Association.

Roop, S. S., Morgan, C. A., Kyte, T. B., Arthur, W. Jr., & Villado, A. J. (2007). *Rail crew resource management (CRM): The business case for CRM training in the railroad industry* (Technical Report 475390-00001). College Station, TX: Texas Transportation Institute.

Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions on Systems, Man, and Cybernetics, 22*, 1296-1308.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin, 100*, 349-363.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103*, 734-760.

Saavedra, R. Early, P., & Van Dyne, L. (1993). Complex interdependence in task-performing groups. *Journal of Applied Psychology, 78*, 61-72

Salas, E., Bowers, C. A., & Edens, E. (Eds.). (2001). *Improving teamwork in organizations: Applications of resource management training*. Mahwah, NJ: Lawrence Erlbaum Associates.

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors, 43*, 641-674.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (Eds.). (1992). *Toward an understanding of team performance and training.* Westport, CT: Ablex.

Salas, E., Fowlkes, J. E., Stout, R. J., Milanovich, D. M., & Prince, C. (1999). Does CRM training improve teamwork skills in the cockpit? Two evaluation studies. *Human Factors, 41*, 326-343.

Salas, E., Kosarzycki, M. P., Tannenbaum, S. I., & Carnegie, D. (2004). Principles and advice for understanding and promoting effective teamwork in organizations. In R. J. Burke & C. Cooper (Eds). *Leading in turbulent times* (pp. 95-120). Malden, MA: Blackwell Publishing.

Salas, E., Prince, C., Bowers, C., Stout, R. J., Oser, R. L, & Cannon-Bowers, J. A. (1999). A methodology for enhancing crew resource management training. *Human Factors, 41*, 161-172.

Salas, E., Rhodenizer, L., & Bowers, C. A. (2000). The design and delivery of crew resource management raining: Exploiting available resources. *Human Factors, 42*, 490-511.

Salas, E., Wilson, K. A., Burke, C. S., & Wightman, D. C. (2006). Does CRM training work? An update, extension, and some critical needs. *Human Factors, 48,* 392-412.

Schmidt, R. A., & Björk, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts in training. *Psychological Science, 3*, 207-217.

Sexton, J. B., Thomas, E. J., & Helmreich, R. L. (2000). Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. *British Medical Journal, 320*, 745-749.

Shaw, J. D., Duffy, M. K., & Stark, E. M. (2000). Interdependence and preference for group work: Main and congruence effects of the satisfaction and performance of group members. *Journal of Management, 26*, 259-279.

Shea, G. P., & Guzzo, R. A. (1987). Groups as human resources. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resource management* (Vol. 5, pp. 323-356). Greenwich, CT: JAI Press.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing reliability. *Psychological Bulletin, 86,* 420-428.

Smith-Jentsch, K. A., Campbell, G. E., Milanovich, D. M., & Reynolds, A. M. (2001). Measuring teamwork mental models to support training needs assessment, development, and evaluation: two empirical studies. *Journal of Organizational Behavior, 22*, 179-194.

Smith-Jentsch, K. A., Mathieu, J. E., & Kraiger, K. (2005). Investigating linear and interactive effects of shared mental models of safety and efficiency in a field setting. *Journal of Applied Psychology, 90*, 523-535.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15*, 72-101.

Stone, R. B., & Babcock G. L. (1988). Airline pilot's perspective. In E.L. Wiener and D.C. Nagel (Eds.), *Human factors in aviation* (pp. 529-560). London, United Kingdom: Academic Press.

Stout, R. E., Salas, E., & Carson, R. (1994). Individual task proficiency and team process behavior: What's important for team functioning? *Military Psychology, 6*, 177-192.

Stout, R. J., Salas, E., & Fowlkes, J. E. (1997). Enhancing teamwork in complex environments through team training. *Group Dynamics: Theory, Research, and Practice, 1*, 169-182.

Stout, R. J., Salas, E., & Kraiger, K. (1997). The role of trainee knowledge structures in aviation team environments. *The International Journal of Aviation Psychology, 7*, 235-250.

Sundstrom, E., De Meuse, K. P., & Futrell, D. (1990). Work teams: Applications and effectiveness. *American Psychologist, 45*, 120-133.

Tannenbaum, S. I., Beard, R. L., & Salas, E. (1992). Team building and its influence on team effectiveness: An examination of conceptual and empirical developments. In K. Kelley (Ed.), *Issues, theory, and research in industrial/organizational psychology*. Amsterdam: Elsevier.

Taylor, J. C., & Patankar, M. S. (2000). The role of communication in the reduction of human error. *Proceedings of the 14th Annual Human Factors in Aviation Maintenance Symposium* (pp. 1-27). Vancouver, B.C.

Tesluk, P., Mathieu, J. E., Zaccaro, S. J., & Marks, M. (1997). Task and aggregation issues in the analysis and assessment of team performance. In M T. Brannick, C. Prince, and E. Salas (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 197-224). Mahwah, NJ: Erlbaum.

Thompson, J. D. (1967). *Organizations in action: Social science bases of administrative theory.* New York, NY: McGraw-Hill.

Tjosvold, D. (1986). The dynamics of interdependence in organizations. *Human Relations, 39*, 517-540.

Tziner, A., Eden, D. (1985). Effects of crew composition on crew performance: Does the whole equal the sum of its parts? *Journal of Applied Psychology, 70*, 85-93.

United Transportation Union (2006). Electronic Train Management System (ETMS) Waiver Hearing; Testimony of United Transportation Union, Paul Thompson, Rick Marceau and James Stern. Retrieved April 25th, 2008 from www.ble-t.org/pr/pdf/utu_etms_testimony.pdf

Urban, J. M., Bowers, C. A., Cannon-Bowers, J. A., & Salas, E. (1995). The importance of team architecture in understanding team processes. In M. Beyerlein, D. A. Johnson, & S. T. Beyerlein (Eds.), *Advances in interdisciplinary studies of work teams* (Vol. 2, pp. 205-228). Greenwich, CT: JAI Press.

U.S. Department of Labor (2006). O*NET version 8.0. Washington, DC. Retrieved November 26, 2006, from the World Wide Web: http://online.onetcenter.org /link/summary/53-4031.00.

Van de Ven, A. H., Delbecq, A., & Koenig, R. (1976). Determinants of coordination modes within organizations. *American Sociological Review, 41*, 322-338.

Van de Ven, A. H., & Ferry, D. L. (1980). *Measuring and assessing organizations*. New York: John Wiley & Sons.

Van der Vegt, G., Emans, B., & Van de Vliert, E. (1998). Motivating effects of task and outcome interdependence in work teams. *Group and organization management, 23*, 124-143.

Van der Vegt, G., Emans, B., & Van de Vliert, E. (2000). Team member's affective responses to patterns of intragroup interdependence and job complexity. *Journal of Management, 26*, 633-655.

Van der Vegt, G., Emans, B., & Van De Vliert, E. (2001). Patterns of interdependence in work teams: A two-level investigation of the relations with job and team satisfaction. *Personnel Psychology, 54*, 51-69.

Von Thaden, T. L., Steelman, K. S. (2005). Classifying crew performance failures in commercial aviation accidents: Can we get the numbers right? *Proceedings of the Safety Across High-Consequence Industries Conference.* St. Louis, MO.

Wageman, R. (1995). Interdependence and group effectiveness. *Administrative Science Quarterly, 40*, 145-180.

Wageman, R. (2001). The meaning of interdependence. In M.E. Turner (Ed.), *Groups at work: Advances in theory and research* (pp. 197-217). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wageman, R., Hackman, J. R., & Lehman, E. (2005). Team diagnostic survey: Development of an instrument. *Journal of Applied Behavioral Science, 41*, 373-398.

Webber, S. S., Chen, G., Payne, S. C, Marsh, S. M., & Zaccaro, S. J. (2000). Enhancing team mental model measurement with performance appraisal practices. *Organizational Research Methods, 3*, 307-322.

Webber, S. S., & Klimoski, R. J. (2004). Crews: A distinct type of work team. *Journal of Business and Psychology, 18*, 261-279.

Weick, K. (1979). *The social psychology of organizing.* Reading, MA: Addison-Wesley.

APPENDIX A

SAMPLE SIZES, OBSERVED POWER & EFFECT SIZES

Hypothesis	*N*	*Observed Power*	*Effect size (Partial η^2)*
1. DV = PTI, IV = Training (T1, T2)			
1A = Acc Team-relatedness	*63*	.42	.05
1B = Sim Team-relatedness	*62*	.85	.13
1C = Acc Team Workflow Pattern	*64*	.35	.04
1D = Sim Team-relatedness	*63*	.05	.00
2. DV = PTI at T1, IV = Crew Type (LOC, MOW).			
2A = Acc Team-relatedness	*63*	.12	.01
2B = Sim Team-relatedness	*62*	.12	.02
2C = Acc Team Workflow Pattern	*64*	.05	.00
2D = Sim Team Workflow Pattern	*63*	.08	.00
3. DV = PTI, IV = Training (T1, T2) × Crew Type (LOC, MOW).			
3A = Acc Team-relatedness	*63*	.08	.01
3B = Sim Team-relatedness	*62*	.78	.11
3C = Acc Team Workflow Pattern	*64*	.32	.04
3D = Sim Team Workflow Pattern	*63*	.11	.01
4. DV = PTI, IV = Training (T1, T2)			
4A = Acc Team-relatedness	*63*	.56	.07
4B = Sim Team-relatedness	*62*	.33	.04
4C = Acc Team Workflow Pattern	*64*	.16	.01
4D = Sim Team Workflow Pattern	*63*	.16	.02
5. DV = PTI, IV at T1, IV = IPE (HIGH, LOW)			
5A = Acc Team-relatedness	*63*	.06	.00
5B = Sim Team-relatedness	*62*	.07	.00
5C = Acc Team Workflow Pattern	*64*	.06	.00
5D = Sim Team Workflow Pattern	*63*	.06	.00

Hypothesis	*N*	*Observed Power*	*Effect size (Partial η^2)*
6. DV = PTI, IV = Training (T1, T2) × IPE (HIGH, LOW).			
6A = Acc Team-relatedness	*63*	.56	.02
6B = Sim Team-relatedness	*62*	.33	.06
6C = Acc Team Workflow Pattern	*64*	.16	.07
6D = Sim Team Workflow Pattern	*63*	.16	.00
8. DV = PTW (RMAQ), IV = Training (T1, T2)			
8A = Acc RMAQ	*140*	1.00	.17
8B = Sim RMAQ	*108*	.54	.04
9. DV = PTW at T1, IV = Crew Type			
9A = Acc RMAQ	*140*	.37	.02
9B = Sim RMAQ	*108*	.09	.00
10. DV = PTW, IV = Training (T1, T2) × Crew Type (LOC, MOW).			
10A = Acc RMAQ	*140*	.44	.02
10B = Sim RMAQ	*108*	.05	.00
11. DV = PTW (RMAQ), IV = Training (T1, T2)			
11A = Acc RMAQ	*71*	.92	.14
11B = Sim RMAQ	*70*	.67	.08
12. DV = PTW at T1, IV = IPE			
12A = Acc RMAQ	*71*	.07	.00
12B = Sim RMAQ	*70*	.05	.00
13. DV = PTW, IV = Training (T1, T2) × IPE (HIGH, LOW).			
13A = Acc RMAQ	*71*	.26	.03
13B = Sim RMAQ	*70*	.21	.02
Extremely small sample precluded any quantitative statistical analysis on Hypotheses 14-26.			

Note. DV = dependent variable, IV = independent variable, PTI = perceptions of task interdependence, Acc = accuracy, Sim = similarity, PTW = perceptions of teamwork, IPE = interpositional experience.

APPENDIX B

MEASURE OF TEAM-RELATEDNESS

OVERALL TEAM–RELATEDNESS: The extent to which you can not successfully perform your job alone without other members of the team.

Using the scale below, please shade in the number corresponding to the level of team–relatedness that BEST describes your **OVERALL job as a ____________________________.**
(fill in your specific job)

① = This job can be performed alone
② = This job can be performed alone with little difficulty
③ = This job can be performed alone with some difficulty
④ = This job can be performed alone only with great difficulty
⑤ = This job **CANNOT** be perform alone under any circumstances

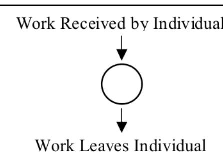
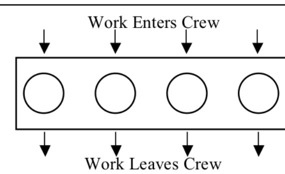
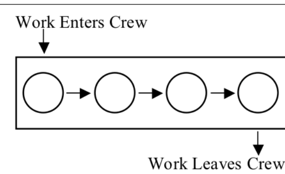
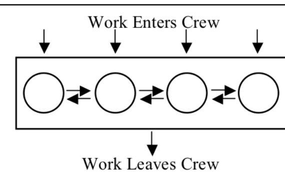
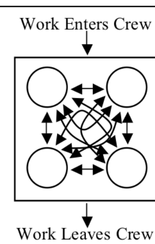
	Overall Job	Team–Relatedness
1.	OVERALL job as a __________________________. *(fill in your specific job)*	① ② ③ ④ ⑤

APPENDIX C

MEASURE OF TEAM WORKFLOW PATTERN

<u>OVERALL CREW WORKFLOW PATTERN</u>: The paths by which work and/or information flows through your train crew in order to optimally perform your overall job as a ________________________.
(fill in your specific job)

The chart below presents five crew workflow patterns as well as a description and diagram for each pattern. Please rank order the five patterns (1 = high, 5 = low) in terms of the extent to which they are descriptive of the work activities in your job. That is, **the pattern that BEST describes the work activities in your job would be ranked number 1 and the pattern that is least descriptive of your work activities would be ranked number 5.**

Description of Crew Workflow Pattern	Illustration	RANK (1-5)
Work and activities are ***<u>NOT</u>*** performed as a member of a crew; they are performed alone outside the context of the crew. Work and activities are performed by an individual working ***<u>ALONE</u>***, ***<u>NOT</u>*** in a crew.		___
Work and activities are performed separately by all crew members and work does not flow between members of the crew.		___
Work and activities flow from one member to another in the crew, but mostly in one direction.		___
Work and activities flow between crew members in a back–and–forth manner over a period of time.		___
Work and activities come into the crew and members must diagnose, problem solve, and/or collaborate as a crew in order to accomplish the crew's task.		___

APPENDIX D

RAILROAD MANAGEMENT ATTITUDES QUESTIONNAIRE (RMAQ)

Please indicate the degree to which you agree with the following statements using the scale to the right .	1 = Strongly Disagree 2 = Disagree 3 = Slightly Disagree 4 = Neither agree nor Disagree 5 = Slightly Agree 6 = Agree 7 = Strongly Agree

1.	I should let other members of my train crew know when my workload is becoming (or about to become) excessive.	① ② ③ ④ ⑤ ⑥ ⑦
2.	The engineer should verbalize plans for procedures or maneuvers and should be sure that the information is understood and acknowledged by other members of his/her train crew.	① ② ③ ④ ⑤ ⑥ ⑦
3.	Even when fatigued, I perform effectively during critical times on a trip.	① ② ③ ④ ⑤ ⑥ ⑦
4.	How I do my work does not influence anyone outside my train crew.	① ② ③ ④ ⑤ ⑥ ⑦
5.	Debriefing the trip can be a valuable learning experience.	① ② ③ ④ ⑤ ⑥ ⑦
6.	Engineers and conductors should be aware of and sensitive to the personal problems of other members of their train crew.	① ② ③ ④ ⑤ ⑥ ⑦
7.	Personal problems can adversely affect my performance.	① ② ③ ④ ⑤ ⑥ ⑦
8.	A debriefing and critique of procedures and decisions after each trip is an important part of developing and maintaining effective crew coordination in the locomotive.	① ② ③ ④ ⑤ ⑥ ⑦
9.	Train crew members should monitor each other for signs of stress or fatigue.	① ② ③ ④ ⑤ ⑥ ⑦
10.	Train crew members should mention their stress or physical problems to each other before or during a trip.	① ② ③ ④ ⑤ ⑥ ⑦
11.	I am less effective when stressed or fatigued.	① ② ③ ④ ⑤ ⑥ ⑦
12.	The pre-trip crew briefing is important for safety and for effective crew management.	① ② ③ ④ ⑤ ⑥ ⑦
13.	Effective communication between a train crew and an outside element (e.g. Train crew & dispatch) is as important as communication within a crew.	① ② ③ ④ ⑤ ⑥ ⑦
14.	Effective crew coordination requires train crew members to take into account the personalities of other members of their train crew.	① ② ③ ④ ⑤ ⑥ ⑦
15.	Train crew members should encourage each other to question procedures during normal operations and emergencies.	① ② ③ ④ ⑤ ⑥ ⑦
16.	Train crew members should alert others to their actual or potential work overloads.	① ② ③ ④ ⑤ ⑥ ⑦
17.	Good communications and crew coordination are as important as technical proficiency for the safety of travel.	① ② ③ ④ ⑤ ⑥ ⑦
18.	It is important to avoid negative comments about the procedures and techniques of other members of your train crew.	① ② ③ ④ ⑤ ⑥ ⑦
19.	My performance is not adversely affected by working with an inexperienced or less capable member of my train crew.	① ② ③ ④ ⑤ ⑥ ⑦
20.	Casual, social conversation in the locomotive cab during periods of low workload can improve train crew coordination.	① ② ③ ④ ⑤ ⑥ ⑦
21.	Most of the actions that I take do not influence or involve other people.	① ② ③ ④ ⑤ ⑥ ⑦
22.	My decision-making ability is as good in emergencies as in routine traveling conditions.	① ② ③ ④ ⑤ ⑥ ⑦

	1 = Strongly Disagree 2 = Disagree 3 = Slightly Disagree 4 = Neither agree nor Disagree 5 = Slightly Agree 6 = Agree 7 = Strongly Agree

23.	Because individuals function less effectively under high stress, good crew coordination is more important in emergency and abnormal situations.	① ② ③ ④ ⑤ ⑥ ⑦
24.	Proper Crew Resource Management by a crew can act as a fatigue counter-measure for individual crew members.	① ② ③ ④ ⑤ ⑥ ⑦
25.	I am more likely to make judgment errors in an emergency.	① ② ③ ④ ⑤ ⑥ ⑦
26.	A true professional does not make mistakes.	① ② ③ ④ ⑤ ⑥ ⑦
27.	Train crew members should avoid disagreeing with others because conflict creates tension and reduces crew effectiveness.	① ② ③ ④ ⑤ ⑥ ⑦
28.	A truly professional train crew member can leave personal problems behind when on the job.	① ② ③ ④ ⑤ ⑥ ⑦

APPENDIX E

CONVERSION OF TEAM WORKFLOW PATTERN SCALE TO PURELY IPSATIVE RANK-ORDERED SCALE

Example 1

	Team workflow pattern				
	No team job	Additive	Sequential	Reciprocal	Intensive
Uncorrected Rating	2	5	3	1	1
	↓	↓	↓	↓	↓
Corrected Ranking	3	5	4	1.5	1.5

Example 2

	Team workflow pattern				
	No team job	Additive	Sequential	Reciprocal	Intensive
Uncorrected Rating	2	3	3	4	4
	↓	↓	↓	↓	↓
Corrected Ranking	1	2.5	2.5	4.5	4.5

APPENDIX F

PRELIMINARY ANALYSIS OF 2-YEAR FOLLOW-UP DATA

Although the small sample size and low power precluded me from running any quantitative statistical analysis on the follow-up measures, means, standard deviations, and effect sizes are presented below. Specifically, means, standard deviations, and effect size estimates are discussed regarding their directionality (i.e., +, -), and magnitude of effect (i.e., *d*). Furthermore, because descriptive statistics based on repeated measures would decrease the sample size further, unlike the tables in the body of the paper, descriptive statistics displayed in the following tables are not based on repeated measures.

Two Year Follow-up on Knowledge Retention (Hypothesis 14)

Hypothesis 14 posited that participants' knowledge of CRM concepts would be higher immediately following CRM training (Time 2) than two years after training (Time 3). The mean number of knowledge test items correctly answered was 26.23 (SD = 4.38, n = 119) before training and 26.40 (SD = 3.20, n = 10) after training. Contrary to the hypothesis, the means suggests that there is no decrease in knowledge after a 2-year retention interval.

Two Year Follow-up Regarding Change in Participants' Perceptions of Task Interdependence (Hypothesis 15 - 18)

As a result of the follow-up survey's low return rate, only one or two participants in a MOW division had follow-up data. This precluded the use of my previous operationalization of similarity in which MOW crew members' perceptions were compared to a referent group consisting of other MOW participants in their specific division (see pg. 50-51). To remedy this, all subsequent analyses regarding the similarity of MOW crew members, utilized a referent group which included all other MOW participants, whether in their division or not. Similar to previous hypotheses, a summary of the results for the research hypotheses pertaining to the change in participants' perceptions of task interdependence after the 2-year retention interval are presented in Table F1. However, because no statistical analyses will be conducted to formally test the hypotheses, Table F1 simply indicates if the direction of change is consistent with the specific hypothesis and the effect size is large ($d > .80$).

Hypothesis 15 posited that across crew type, there would be a decrease in the accuracy and similarity of participants' perceptions of task interdependence from Time 2 (immediately after training) to Time 3 (two years later). A preliminary assessment of the means displayed in Table F2 suggests that the accuracy and similarity of both team-relatedness and team workflow pattern actually increased from Time 2 to Time 3. Thus, the direction on change was not consistent with Hypothesis 15.

Table F1

Hypotheses Pertaining to Two Year Follow-up on Perceptions of Task Interdependence

Hypothesis and dependent variable	Level of independent variable and proposed effect on	Δ consistent with hypothesis & d > .80[A]
Hypothesis 15 - main effect for Time (when crossed with Crew Type)		
ACC Team-Relatedness	Time 2 > Time 3	No
ACC Team Workflow Pattern	Time 2 > Time 3	No
SIM Team-Relatedness	Time 2 > Time 3	No
SIM Team Workflow Pattern	Time 2 > Time 3	No
Hypothesis 16 – interaction (Time × Crew Type)		
Decrease in ACC Team-Relatedness from Time 2 to Time 3	LOC > MOW	No
Decrease in ACC Team Workflow Pattern from Time 2 to Time 3	LOC > MOW	No
Decrease in SIM Team-Relatedness from Time 2 to Time 3	LOC > MOW	No
Decrease in SIM Team Workflow Pattern from Time 2 to Time 3	LOC > MOW	No
Hypothesis 17 - simple effect for Crew Type (at Time 3)		
ACC Team-Relatedness	MOW > LOC	No
ACC Team Workflow Pattern	MOW > LOC	No
SIM Team-Relatedness	MOW > LOC	No
SIM Team Workflow Pattern	MOW > LOC	No
Hypothesis 18 - main effect for Time (when crossed with IPE)		
ACC Team-Relatedness	Time 2 > Time 3	No
ACC Team Workflow Pattern	Time 2 > Time 3	No
SIM Team-Relatedness	Time 2 > Time 3	No
SIM Team Workflow Pattern	Time 2 > Time 3	No
Hypothesis 19 – interaction (Time × IPE)		
Decrease in ACC Team-Relatedness from Time 2 to Time 3	Lower IPE > Higher IPE	No
Decrease in ACC Team Workflow Pattern Time 2 to Time 3	Lower IPE > Higher IPE	No
Decrease in SIM Team-Relatedness from Time 2 to Time 3	Lower IPE > Higher IPE	No
Decrease in SIM Team Workflow Pattern Time 2 to Time 3	Lower IPE > Higher IPE	No
Hypothesis 20 - simple effect for IPE (at Time 3)		
ACC Team-Relatedness	Higher IPE > Lower IPE	No
ACC Team Workflow Pattern	Higher IPE > Lower IPE	Yes
SIM Team-Relatedness	Higher IPE > Lower IPE	No
SIM Team Workflow Pattern	Higher IPE > Lower IPE	Yes

Note. ACC=accuracy; SIM=similarity; LOC=locomotive crew; MOW=maintenance of way crew. IPE=interpositional experience. [A] indicates if means are consistent with hypothesis and magnitude of effect (i.e., d) is greater than .80.

Table F2
Accuracy and Similarity on Perceptions of Task Interdependence (i.e., D) by Crew Type Across Time 2 and 3

	Time 2			Time 3			*d*
	N	Mean	*SD*	*N*	Mean	*SD*	
ACCURACY							
Team-relatedness							
Overall	72	1.10	1.26	14	0.36	0.74	.72
LOC	10	0.40	0.84	8	0.13	0.35	.42
MOW	62	1.21	1.28	6	0.67	1.03	.46
Team workflow pattern							
Overall	76	3.04	2.18	15	1.69	2.22	.61
LOC	12	3.86	2.30	9	1.34	2.23	1.11
MOW	64	2.89	2.14	6	2.22	2.30	.30
SIMILARITY							
Team-relatedness							
Overall	71	1.05	0.78	14	0.65	0.51	.61
LOC	9	0.67	0.58	8	0.33	0.31	.73
MOW	62	1.10	0.79	6	1.07	0.41	.05
Team workflow pattern							
Overall	75	2.84	0.86	15	2.30	1.22	.52
LOC	11	2.83	0.99	6	2.03	1.11	.76
MOW	64	2.84	0.84	6	2.71	1.36	.12

Note. Lower mean indicates increased accuracy (or similarity). Negative *d* indicates decreased accuracy or similarity (consistent with hypothesis). LOC = locomotive crew members; MOW = maintenance of way crew members.

Moreover, Hypothesis 16 posited that decreases in the accuracy and similarity of participants' perceptions of task interdependence would be greater for locomotive compared to MOW crew members. However the descriptive statistics displayed in Table F2 suggest that not only does the accuracy and similarity of perceptions of task interdependence increase for all crew members over the 2-year retention interval, but the effect size of the increases are greater for locomotive compared to MOW crew members. This resulted in locomotive crew members having greater accuracy and similarity of perceptions of task interdependence after the retention interval (Time 3) than MOW crew members, which is the opposite of what was predicted by Hypothesis 17. Thus, the direction of change was not consistent with Hypotheses 16 or 17.

The descriptive statistics shown in Table F3 display similar results with regards to interpositional experience. Hypothesis 18 posited that for all levels of interpositional experience, there would be a decrease in the accuracy and similarity of participants' perceptions of task interdependence over the 2-year retention interval. However, preliminary assessment of the descriptive statistics in Table F3 suggests that overall, the accuracy and similarity of both team-relatedness and team workflow pattern actually increase from Time 2 to Time 3. Hypothesis 19 suggested that the posited decrease in the level of accuracy and similarity of perceptions of task interdependence from Time 2 to Time 3 would be greater for participants who have lower compared to higher levels of interpositional experience. Although the data are not consistent with the overall direction of change suggested by Hypothesis 19 (i.e., decrease), the pattern of results is. For example, as seen in Table F3, crew members with higher levels of interpositional experience showed a medium to large increase (d ranged from .43 to 1.28) in the accuracy and similarity of their perceptions. However, crew members with lower levels of interpositional experience showed either decreases (d = -.31), no change (d = -.02, .03) or smaller increases (d = .20) in the accuracy and similarity of their perceptions than their higher interpositional experience counterparts.

Finally, Hypothesis 20 predicted that the accuracy and similarity of perceptions of task interdependence at Time 3 would be greater for participants who have higher levels of interpositional experience, compared to participants with lower levels of interpositional experience. Descriptive statistics shown in Table F3 suggest this is the case for the accuracy and similarity of team workflow pattern, but not team-relatedness. Specifically, as indicated in Table 8, the direction of means for both the accuracy and similarity of team workflow pattern were consistent with Hypothesis 20 (i.e., higher interpositional experience > lower interpositional experience), and the magnitude of the effects were sufficiently large (*d* > .80) to take note.

Table F3
Accuracy and Similarity of Perceptions of Task Interdependence (i.e., D) by Interpositional Experience Across Time 2 and 3

	Time 2			Time 3			*d*
	N	Mean	*SD*	*N*	Mean	*SD*	
ACCURACY							
Team-Relatedness							
Overall	72	1.10	1.26	6	0.67	1.03	.37
Low IPE	33	0.88	0.96	3	0.67	1.15	.20
High IPE	39	1.28	1.45	3	0.67	1.15	.47
Team Workflow Pattern							
Overall	76	3.04	2.18	7	2.29	2.75	.30
Low IPE	35	3.33	2.11	4	3.39	3.16	-.02
High IPE	41	2.80	2.23	3	0.82	1.41	1.06
SIMILARITY							
Team-Relatedness							
Overall	71	1.05	0.78	6	0.86	0.65	.26
Low IPE	32	0.82	0.58	3	0.80	0.80	.03
High IPE	39	1.24	0.87	3	0.91	0.64	.43
Team Workflow Pattern							
Overall	75	2.84	0.86	7	2.78	1.64	.05
Low IPE	34	2.93	0.83	4	3.41	2.02	-.31
High IPE	41	2.77	0.89	3	1.93	0.26	1.28

Note. Lower means indicate increased accuracy (or similarity). Negative *d* indicates decreased accuracy or similarity (consistent with hypothesis). IPE = interpositional experience.

Two Year Follow-up Regarding Change in Participants' Perceptions of Teamwork (Hypotheses 21 – 26)

A summary of the results for the research hypotheses pertaining to the change in participants' perceptions of teamwork after the 2-year retention interval are presented in Table F4. However, because no statistical analyses will be conducted to formally test the hypotheses, similar to Table F1, Table F4 simply indicates if the direction of change is consistent with the specific hypothesis and the effect size is large ($d > .80$). Hypothesis 21 predicted that overall there would be a decrease in the accuracy and similarity of participants' perceptions of teamwork from Time 2 (immediately after training) to Time 3 (2-years later). A preliminary assessment of the overall means shown in Table F5 suggest this is not the case but in fact, accuracy and similarity increase slightly. Hypothesis 22 stated that the posited decreases in accuracy and similarity of participants' perceptions of teamwork from Time 2 to Time 3 would be greater for participants who belong to locomotive compared to MOW crews. Although the data are not consistent with the overall direction of change suggested in Hypothesis 22 (i.e., decrease), the pattern of results is. Specifically, locomotive crew members showed a decrease in the accuracy and similarity of perceptions of teamwork from Time 2 to Time 3, whereas MOW crew members showed an increase.

Finally, Hypothesis 23 predicted that the accuracy and similarity of perceptions of teamwork at Time 3 would be greater for MOW participants than locomotive participants. Descriptive statistics shown in Table F5 suggest this is the case for the similarity of crew members' perceptions teamwork, but not the accuracy of those perceptions. Specifically, the direction of means for both the accuracy and similarity of teamwork were consistent with Hypothesis 23 (i.e., MOW > locomotive), but the magnitude of the effect size was sufficiently large ($d > .80$) only for mean differences in the similarity of perceptions of teamwork. Thus although the effect size was sufficiently large ($d > .80$) and the direction of change was consistent with Hypothesis 23a, taken together, data regarding the influence of crew type on the long-term retention of participants' perceptions of teamwork were not consistent with Hypotheses 21-23.

Table F4

Hypotheses Pertaining to Two Year Follow-up on Perceptions of Teamwork

Hypothesis and dependent variable	Level of independent variable and proposed effect on	Δ consistent w/ hypothesis & $d > .80^{A}$
Hypothesis 21 - main effect for Time (when crossed with Crew Type)		
ACC Teamwork	Time 2 > Time 3	No
SIM Teamwork	Time 2 > Time 3	No
Hypothesis 22 – interaction (Time × Crew Type)		
Decrease in ACC Teamwork from Time 2 to Time 3	LOC > MOW	No
Decrease in SIM Teamwork from Time 2 to Time 3	LOC > MOW	No
Hypothesis 23 - simple effect for Crew Type (at Time 3)		
ACC Teamwork	MOW > LOC	No
SIM Teamwork	MOW > LOC	Yes
Hypothesis 24 - main effect for Time (when crossed with IPE)		
ACC Teamwork	Time 2 < Time 3	No
SIM Teamwork	Time 2 < Time 3	No
Hypothesis 25 – interaction (Time × IPE)		
Decrease in ACC Teamwork from Time 2 to Time 3	Lower IPE > Higher IPE	No
Decrease in SIM Teamwork from Time 2 to Time 3	Lower IPE > Higher IPE	No
Hypothesis 26 - simple effect for IPE (at Time 3)		
ACC Teamwork	Higher IPE > Lower IPE	No
SIM Teamwork	Higher IPE > Lower IPE	No

Note. ACC = accuracy; SIM = similarity; LOC = locomotive crew; MOW = maintenance of way crew. IPE=interpositional experience. [A] indicates if means are consistent with hypothesis and magnitude of effect (i.e., d) is greater than .80.

Table F5

Accuracy and Similarity of Perceptions of Teamwork (i.e., D) by Crew Type Across Time 2 and 3

	Time 2			Time 3			*d*
	N	Mean	*SD*	*N*	Mean	*SD*	
Accuracy of Teamwork							
Overall	155	10.21	3.01	15	9.82	2.16	.15
LOC	83	10.08	3.03	9	10.18	2.55	-.04
MOW	72	10.37	3.01	6	9.29	1.46	.46
Similarity of Teamwork							
Overall	123	8.14	2.48	15	7.64	1.74	.23
LOC	53	7.95	2.47	9	8.28	1.92	-.15
MOW	70	8.28	2.49	6	6.67	0.85	.87

Note. Lower mean indicates increased accuracy (or similarity). Negative *d* indicates decreased accuracy or similarity (consistent with hypotheses). LOC = locomotive; MOW = maintenance of way. Overall means are unweighted.

The descriptive statistics shown in Table F6 display similar results regarding interpositional experience. Hypothesis 24 posited that across all levels of interpositional experience, there would be a decrease in the accuracy and similarity of participants' perceptions of teamwork. However, preliminary assessment of the descriptive statistics suggests that the accuracy and similarity of participants' perceptions of teamwork actually increase from Time 2 to Time 3. Hypothesis 25 suggested that the posited decreases in the levels of accuracy and similarity of participants' perceptions of teamwork would be greater for participants who have lower compared to higher levels of interpositional experience. However, the pattern of results is opposite of what was predicted by Hypothesis 25. Specifically, participants with lower levels of interpositional experience showed a greater increase in the accuracy and similarity of perceptions of teamwork compared to participants with higher levels of interpositional experience. As a result, participants with lower levels of interpositional experience had more accurate and similar perceptions of teamwork at Time 3 than participants with higher levels of interpositional experience. This is the opposite of what was predicted by Hypothesis 26.

Table F6

Accuracy and Similarity of Perceptions of Teamwork (i.e., D) by Interpositional Experience Across Time 2 and 3

	Time 2			Time 3			*d*
	N	Mean	*SD*	*N*	Mean	*SD*	
Accuracy of Teamwork							
Overall	85	10.40	2.96	7	9.59	1.27	.36
Low IPE	39	10.11	2.88	4	8.75	0.46	.66
High IPE	46	10.64	3.02	3	10.71	1.12	-.03
Similarity of Teamwork							
Overall	84	8.21	2.42	7	7.09	1.41	.57
Low IPE	38	8.15	2.37	4	6.46	1.16	.91
High IPE	46	8.27	2.49	3	7.93	1.45	.17

Note. Lower score indicates increased accuracy (or similarity). Negative *d* indicates decreased accuracy or similarity (consistent with hypotheses). IPE=interpositional experience. Overall means are unweighted.

VITA

Tobin Bruce Kyte received a Bachelor of Science degree in psychology from Northern Arizona University in 1996. He received his Master of Arts degree in industrial/organizational psychology from New York University in 1999. His Doctor of Philosophy in psychology with a specialization in industrial/organizational psychology at Texas A&M University was completed in August of 2008. Dr. Kyte has completed internships at Avon Inc. and PepsiCo Inc. Furthermore, he has applied experience conducting job analyses as well as developing, implementing and evaluating performance appraisal systems and training programs at the Department of Multicultural Services (Texas A&M University) and the Texas Transportation Institute. He also has worked as an independent consultant in Austin Texas.

Dr. Tobin Bruce Kyte can be contacted through the Department of Psychology, Texas A&M University, College Station, TX., 77843-4235. His email address is: tobykyte@gmail.com.