COLLECTIVE EFFICACY IN A HIGH-FIDELITY SIMULATION OF AN AIRLINE OPERATIONS CENTER

by

Shanna Jinkerson

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts in Psychology

Middle Tennessee State University
2013

Thesis Committee:

Richard G. Moffett III, Ph.D., Chair

Glenn E. Littlepage, Ph.D., Committee Member,

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the many people who assisted me in the process of writing my thesis. First of all, I would like to thank my thesis advisor Rick Moffett for his dedication to the development of this project. He understood my priorities for this project, as well as professionally beyond it, and he helped me to reach my personal goals. Additionally, Glenn Littlepage was incredibly helpful in the data analysis process, and I am very thankful for all of his guidance. Amber Fritsch was also a key part of me completing this thesis. She helped immeasurably with organizing the data, as well as serving as a subject matter expert concerning the simulations.

I would also like to thank my parents for supporting me through this project, as well as all of my professional and educational pursuits. Finally, I would like to thank all of my colleagues in this program. They were always there for me, even after they graduated, and I know some of the friendships I made will last a lifetime.

ABSTRACT

This study investigated the relationships between collective efficacy, teamwork, and team performance. Participants were placed into teams, where they worked together in a high-fidelity simulation of an airline operations center. Each individual was assigned a different role to represent different jobs within an airline (Flight Operations Coordinator, Crew Scheduling, Maintenance, Weather, Flight Scheduling, or Flight Planning.) Participants completed a total of three simulations with an After Action Review between each. Within this setting, both team performance and teamwork behaviors were shown to be positively related to expectations for subsequent performance (collective efficacy). Additionally, teamwork and collective efficacy were not shown to be concomitantly related to subsequent team performance. A chi-square test was used to evaluate existence of performance spirals, and they were not supported. The results of this study were likely impacted by lack of power, as well as a lack of consistency across the three simulations.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER I

INTRODUCTION

Organizations in the public, private, and military sectors increasingly utilize teams (Goodwin, Burke, Wildman, & Salas, 2009; Kozlowski & Ilgen, 2006), and thousands of studies have been conducted to research teams and team functioning (Guzzo & Shea, 1992; Hackman, 1987; Kozlowski & Ilgen, 2006). Even so, more research is being done every day, as teams are still not fully understood. For teams to be effective in complex environments, they must be able to engage in both taskwork and teamwork behaviors (Krokos, Backer, Alonso, & Day, 2009). Global competition requires teams to be flexible and able to adapt to a continuously changing environment (Kozlowski & Bell, 2003).

Flexibility within teams is particularly important for the aerospace industry. For example, airline flight operations teams must work together seamlessly to oversee critical aviation functions. Many times, flight operations teams are physically collocated within flight operations centers, command centers, and control centers, known as "the hub of communication, information, and coordination" (Sanders, 2013, p. 3). Within these centers, teams composed of multiple aviation disciplines must collaborate to oversee safe and efficient operations of the airline, as well as each individual airplane. Flight operations teams input and monitor information concerning weather, fuel, airplane weight and balance, maintenance problems and routine checks, and schedules of crew members. Additionally, flight operations teams must manage any problems that occur during flight, such as medical emergencies or terrorist attacks (Sanders, 2013).

Team Performance

Team performance combines the success of individual taskwork with team interactions (Dorsey et al., 2009). When completing a highly interdependent task, individual actions are only successful when integrated as a team. High levels of team performance are critical to the success of airlines. Flight operations teams must work together to coordinate tasks and information under immense time pressure. Failing to do so can have serious repercussions, ranging from causing flight delays to jeopardizing safety (Cannon-Bowers, Salas, & Converse, 1993; NEXTOR, 2010). Even if passengers are safe, flight delays and cancellations have more negative repercussions than frustrating passengers; they have also been approximated to cause losses of over $31 billion per year in the United States. The estimated delay cost breakdown is $8.3 billion to airlines, $16.7 billion to passengers, $2.2 billion from lost demand, and a $4 billion impact on GDP.

While delays clearly have a direct negative impact on airlines, the indirect effects can also be costly. Delays can cause passengers to choose to use other airlines in the future, or to simply not fly at all. Delays can cause airlines to need to raise ticket cost, which can further deter potential passengers and raise costs for businesses that use air transportation. However, some delays cannot and should not be avoided, due to problems that can put the airline or passengers in danger (NEXTOR, 2010). Even so, calculating delay times during a flight operations center simulation is one objective way to measure team performance.

Team performance can also be measured by outside observers or task participants, or the simulation equipment itself can record performance data (Krokos et al., 2009). One type of rating scale that observers can use is a behaviorally anchored rating scale

(BARS). This type of scale focuses on performance quality, and it provides behavioral examples with corresponding numerical values for observers to select in rating participants (Pulakos, 2009). BARS allow raters to understand what performance at the top, midpoint, and bottom of each scale looks like (Hauenstein, Brown, & Sinclair; 2010). By providing anchors, BARS scales can increase inter-rater reliability by increasing common understanding of the different dimensions (Smith & Kendall, 1963).

Teamwork Behavior

Teamwork behaviors help teams to reach their goals and to accomplish given tasks (Tasa, Taggar, & Seijts, 2007). Marks, Mathieu, and Zaccaro (2001) categorize teamwork behavior into three types of processes: interpersonal processes, action phase processes, and transition phase processes. These researchers state that interpersonal processes refer to members of the group engaging in conflict management, encouraging and supporting each other, and regulating emotions as necessary. Action processes refer to taskwork, such as construction workers putting up walls. Transition processes refer to planning for the completion of a task or evaluation of progress toward a goal, such as construction workers reviewing blueprints and prioritizing which tasks should be completed first (Marks et al., 2001).

In the context of this study, the simulations represent the action phases, during which teams make decisions and actively track their progress toward the goal of running an efficient airline. Each simulation is followed up by an After Action Review, which represents the transition phase. The military has used After Action Reviews for decades, and they are increasingly being used in aviation and other workplace settings (Tannenbaum & Cerasoli, 2013). During the After Action Review, each team is debriefed about their past performance. In the After Action Review, a facilitator leads a discussion of past performance by giving specific feedback and encouraging future planning. Participants discuss the successes and failures of the simulation, what behaviors added to the successes and failures, and how to improve in upcoming simulation(s). Interpersonal processes are important throughout the action processes in the simulations and the transition processes in the After Action Reviews, as indicated in Marks et al. (2001). Each type of teamwork behavior is integral to the success of a flight operations team.

Collective Efficacy

The construct collective efficacy developed as a different level of self-efficacy. Self-efficacy “refers to the beliefs in one’s capabilities to organize and execute the courses of action required to produce given attainments” (Bandura, 1997, p. 3). It refers to the degree to which an individual believes he or she can achieve a certain outcome (Bandura, 1997). Self-efficacy refers to the individual, while collective efficacy refers to a group or a team. Collective efficacy refers to a team’s shared perception of whether or not they will be able to complete a specific task (Bandura, 1997). Collective efficacy is important because it is one construct that can affect team functioning (Bandura, 1997).

Collective efficacy is one type of a team capability belief; the other is group potency, which focuses on a shared belief in the team’s ability to perform across any task. Group potency is a global evaluation referencing the team’s general capability, and it is measured using items such as “This team has confidence in itself,” and “No task is too tough for this team” (Guzzo, Yost, Campbell, & Shea, 1993). Alternatively, collective efficacy refers to a team’s shared perception of whether or not they will be able to complete a specific task, rather than any task they might encounter (Bandura, 1997). Both group potency and collective efficacy reflect the team's confidence in its ability to perform, and some studies have used these terms interchangeably (Jung & Sosik, 2003). However, they reflect important theoretical differences (Gibson, Randel, & Earley, 2000; Gully, Incalcaterra, Joshi, & Beaubien, 2002; Mathieu, Maynard, Rapp, & Gibson, 2008; Stajkovic, Lee, & Nyberg, 2009), and collective efficacy has been shown to fully mediate the relationship between group potency and performance (Stajkovic et al., 2009). Collective efficacy is the construct measured in this study.

Collective efficacy is important because it impacts a team's decisions to take action and put enough effort or enough time to achieve the goal at hand (Stajkovic et al., 2009). Two meta-analyses (Gully et al., 2002; Stajkovic et al., 2009) have shown the connection between collective efficacy and team performance, yet there is minimal research into its antecedents. More research is needed, so that managers and team members can play an active role in creating conditions that build collective efficacy (Tasa et al., 2007).

As researchers have begun to look into the antecedents of collective efficacy, it became apparent that many antecedents of collective and self-efficacy are similar; however, there are additional antecedents for collective efficacy, as team performance requires additional factors (Chen & Bliese, 2002; Gibson, 1999; Gully et al., 2002; Tasa et al. 2007). Specifically, Tasa et al. (2007) showed how collective efficacy is influenced by the aggregate teamwork behaviors in a team.

Another antecedent of collective efficacy is task-relevant knowledge (Tasa et al., 2007). Teams that possess more task-relevant knowledge perform better than teams with less task-relevant knowledge (LePine, Hollenbeck, Ilgen, & Hedlund, 1997). Without task-relevant knowledge, teams do not have the same degree of confidence in their ability to perform, meaning that they are unlikely to develop high levels of collective efficacy (Tasa et al., 2007). Task-relevant knowledge is not measured in this study, but a certain level of knowledge can be assumed, as participants in the aerospace simulation were undergraduate seniors majoring in aerospace, as well as recipients of specific training in the jobs in the simulation.

Collins and Parker (2010) further categorize collective efficacy into two types: team outcome efficacy and team process efficacy. Team outcome efficacy focuses on the team's confidence in their ability to achieve a certain quantity or quality of outputs. Alternatively, team process efficacy is a team's confidence in their ability to work together on processes for the entire existence of the team, including establishing goals, coordinating tasks, and managing interpersonal relationships. The importance of team process efficacy is reflected in Gersick (1988)'s punctuated equilibrium model, which states that group progress is led by members' awareness of time and deadlines. The emphasis in time and deadlines in Gersick's (1988) model is reflected in team process efficacy's emphasis on establishing goals and coordinating tasks (Collins & Parker, 2010).

Distinguishing between team outcome efficacy and team process efficacy supports the distinction between taskwork and teamwork (Marks et al., 2001). Collins and Parker's (2010) study showed that team outcome efficacy was a stronger predictor or team performance, but it did not support team citizenship. In their study, team process efficacy did not impact team performance as strongly for a relatively short term project, but it may have a greater effect for longer projects. This study does not differentiate between team outcome efficacy and team process efficacy. However, Collins and Parker's (2010) measure of team process efficacy is related to this study's measure of teamwork; their measure was derived using concepts from Marks et al.'s (2001) discussion of teamwork, while this study directly used Marks et al.'s (2001) teamwork measure.

Marks et al. (2001) discuss how emergent states occur as a result of team experiences, and these emergent states become inputs in following outcomes and processes. Collective efficacy can be considered an emergent state, formed by individual team member behavior. Mean levels of collective efficacy change over time as the team works together, completes tasks, and gathers feedback about each other and about the context, task, and process (Bandura, 1997; Gibson, 1999; Lindsley, Brass, & Thomas, 1995). When team members observe individuals engaging in behaviors that are beneficial to the team, their collective efficacy should rise, as they see that their team is capable (Tasa et al., 2007). Likewise, when team members see individuals behave in ways that are destructive or not supportive of the team, their collective efficacy should decrease (Tasa et al., 2007).

Collective Efficacy and Teamwork Behavior

Gibson (1999) theorized that "group efficacy forms as group members collectively acquire, store, manipulate, and exchange information about each other and about their task, context, process, and prior performance" (p. 138). Tasa et al. (2007) further showed that between-group differences in teamwork behaviors can affect the formation of both team performance and collective efficacy. Specifically, they stated that individuals form their collective efficacy beliefs based on observed behaviors and interactions of fellow team members. For example, they found that individuals' collective efficacy beliefs increased based on observations of positive teamwork behaviors. These researchers discovered that individual collective efficacy predicted teamwork behavior, and teamwork behavior was positively related to subsequent collective efficacy.

Additionally, collective efficacy has also been shown to moderate the relationship between agreeableness and interpersonal teamwork behavior (Tasa, Sears, & Schat, 2011). Individuals high in the personality trait of agreeableness are naturally inclined to engage in interpersonal teamwork behavior, yet they are even more likely to do so when surrounded by environment of high collective efficacy (Tasa et al., 2011).

Collective Efficacy and Team Performance

Collective efficacy is one antecedent of team performance (Gully et al., 2002; Mathieu, Rapp, Maynard, & Mangos, 2010; Stajkovic, Lee, & Nyberg, 2009; Tasa et al., 2007). A meta-analysis by Stajkovic et al. (2009) showed that the relationship between collective efficacy and team performance was positive (r = .37) and significant (p < .01). However, this only shows part of the relationship, as collective efficacy and team performance is a cyclic relationship. Collective efficacy affects team performance, and team performance affects later collective efficacy, known as an efficacy-performance spiral (Lindsley et al., 1995; Tasa et al., 2007). Each variable alternates between being the cause and the effect. Spirals can either be downward, upward, or self-correcting. In a downward spiral, low collective efficacy leads to poor performance, which in turn leads to low collective efficacy. In an upward spiral, the opposite is true. In a self-correcting spiral, a decrease in performance or collective efficacy is followed by an increase in either performance or collective efficacy. This self-correction indicates that the group is able to use their experience to make adjustments and improve (Lindsley et al., 1995).

A summary of expected relationships can be found in Appendix A.

Hypothesis 1: Team performance and teamwork behaviors are positively related to expectations for subsequent performance (collective efficacy).

Hypothesis 2: Collective efficacy and teamwork behaviors are positively related to subsequent team performance.

Research Question 1: Is the change in team performance concomitantly related to subsequent changes in collective efficacy?

Research Question 2: Is the change in collective efficacy concomitantly related to subsequent changes in team performance?

CHAPTER II

METHODS

Participants

Participants in this study were students in an aerospace capstone class. Students in this class were undergraduate senior aerospace majors. These participants had completed coursework in aerospace specializations, and they entered this study with task-relevant knowledge. One component of this class was the simulation used in this study, but students were allowed to withhold their data from analysis without penalty. The class professor assigned participants to teams of approximately ten participants each. Teams were formed based on the major concentrations of students (e.g., aerospace administration, aerospace technology, flight dispatch, maintenance management, and professional pilot). However, certain team members were excluded from this study because they were not physically collocated in the same room, yielding team sizes ranging between three and seven team members. This study uses archival data collected for other research, and this study only uses a subset of that data. The data were collected across three semesters for a total of 15 groups.

Procedures

Participation in this simulation was necessary for students to receive course credit. Participants were asked to roleplay as airline employees, working at their first job. The simulation began with students participating in an organizational socialization process (i.e., onboarding) to the fictional company Universal E-Lines. Lab staff, roleplaying representatives of Universal E-Lines, gave a one hour PowerPoint presentation, which introduced participants to the company, a small southeastern United States airline with 30 aircraft, two hub airports, and 14 spoke airports. Participants were also acquainted with the organizational goals of Universal E-lines, which are to "Ensure that no aircraft errors result in the loss of human life, have minimal delays, provide as much profit for the airline as possible, and maintain the professional image of Universal E-lines" (NASA FOCUS Lab, 2013). Each participant learned about the layout of the simulation workspaces. He or she was then assigned to a team and then participants were further appointed to one of six specific airline roles within the team: Flight Operations Coordinator, Crew Scheduling, Maintenance, Weather, Flight Scheduling, or Flight Planning (see Appendix B for job descriptions of each role).

Following the onboarding, subject matter experts provided a one-hour long individual training for each participant on how to complete their specific roles. Participants were also given an overview of all other simulation airline roles. This training further included an introduction to lab resources and technology. For example, all participants had computers with two screens, which gave them access to the software necessary for each unique role. The training allowed them to learn the software used by their role. Participants were also able to practice at their workstations, helping them to prepare for the simulations. Additionally, there were flat panel displays on the walls that were viewable by all participants. These screens showed aircraft radar, weather maps, departure and arrival times, and the performance measures of delay time and lost revenue. Furthermore, participants learned about the three available communication channels: verbal face-to-face, audio using headsets, or text using an instant messenger system on their computers (see Appendix C for pictures of the lab workspaces). This training helped to prepare students with the technical knowledge required to participate in the simulations.

Each simulation role had specific tasks to complete. The Flight Operations Coordinator oversees the flight operations team and is responsible for clearing all arrivals and departures. This role must take in information from all the other positions and integrate it in making decisions. Crew Scheduling tracks duty times of pilots and flight crews, making sure that they stay within the set number of hours they are legally allowed to work. Maintenance monitors necessary repairs, both planned and unplanned, ensuring that aircraft are airworthy. Weather tracks current meteorological data, sharing information about any weather-related safety concerns. Flight Scheduling closely assists the Flight Operations Coordinator by tracking aircraft arrivals and departures, as well as helping to stay on schedule. Flight Planning reviews weight data regarding fuel, cargo, and passengers for each flight.

Other participants in this simulation were assigned other roles (Ramp Tower Coordinator, Pseudo Pilot, Flight Simulator Pilot), but they are not included in this study because they were not physically collocated in the same room, making it more difficult for team level emergent states, such as collective efficacy, to form. Additionally, excluding these positions allowed the study to focus on one team, rather than a multi-team system.

Simulation. Each team participated in three flight simulations, in which the team was responsible for overseeing the airline to ensure that as many aircraft depart and arrive on time as possible. Each simulation lasted roughly two and a half hours, during which students manage up to 60 flight events (takeoffs and landings). Participants were observed by lab staff that assessed teams using a behaviorally anchored rating scale (BARS) of team performance. Throughout the simulation, teams were subjected to events (i.e., triggers) that could affect the functioning of the airline. These triggers included such things as maintenance problems and health emergencies. The triggers add to the fidelity of the simulation, as airlines have issues to deal with on a day-to-day basis. These were developed by aerospace professors, and lab staff determined which ones to give which teams, as the triggers varied across simulations. Different triggers were used to decrease treatment contamination, as participants in the study were able to interact with other teams during their class, and each team needed to be able to react to the simulation without expectations of what other participants experienced.

The simulation used in this study was appropriate for several reasons. Similar to the simulation used by Tasa et al. (2007) to study collective efficacy, the task was highly complex, requiring the teams to constantly monitor multiple factors affecting performance. However, using Steiner's (1972) task typology, Tasa et al. (2007) acknowledged that the task used in their study could be completed as a disjunctive task, meaning that "only one person needs to perform well in order for the team to succeed" (p. 24). Alternatively, the simulation used in this study was a conjunctive task; "everyone's success is determined by the effectiveness with which the least proficient member operates" (Steiner, 1972, p. 28). While each team member had unique tasks, their efforts must combine to reach common goals. If any one individual chose to not participate or performed poorly, the flight operations team performance would suffer.

Because of the conjunctive nature of this task, it is also considered interdependent. When tasks are highly interdependent, it becomes harder for individuals to cognitively separate themselves from the team. As such, researchers have theorized that collective efficacy has a greater impact on performance in highly interdependent teams (Lindsley et al., 1995), and task interdependence has been shown to moderate the relationship between collective efficacy and performance (Gibson, 1999; Gully et al., 2002; Stajkovic et al., 2009).

This simulation also provides powerful training potential, which would not be possible with the actual task, as it would not be safe or practical (Stout, Cannon-Bowers, & Salas, 1996). Additionally, the simulation has high fidelity, increasing the extent to which results can be generalized beyond the simulation.

After Action Review. Following each simulation, participants received an After Action Review form to complete and bring to the After Action Review meeting the following week. In the After Action Review, a facilitator led a discussion of past performance by giving specific feedback and encouraging future planning. This discussion helped teams to debrief their past performance, as participants discussed the successes and failures of the simulation, what behaviors added to the successes and failures, and how to improve in upcoming simulation(s). This After Action Review gave participants an opportunity to engage in uninterrupted transition processes, as explained in Marks et al. (2001), and these transition processes are assessed as one part of the teamwork measure used in this study.

Measures

Team performance. Similar to Tasa et al. (2007), the first instance of team performance is referred to as initial team performance. This distinction is made because teams were still acclimating to the simulation, while the second and third team performance measures better reflect actual performance. This concept is also supported by a study of self-efficacy and performance, in which many participants lowered their self-efficacy after the first performance, indicating that they may have initially overestimated their abilities before fully understanding the task (Shea & Howell, 2000).

Team performance was measured with one objective measure (delay time) and one subjective measure, a rubric completed by observers. These two measures were used because they represent different aspects of team performance. Both measures (delay time and the observer rubric) were gathered for each of the three simulations.

Delay time is used as a performance measure because the main goal that simulation participants are trying to reach is for all aircraft to leave on time and arrive on time. Delay time is a continuous time variable that is measured in number of hours, and performance improvement is indicated by decreases in delay times. Delay time is calculated on a flight-by-flight basis using a formula that subtracts the time difference between the scheduled arrival and actual arrival times of each flight.

It is a sufficient measure of performance because all members of the flight operations team must be adequately completing their individual tasks for the delay time to be low; it pools the success of the team into one cohesive score. However, it is an imperfect measure. Participants were responsible for timestamping as each aircraft departs and arrives, yet in some cases, lab staff observed extreme errors in timestamping, which required lab staff to retroactively try to fix the worst participant errors to their best estimate. As such, from a quantitative perspective, delay time should not be the only measure of performance, which further emphasizes the importance of the observer rubric.

The rubric completed by outside observers included 10 items on a behaviorally anchored rating scale (BARS) on a scale of one to seven. This rubric includes three subscales: four problem solving items, two coordination items, and three information utilization items. Each item was rated on a scale of one to seven, where a rating of one indicates "Trainee Level," and a rating of seven indicates "Professional Level." For the purposes of this study, only the problem solving subscale was included as a performance measure because the other two subscales focused more on team processes and less on actual performance. The problem solving subscale is presented in Appendix D.

Teamwork. Teamwork was measured using Marks et al.'s (2001) team process taxonomy (see Appendix E). This measure uses 30 self-report items to measure extent on a unipolar Likert scale from one to five. A rating of one indicates "Not at all," and a rating of five indicates "To a very great extent." Participants were given this measure three times, once after each flight simulation. This measure provides an overall total scale score, as well as three subscores for action processes, transition processes, and interpersonal processes. For the purposes of this study, subscales will not be directly investigated. A reliability score was calculated for the overall teamwork scale using Cronbach's α and is presented in the Results section.

Teamwork Action Processes. This subscale has twelve items, and it focuses on four dimensions: monitoring progress toward goals, systems monitoring, team monitoring and backup responses, and coordination activities.

Teamwork Transition Processes. This subscale has nine items, and it focuses on three dimensions: mission analysis, goal specification, and strategy formulation and planning.

Teamwork Interpersonal Processes. This subscale has nine items, and it focuses on three dimensions: conflict management, motivating and confidence building.

Collective efficacy. Collective efficacy was measured using a 10-item scale based on a measure used by Quiñones (1995) and is presented in Appendix F. This scale measures agreement on a one to five scale, where a rating of one indicates "Strongly Disagree" and a rating of five indicates "Strongly Agree." Six of the ten items on this scale are reverse scored. This original measure was used to quantify self-efficacy by investigating participants' beliefs about their individual future performance in a training simulation. The measure used in this study modified the original measure by focusing specifically on the simulation and on participants' beliefs about their team's future performance. Participants were given this measure four times, once after the onboarding and then following each flight simulation (see Appendix A for timeline and sequence of all measures).

CHAPTER III

RESULTS

Missing Data Issues

Missing data was a concern for items measuring teamwork behavior and collective efficacy at the individual level and for aggregated observer rubric ratings of team performance. These two situations necessitated different approaches to how best address the missing data issues described below.

Teamwork behavior and collective efficacy. Prior to adjusting for missing data, 25 of 92 individuals were missing responses to at least one item on either the measure of teamwork behavior or collective efficacy. In order to preserve as many ratings as possible, individuals missing three or less item responses were adjusted using mean substitution. Within-subject means for each homogeneous scale were substituted for these item-level missing data, as explained by Switzer and Roth (2002).

No participants were missing four to nine item responses. However, other individuals were missing 10 or more items, indicating that they had no responses for an entire construct. In these cases, the individual was list-wise deleted from all analyses and not aggregated to the team level.

Team performance. Following aggregation, two of 17 teams were missing all observer rubric ratings for Simulation 1. As such, these two teams were list-wise deleted from all analyses, yielding an *n* of 15 teams for all analyses.

Data Aggregation Issues

Following the methods of Tasa et al. (2007), both teamwork behavior and collective efficacy were measured at the individual level and aggregated to the team level. Following Tasa et al.'s method, intraclass correlations (ICCs) were calculated to determine team-level reliability. ICC(1) was calculated using the Bartko formula for each scale to reflect the level of response variability at the individual level that is ascribed to team membership (Bliese, 2000). James (1982) reviewed multiple studies that used ICC(1) values, and they ranged from .00 to .50 with a median value of .12. Therefore, this study considers ICCs of .12 or higher as justification for aggregation.

Teamwork behavior. Chan (1998) theorized that using the mean of individual ratings is appropriate when the construct is being measured at the team level. As such, the means of individual teamwork ratings were aggregated to represent team-level teamwork behaviors. This method does not look at the teamwork behaviors of each individual within the team, nor does it imply that every team member displayed the same amount of teamwork behaviors. Rather, this method focuses on overall interpretations of teamwork behaviors (Tasa et al., 2007). The teamwork behavior scale was appropriate for aggregation in two of three simulations, according to the ICC(1) values. The ICC for teamwork in the first simulation was not statistically significant, ICC(1) = .12, $p = .06$. However, these scores were aggregated for this study because the construct is being conceived at the team level and thus should be aggregated. Specific ICC(1) values can be found in Table 1.

Collective efficacy. There are two primary ways that collective efficacy has been historically measured. The first is using group responses to a single questionnaire answered collectively by the team. The second is aggregation of individual responses, averaging them into one response per team. By definition, efficacy perceptions lie within the individual, so some researchers suggest that the aggregation method is a better reflection of how collective efficacy functions (Gully et al., 2002; Stajkovic et al., 2009). Consequently, the aggregation method was used in this study. The collective efficacy scale was appropriate for aggregation all four times it was measured, according to the ICC(1) values. Specific ICC(1) values can be found in Table 1.

Table 1

ICC(1) for Individual Collective Efficacy and Teamwork

Measure	*ICC(1)*	*F*	*p*
Collective Efficacy			
Estimated	.17	2.07	.020*
Simulation 1	.24	2.70	.002**
Simulation 2	.33	3.68	.000**
Simulation 3	.19	2.24	.011**
Teamwork			
Simulation 1	.12	1.71	.062
Simulation 2	.22	2.51	.004**
Simulation 3	.28	3.08	.001**

Note. * indicates correlation is significant at the .05 level (2-tailed). ** indicates correlation is significant at the .01 level (2-tailed). ICC(1) was calculated using the Bartko formula: ICC(1) = (MSB – MSW) / [MSB + ((k -1) * MSW)]

Inter-rater Reliability

As explained by Bliese (2000), ICC(1) was also used to test inter-rater reliability among raters completing the observer rubric measure of team performance. ICCs were evaluated using Cicchetti's (1994) guidelines for determining inter-rater reliability, and

ICC(1) strength ranged from fair (.40 to .59) to good (.60 to .74). Specific ICC(1) values can be found in Table 2.

Table 2

ICC(1) for Inter-rater Reliability of Observer Rubric Team Performance

Measure	*ICC(1)*	*F*	*p*
Team Performance			
Simulation 1 Observer Rubric	.44*	10.57	.000
Simulation 2 Observer Rubric	.68**	25.94	.000
Simulation 3 Observer Rubric	.59*	18.43	.000

Note. * indicates ICC(1) inter-rater reliability strength is fair. ** indicates ICC(1) inter-rater reliability strength was good. ICC(1) was calculated using the Bartko formula: ICC(1) = (MSB – MSW) / [MSB + ((k -1) * MSW)]

Descriptive Statistics

Table 3 contains descriptive statistics for all measures across all simulations. With the exception of the first simulation being lower than the estimated level, collective efficacy increased each time it was measured. The lowest collective efficacy score was observed following the first simulation (*M* = 4.17, *SD* = 0.42), and the highest collective efficacy score was observed following the third simulation (*M* = 4.35, *SD* = 0.46). Teamwork increased each time it was measured. The lowest teamwork score was observed following the first simulation (*M* = 3.64, *SD* = 0.39), and the highest teamwork score was observed following the third simulation (*M* = 4.08, *SD* = 0.46).

Table 3

Means and Standard Deviations for Aggregated Collective Efficacy, Teamwork, and Team Performance

Measure	*n*	*M*	*SD*
Collective Efficacy			
Estimated	15	4.18	0.35
Simulation 1	15	4.17	0.42
Simulation 2	15	4.25	0.49
Simulation 3	15	4.35	0.46
Teamwork			
Simulation 1	15	3.64	0.39
Simulation 2	15	3.94	0.40
Simulation 3	15	4.08	0.46
Team Performance			
Simulation 1 Delay Time	15	10.02	4.33
Simulation 2 Delay Time	15	5.49	2.16
Simulation 3 Delay Time	15	5.33	4.27
Simulation 1 Observer Rubric	15	3.32	0.89
Simulation 2 Observer Rubric	15	3.69	1.26
Simulation 3 Observer Rubric	15	4.24	1.25

Note. Collective efficacy is measured on a 5-point agreement scale, where a rating of 1 indicates "Strongly Disagree" and a rating of 5 indicates "Strongly Agree." Teamwork is measured on a 5-point extent scale, where a rating of 1 indicates "Not at all" and a rating of 5 indicates "To a very great extent." Team performance as measured by delay time is a continuous time variable measured in number of hours. Team performance as measured by the observer rubric is rated on a 7-point behaviorally anchored rating scale, where a rating of 1 indicates "Trainee Level" and a rating of 7 indicates "Professional Level."

Team performance also improved each time it was measured with both the objective (delay time) and subjective (observer rubric) measures. Delay time decreased each time it was measured, indicating performance improvement, as delay time should be as low as possible. The poorest delay time performance was observed during the first simulation (*M* = 10.02, *SD* = 4.33), and the best delay time performance was observed following the third simulation (*M* = 5.33, *SD* = 4.27). The lowest observer rubric rating was observed during the first simulation (*M* = 3.32, *SD* = 0.89), and the highest observer rubric rating was observed following the third simulation (*M* = 4.24, *SD* = 1.25).

Scale Reliability

Cronbach's alpha was used to evaluate scale reliability of each measure following aggregation. The reliability for estimated collective efficacy was α = .93, Simulation 1 collective efficacy scale reliability was α = .93, Simulation 2 collective efficacy scale reliability was α = .96, and Simulation 3 collective efficacy scale reliability was α = .96. As such, all collective efficacy scales were reliable. The reliability for Simulation 1 teamwork scale reliability was α = .98, Simulation 2 teamwork scale reliability was α = .98, and Simulation 3 teamwork scale reliability was α = .99. As such, all teamwork scales were reliable. The reliability for the Simulation 1 team performance scale as measured by the observer grading rubric was α = .98, the Simulation 2 observer rubric reliability was α = .99, and the Simulation 3 observer rubric reliability was α = .99. As such, each usage of the observer rubric was reliable. Table 4 also contains all reliability coefficients on the diagonal. No coefficient was calculated for delay time, as it is a single continuous time variable. Because of this, there are no individual items to assess for internal consistency.

Table 4

Correlations and Coefficient Alphas for Aggregated Collective Efficacy, Teamwork, and Team Performance

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Est. CE	**.93**												
2. S1 CE	.50	**.93**											
3. S2 CE	.35	.80**	**.96**										
4. S3 CE	.72**	.80**	.66**	**.96**									
5. S1 TW	.28	.38	.37	.30	**.98**								
6. S2 TW	.25	.68**	.76**	.58*	.81**	**.98**							
7. S3 TW	.68**	.69**	.52*	.91**	.48	.60*	**.99**						
8. S1 DT	.09	-.29	-.01	-.18	-.37	-.37	-.24	--					
9. S2 DT	.07	-.23	-.54*	-.23	-.05	-.43	-.20	.06	--				
10. S3 DT	-.32	.12	-.09	-.69**	.12	.17	-.64**	.17	.04	--			
11. S1 OR	.03	.46	.38	.24	-.01	.23	.08	-.37	-.14	.00	**.98**		
12. S2 OR	.14	.59*	.74**	.51	.47	.80**	.42	-.44	-.40	-.17	.34	**.99**	
13. S3 OR	.27	.43	.44	.65**	.01	-.36	.52*	-.36	-.26	-.65**	.18	.74**	**.99**

Note. n = 15. Coefficient alphas are presented in boldface along the diagonal. * indicates correlation is significant at the .05 level (2-tailed). ** indicates correlation is significant at the .01 level (2-tailed). Est. CE = Estimated Collective Efficacy Measure; S1 CE = Simulation 1 Collective Efficacy Measure; S2 CE = Simulation 2 Collective Efficacy Measure; S3 CE = Simulation 3 Collective Efficacy Measure; S1 TW = Simulation 1 Teamwork Measure; S2 TW = Simulation 2 Teamwork Measure; S3 TW = Simulation 3 Teamwork Measure; S1 DT = Simulation 1 Delay Time Measure of Team Performance; S2 DT = Simulation 2 Delay Time Measure of Team Performance; S3 DT = Simulation 3 Delay Time Measure of Team Performance; S1 OR = Simulation 1 Observer Rubric Measure of Team Performance; S2 OR = Simulation 2 Observer Rubric Measure of Team Performance; S3 OR = Simulation 3 Observer Rubric Measure of Team Performance. Delay times that are negatively correlated indicate an increase in performance, as delay times should be as low as possible.

Testing Hypothesis 1

Hypothesis 1 states, "Team performance and teamwork behaviors are positively related to expectations for subsequent performance (collective efficacy)." Team performance as measured by delay time was considered first, followed by team performance as measured by the observer rubric.

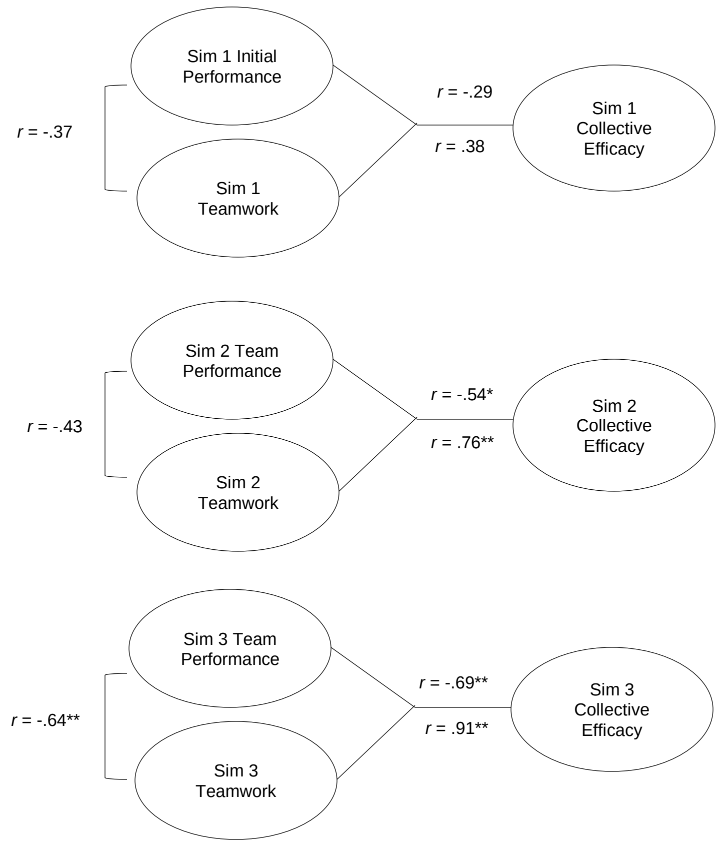


Figure 1. Hypothesis 1: Correlations between Team Performance as Measured by Delay Time, Teamwork, and Collective Efficacy.

Note. n = 15. * indicates correlation is significant at the .05 level (2-tailed). ** indicates correlation is significant at the .01 level (2-tailed). Sim = Simulation. Delay times that are negatively correlated indicate an increase in performance, as delay times should be as low as possible.

Correlations. To begin evaluating Hypothesis 1, relevant bivariate correlations were calculated, as seen in Figures 1 and 2. (All bivariate correlations can be found in Table 4.) No relevant correlations from Simulation 1 were significant. Even so, the correlation between Simulation 1 teamwork and Simulation 1 collective efficacy reached a medium effect size (r = .38, p = .16), as defined by Cohen (1988). The correlation between Simulation 1 team performance as measured by the observer rubric and Simulation 1 collective efficacy also reached a medium effect size, r = .46, p = .08. Additionally, the correlation between Simulation 1 delay time team performance and Simulation 1 collective efficacy nearly reached a medium effect size, r = -.29, p = .29.

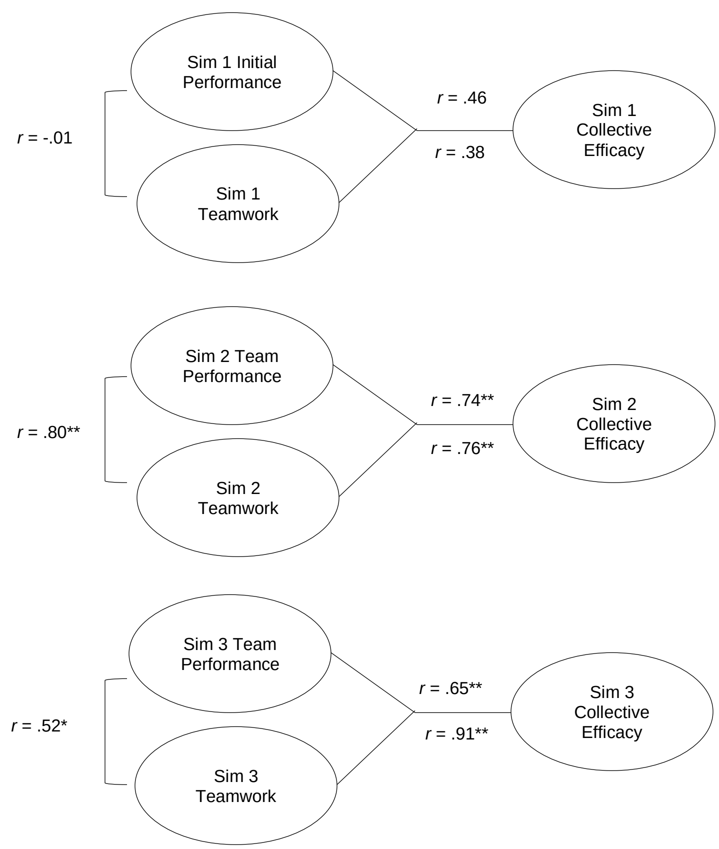


Figure 2. Hypothesis 1: Correlations between Team Performance as Measured by the Observer Rubric, Teamwork, and Collective Efficacy.

Note. n = 15. * indicates correlation is significant at the .05 level (2-tailed). ** indicates correlation is significant at the .01 level (2-tailed). Sim = Simulation.

However, all relevant Simulation 2 variables were significantly related to each other. Simulation 2 delay time team performance was negatively related to Simulation 2 collective efficacy, $r = -.54$, $p = .020$, which indicates better delay time performance is associated with greater collective efficacy. Simulation 2 team performance as measured by the observer rubric was also related to Simulation 2 collective efficacy, $r = .74$, $p = .002$. Simulation 2 teamwork was significantly correlated with Simulation 2 collective efficacy, $r = .76$, $p = .001$.

Additionally, all relevant Simulation 3 variables were significantly related to each other. Simulation 3 delay time team performance was negatively related to Simulation 3 collective efficacy, $r = -.69$, $p = .005$, which indicates better delay time performance is associated with greater collective efficacy. Simulation 3 team performance as measured by the observer rubric was also related to Simulation 3 collective efficacy, $r = .65$, $p = .009$. Also, Simulation 3 teamwork was significantly correlated with Simulation 3 collective efficacy, $r = .91$, $p < .001$.

Regressions. Hypothesis 1 was further evaluated with six multiple regressions, using the combinations of variables seen in Figures 1 and 2. Regressions were conducted using the enter method to determine if either of the predictors was able to explain variance over and above the other predictor.

Simulation 1. Neither of the Simulation 1 regression analyses was significant. Simulation 1 teamwork and Simulation 1 delay time team performance did not account for a significant proportion of variance in Simulation 1 collective efficacy, $R^2 = .17$, $F(1, 13) = 1.25$, $p = .321$. Additionally, Simulation 1 teamwork and Simulation 1 team performance as measured by the observer rubric did not account for a significant proportion of variance in Simulation 1 collective efficacy, $R^2 = .36$, $F(1, 13) = 3.44$, $p = .07$.

Simulation 2. Simulation 2 teamwork and Simulation 2 delay time team performance explained a significant proportion of variance in Simulation 2 collective efficacy, $R^2 = .63$, $F(1, 13) = 10.27$, $p = .003$. Simulation 2 teamwork was a significant predictor of Simulation 2 collective efficacy, $b = .76$, $t(13) = 4.21$, $p = .001$; however, Simulation 2 delay time team performance was not.

Simulation 2 teamwork and Simulation 2 team performance as measured by the observer rubric explained a significant proportion of variance in Simulation 2 collective efficacy, $R^2 = .63$, $F(1, 13) = 10.06$, $p = .003$. Simulation 2 teamwork was as a significant predictor of Simulation 2 collective efficacy, $b = .76$, $t(13) = 4.21$, $p = .001$; however, Simulation 2 team performance as measured by the observer rubric was not.

Simulation 3. Simulation 3 teamwork and Simulation 3 delay time team performance explained a significant proportion of variance in Simulation 3 collective efficacy, $R^2 = .84$, $F(1, 13) = 31.99$, $p < .001$. Simulation 3 teamwork was a significant predictor of Simulation 3 collective efficacy, $b = .91$, $t(13) = 4.21$, $p < .001$; however, Simulation 3 delay time team performance was not. Simulation 3 teamwork explained a significant proportion of variance in Simulation 3 collective efficacy, $R^2 = .82$, $F(1, 13) = 61.03$, $p < .001$.

Simulation 3 teamwork and Simulation 3 team performance as measured by the observer rubric explained a significant proportion of variance in Simulation 3 collective efficacy, $R^2 = .87$, $F(1, 13) = 39.27$, $p < .001$. Simulation 3 teamwork was a significant predictor of Simulation 3 collective efficacy, $b = .91$, $t(13) = 4.21$, $p < .001$; however, Simulation 3 team performance as measured by the observer rubric was not.

Hypothesis 1 was supported, as teamwork and team performance (delay time and observer rubric) were significantly related to collective efficacy. However, only the teamwork measure was a significant predictor of collective efficacy.

Testing Hypothesis 2

Hypothesis 2 states, “Collective efficacy and teamwork behaviors are positively related to subsequent team performance.” Team performance as measured by delay time was considered first, followed by team performance as measured by the observer rubric.

Correlations. To begin evaluating this hypothesis, relevant bivariate correlations were calculated, as seen in Figures 3 and 4. (All bivariate correlations can be found in Table 4.) Team performance as measured by delay time was considered first, followed by team performance as measured by the observer rubric. Only one correlation relevant to Hypothesis 2 was significant. Simulation 1 collective efficacy was significantly related to Simulation 2 team performance as measured by the observer grading rubric, $r = .59$, $p = .020$.

When investigating effect sizes, all of the delay time team performance correlations were small. However, the correlation between Simulation 1 teamwork and Simulation 2 team performance as measured by the observer rubric reached a medium effect size ($r = .47$, $p = .08$), as defined by Cohen (1988). The correlation between Simulation 2 teamwork and Simulation 3 team performance as measured by the observer rubric also reached a medium effect size ($r = .37$, $p = .18$). Additionally, the correlation between Simulation 2 collective efficacy and Simulation 3 team performance as measured by the observer rubric reached a medium effect size, *r* = .43, *p* = .10.

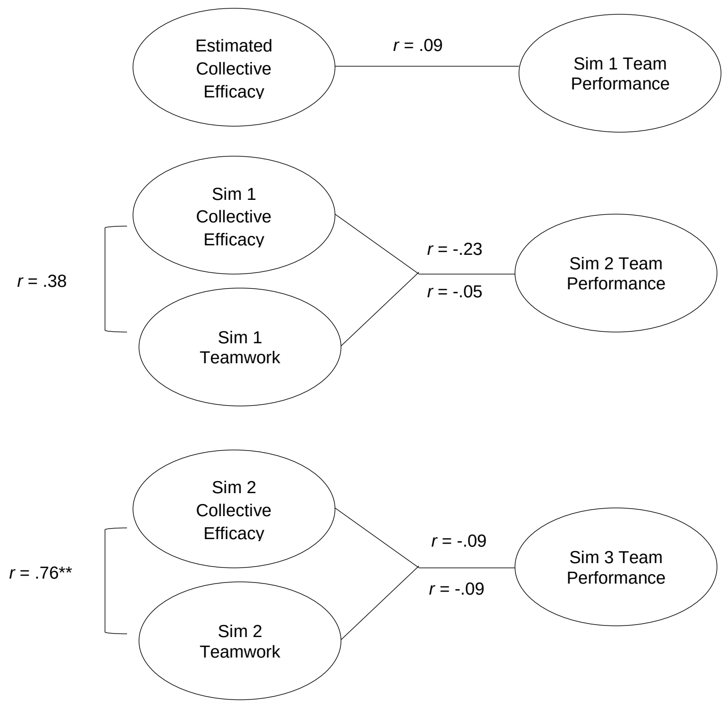


Figure 3. Hypothesis 2: Correlations between Team Performance as Measured by Delay Time, Teamwork, and Collective Efficacy.

Note. *n* = 15. * indicates correlation is significant at the .05 level (2-tailed). ** indicates correlation is significant at the .01 level (2-tailed). Sim = Simulation. Delay times that are negatively correlated indicate an increase in performance, as delay times should be as low as possible.

Regressions. Hypothesis 2 was further evaluated with six regressions, using the combinations of variables seen in Figures 3 and 4. Estimated collective efficacy was not a significant predictor of Simulation 1 team performance for neither delay time, R^2 = .01, *F*(1, 13) = .10, *p* =.75, nor the observer rubric, R^2 = .001, *F*(1, 13) = .02, *p* = .90. Simulation 1 collective efficacy and simulation 1 teamwork were not significant predictors of Simulation 2 delay time team performance, $R^2 = .23$, $F(1, 13) = .33$, $p = .72$. Additionally, Simulation 2 collective efficacy and simulation 2 teamwork were not significant predictors of Simulation 3 team performance for neither delay time, $R^2 = .01$, $F(1, 13) = .05$, $p = .95$, nor the observer rubric, $R^2 = .19$, $F(1, 13) = 1.44$, $p = .27$.

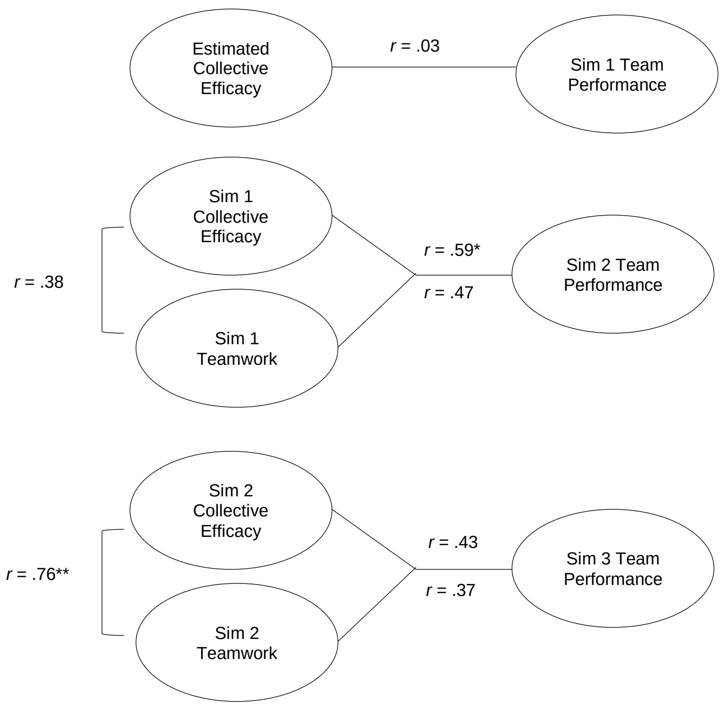


Figure 4. Hypothesis 2: Correlations between Team Performance as Measured by the Observer Rubric, Teamwork, and Collective Efficacy.

Note. $n = 15$. * indicates correlation is significant at the .05 level (2-tailed). ** indicates correlation is significant at the .01 level (2-tailed). Sim = Simulation.

However, Simulation 1 teamwork and Simulation 1 collective efficacy explained a significant proportion of variance in Simulation 2 team performance as measured by the observer rubric, $R^2 = .42$, $F(1, 13) = 4.31$, $p = .039$. Simulation 1 collective efficacy was a significant predictor of Simulation 2 team performance as measured by the observer rubric, *b* = .59, *t*(13) = 2.65, *p* = .020; however, Simulation 1 teamwork was not. Simulation 1 collective efficacy explained a significant proportion of variance in Simulation 2 team performance as measured by the observer rubric, R^2 = .35, *F*(1,13) = 7.01, *p* = .020. Hypothesis 2 was only minimally supported, as collective efficacy was shown to be related to subsequent team performance in only one of three possible instances.

Testing Research Question 1

Research Question 1 states "Is the change in team performance concomitantly related to subsequent changes in collective efficacy?" If performance improved from Simulation 1 to Simulation 2, collective efficacy should improve from Simulation 1 to Simulation 2. If performance continued to improve from Simulation 2 to Simulation 3 and collective efficacy continued to improve from Simulation 2 to Simulation 3, evidence would be provided for a positive performance spiral. Likewise, if performance declined from Simulation 1 to Simulation 2, collective efficacy should decline from Simulation 1 to Simulation 2. If performance continued to decline from Simulation 2 to Simulation 3 and collective efficacy continued to decline from Simulation 2 to Simulation 3, evidence would be provided for a negative performance spiral. Alternatively, if performance-efficacy spirals are not supported, the relationship of the changes between performance and collective efficacy should be mixed. For example, performance would improve, but collective efficacy would decline. Teams that experience mixed results such as these would not provide evidence for existence of performance spirals.

To assess this, each team was evaluated to determine whether or not there was evidence of a performance spiral. A chi-square analysis was conducted to determine whether the observed change was different enough from the expected change to provide statistical significance. The expected frequency was that 50% of teams would show performance spirals (either positive or negative), and 50% of teams would not show performance spirals. Using delay time as the team performance criterion, five cases showed performance spirals, and 10 did not. Using the observer rubric as the team performance criterion, four cases showed performance spirals, and 11 did not. The observed frequency was not significantly different from the expected frequency, as results were not significant for delay time ($\chi^2(1) = .45$) or for the observer rubric ($\chi^2(1) = .40$). As such, Research Question 1 did not find evidence that change in team performance was concomitantly related to subsequent changes in collective efficacy.

Testing Research Question 2

Research Question 2 states "Is the change in collective efficacy concomitantly related to subsequent changes in team performance?" If collective efficacy improves from the estimated level to the Simulation 1 level, performance should improve from Simulation 1 to Simulation 2. If collective efficacy continues to improve from Simulation 1 to Simulation 2 and performance continues to improve from Simulation 2 to Simulation 3, evidence would be provided for a positive performance spiral. Likewise, if collective efficacy declines from the estimated level to the Simulation 1 level, performance should decline from Simulation 1 to Simulation 2. If collective efficacy continues to decline from Simulation 1 to Simulation 2 and performance continues to decline from Simulation 2 to Simulation 3, evidence would be provided for a negative performance spiral. Alternatively, if performance-efficacy spirals are not supported, the relationship of the changes between collective efficacy and performance should be mixed. For example, collective efficacy would improve, but performance would decline. Teams that experience mixed results such as these would not provide evidence for existence of performance spirals.

To assess this, each team was evaluated to determine whether or not there was evidence of a performance spiral. A chi-square analysis was conducted to determine whether the observed change was different enough from the expected change to provide statistical significance. The expected frequency was that 50% of teams would show performance spirals (either positive or negative), and 50% of teams would not show performance spirals. Using delay time as the team performance criterion, three cases showed performance spirals, and 12 did not. Using the observer rubric as the team performance criterion, five cases showed performance spirals, and 10 did not. However, the observed frequency was not significantly different from the expected frequency, as results were not significant for delay time ($\chi^2(1) = .58$) or for the observer rubric ($\chi^2(1) = .40$). As such, Research Question 2 did not find evidence that change in collective efficacy was concomitantly related to subsequent changes in team performance.

CHAPTER IV

DISCUSSION

This study investigated the relationships among collective efficacy, teamwork, and team performance. Each of these is a team-level construct, so it was important to aggregate individual responses to the team level. Based on analyses of agreement, aggregation to the team level was justified. Hypotheses were evaluated using correlations and regressions.

Hypothesis 1

Following aggregation, Hypothesis 1 was investigated. It states, "Team performance and teamwork behaviors are positively related expectations for subsequent performance (collective efficacy)." Hypothesis 1 was supported. While team performance and teamwork were not correlated with collective efficacy during Simulation 1, the hypothesis can still be considered supported, as teams had yet to fully establish teamwork behaviors during the first simulation and team members were still learning how to do the basic tasks of the simulation. Additionally, the ICC(1) level for Simulation 1 teamwork was not significant, ICC(1) = .12, p = .06, which can help to explain why teamwork was not significantly correlated with collective efficacy. During Simulation 1, aggregated constructs were subject to error variance due to individual perceptions. The first After Action Review did not occur until all Simulation 1 measures were conducted, meaning that participants had not yet discussed their individual perceptions and established shared mental models. The After Action Review provided participants the opportunity to discuss the successes and failures of the simulation, what behaviors added to the successes and failures, and how to improve in upcoming simulation(s). Additionally, participants in Simulation 1 were often highly focused on individual tasks, and their perceptions of teamwork behaviors may have been inaccurate, leading to inconsistent ratings. This lack of shared mental models may also partially explain the nonsignificant correlations found in Simulation 1.

After identifying that relationships existed between team performance, teamwork, and collective efficacy, regressions were conducted to determine which predictors were strongest. Teamwork was shown to be a better predictor of collective efficacy than either measure of team performance. Teamwork was a significant predictor, but neither measure of team performance was for Simulation 2 or Simulation 3. Part of the reason that team performance was not not a significant predictor is that teamwork and team performance were highly correlated with each other. As such, team performance was not able to add much prediction beyond the variance explained by teamwork.

Hypothesis 2

Hypothesis 2 states, "Collective efficacy and teamwork behaviors are positively related to subsequent team performance." Hypothesis 2 was only minimally supported, as collective efficacy was shown to be related to subsequent team performance in the second simulation but not in the other two. Further, teamwork during one simulation was not shown to be positively related to subsequent team performance. However, while only one correlation was significant, the relationships between several variables reached a medium effect size as defined by Cohen (1988), and it is possible that more relationships would have been significant with greater statistical power.

Moreover, one possible reason that Hypothesis 2 was less supported than Hypothesis 1 is the timeline of measurement. All variables included in Hypothesis 1 were measured on the same day. For Hypothesis 2, predictor variables were measured three to four weeks prior to the criterion of performance, which can affect their impact. In addition, the simulations can be considered qualitatively different. Many things can change from one simulation to the next, introducing variance due to error. Because of this, teams that had high teamwork and collective efficacy during one simulation may not be able to translate these emergent states into team performance success in the following simulation. However, if each simulation had been exactly the same, it is possible that teamwork and collective efficacy would have been significantly related to team performance.

Some of the factors that differed from one simulation to another included changes in triggers, weather, and occasionally personnel. The triggers used in the simulations were not constant; each team received different triggers to minimize treatment contamination. However, the difficulty of these triggers was not necessarily equal. As such, the performance measures obtained were likely affected by a lack of consistency of the triggers. For example, if a team experienced particularly difficult triggers, their objective performance measure may show a higher delay time, even if their performance was actually better than another team's.

External to the triggers activated by the lab staff, weather was also an uncontrolled variable. Weather can significantly impact the choices made by an airline, which was also the case for the simulations used in this study. Because live weather was used, teams experienced different weather conditions. For example, weather may have caused aircraft to be rerouted on certain days, decreasing the delay time team performance. Additionally, some teams may have experienced change in personnel from one simulation to the next. While this was avoided as much as possible, there were situations where a team member had to be temporarily replaced, as he or she was not available for one of the simulations. Due to the disjunctive nature of the simulation, if only one team member was struggling, overall team performance could suffer. All of these potential factors changing from one simulation to the next made it more difficult to see changes across simulations, which are probably the significant factors in Hypothesis 2 only being minimally supported.

Research Questions 1 and 2

Research Questions 1 and 2 focused on establishing the existence of performance spirals. To assess this, each team was evaluated to determine whether or not there was evidence of a performance spiral. A chi-square analysis was conducted to determine whether the observed change was different enough from the expected change to provide statistical significance. However, the observed frequency was not significantly different from the expected frequency. Analyses regarding Research Question 1 did not find evidence that change in team performance was concomitantly related to subsequent changes in collective efficacy, and Research Question 2 did not find evidence that change in collective efficacy was concomitantly related to subsequent changes in team performance. One problem with these analyses is that assumptions of the chi-square test were violated. Chi-square tests require at least five inputs per cell (20 total), and this study only provided 15 total inputs.

Implications

This study supported the growing body of knowledge that indicates collective efficacy and self-efficacy do not all have the same antecedents. For example, this study shows the importance of teamwork in later collective efficacy, which would not be a factor in self-efficacy.

This study also supported Tasa et al.'s (2007) conclusion that teamwork is significantly related to collective efficacy. These researchers suggested that this connection indicates that organizations should support interventions that encourage teamwork behaviors, as they lead to collective efficacy (Tasa et al., 2007). As such, organizations should work to increase teamwork behaviors, rather than focusing on outcomes only (Tasa et al., 2007). One way to work toward teamwork behaviors is through training programs that seek to improve teamwork behaviors (Tasa et al., 2007). For example, Chen, Donahue, and Klimoski (2004) showed that a university course designed to develop students' teamwork KSAs significantly increased students' teamwork knowledge and skills. The current study also further supports Tasa et al.'s (2007) finding that teamwork behaviors are relevant inputs in the formation of collective efficacy beliefs.

In addition, the current study expands Tasa et al.'s (2007) study by using a different type of task. Using Steiner's (1972) task typology, Tasa et al. (2007) acknowledged that the task used in their study could be completed as a disjunctive task, meaning that "only one person needs to perform well in order for the team to succeed" (p. 24). Alternatively, the simulation used in this study was a conjunctive task; "everyone's success is determined by the effectiveness with which the least proficient member operates" (Steiner, 1972, p. 28). While each team member has unique tasks, their efforts must combine to reach common goals. If any one individual chose to not participate or performed poorly, the flight operations team would not succeed. It is also possible that the conjunctive nature of the simulation used in this study made it more difficult for collective efficacy and teamwork to have an impact on team performance because success also required individual task knowledge and skills that may not have been equivalent across teams.

Limitations

The most significant limitation of this study is the low number of teams assessed. As only 15 teams were used, statistical power was limited, making it difficult to find significance. It is possible that some conclusions drawn by this study are Type II errors. This limitation could be addressed by replicating this study with a larger sample size. Sample size was inhibited by the fact that only aerospace seniors are appropriate participants. In addition, missing data also decreased the analyzed sample.

A second limitation of this study is that collective efficacy was measured prior to teams receiving feedback in After Action Reviews, so the teams may have had higher collective efficacy than was merited, as they may have not fully understood how well they performed. That being said, each participant was able to view a live feed of the arrival and departure times of the simulation aircraft. Delay times were displayed, so participants did have access to some performance feedback. Even so, participants had no standard to compare themselves to, in order to understand their level of success or failure. However, by measuring collective efficacy prior to participating in an After Action Review, this study addressed a limitation of Tasa et al.'s (2007) study, in which specific feedback was given before collective efficacy was measured. In a work setting, collective efficacy continues to affect employees, whether teams receive continuous feedback or not. As such, the timing of the collective efficacy measure in relation to receiving feedback in this study both causes a limitation and addresses a limitation of a previous study.

Another limitation of this study is that the subjective observer performance rating measure raises concerns about unreliability and accuracy. The same raters were not necessarily involved with rating each simulation or each team, and we do not know if all raters used the same standards and frame-of-reference throughout the observations, jeopardizing inter-rater reliability (Krokos et al., 2009). Yet, inter-rater reliability was supported through the use of ICC(1), as noted in Table 2.

Additionally, as previously mentioned, the triggers used in the simulations were not constant, and the difficulty of these triggers was not necessarily equal. Further, weather was an uncontrolled variable. Weather can significantly impact the choices made by an airline, which was also the case for the simulations used in this study. Live weather was used, causing teams to experience different weather conditions. Because neither triggers nor weather were held constant across teams and across simulations, significant error variance is likely inherent in the data.

One final limitation is the potential for same-source bias in the results, as the same team members assessed teamwork and collective efficacy. However, this may not have been a serious issue, as collective efficacy and teamwork were not always significantly correlated with each other. That being said, there is also the potential for testing bias, as this is a longitudinal study, and the same measure is completed multiple times: collective efficacy is measured four times, and teamwork is measured three times. Even so, neither measure of team performance used self-report methods, meaning that they were not susceptible to same-source bias, and relationships were found between collective efficacy and both the objective (delay time) and subjective (observer rubric) measures of team performance. Teamwork was also related to both measures of team performance.

Future Research

Future research of this topic should aim to increase power. The number of teams included in this study was limited, and more predictors may have been found to be significant with more power. Future studies would also benefit from using equally difficult triggers across all simulations. The varied difficulty of simulations makes it challenging to draw any solid conclusions about team performance, as conditions were not equal across teams and across time.

Additionally, the current study seemed to indicate that the timing of the After Action Review could have been a factor in the self-report ratings of teamwork and collective efficacy perceptions. As such, a follow-up study could investigate if the time of the After Action Review significantly impacted the teamwork or collective efficacy perceptions. In the first group, collective efficacy and teamwork perceptions would be assessed prior to the After Action Review. In the second group, the After Action Review would occur prior to assessing collective efficacy and teamwork perceptions. Villado and Arthur (2013) found that compared to teams who did not complete an After Action Review, those who did had higher levels of team performance, team efficacy, openness of communication, and cohesion. However, this proposed study would go beyond Villado and Arthur's (2013) study, as both groups would participate in an After Action Review, yet at different times. This proposed study would help to provide more information about best practices regarding After Action Reviews.

REFERENCES

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein, & S. W. J. Kozlowksi (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349-381). San Francisco: Jossey-Bass.

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

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology, 83*, 234-246. doi:10.1037/0021-9010.83.2.234

Chen, G. L., & Bliese, P. D. (2002). The role of different levels of leadership in predicting self- and collective efficacy: Evidence for discontinuity. *Journal of Applied Psychology, 87*, 549-556.

Chen, G., Donahue, L. M., & Klimoski, R. J. (2004). Training undergraduates to work in organizational teams. *Academy of Management Learning and Education, 3*, 27-40.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290.

Cohen, J. (1998). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Collins, C. G., & Parker, S. K. (2010). Team capability beliefs over time: Distinguishing between team potency, team outcome efficacy, and team process efficacy. *Journal of Occupational and Organizational Psychology, 83*, 1003-1023. doi:10.1348/096317909X484271

Dorsey, D., Russell, S., Keil, C., Campbell, G., Van Buskirk, W., & Schuck, P. (2009). Measuring teams in action: Automated performance measurement and feedback in simulation-based training. In E. Salas, G. F., Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 383-408). New York: Routledge.

Gersick, C. J. G. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal, 31*, 9-41. doi:10.2307/256496

Gibson, C. B. (1999). Do they do what they believe they can? Group efficacy and group effectiveness across tasks and cultures. *Academy of Management Journal, 42,* 138-52. doi:10.2307/257089

Gibson, C. B., Randel, A. E., & Earley, P. C. (2000). Understanding group efficacy: An empirical test of multiple assessment measures. *Group and Organization Management, 25*, 67-97. doi: 10.1177/1059601100251005

Goodwin, G. F., Burke, C. S., Wildman, J. L., & Salas, E. (2009). Team effectiveness in complex organizations: An overview. In E. Salas, G. F., Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 383-408). New York: Routledge.

Gully, S. M., Incalcaterra, K. A., Joshi, A., & Beaubien, J. M. (2002). A meta-analysis of team-efficacy, potency, and performance: Interdependence and level of analysis as moderators of observed relationships. *Journal of Applied Psychology, 87*, 819-832. doi:10.1037//0021-9010.87.5.819

Guzzo, R. A., & Shea, G. P. (1992). Group performance and intergroup relations in organizations. In Dunnette, M. D., & Hough, L. M. (Eds.), *Handbook of industrial and organizational psychology* (2nd ed.). (pp. 296-313). Palo Alto, CA: Consulting Psychologists Press.

Guzzo, R. A., Yost, P. R., Campbell, R. J., & Shea, G. P. (1993). Potency in groups: Articulating a construct. *British Journal of Social Psychology, 32*, 87-106.

Hackman, J. R. (1987). The design of work teams. In Lorsch, J. W. (Ed.), *Handbook of organizational behavior* (pp. 315-342). Englewood Cliffs, NJ: Prentice-Hall.

Hauenstein, N. M. A., Brown, R. D., & Sinclair, A. L. (2010). BARS and those mysterious, missing middle anchors. *Journal of Business Psychology, 25*, 663-672. doi:10.1007/s10869-010-9180-7

James, L. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67,* 219-229. doi:10.1037/0021-9010.67.2.219

Jung, D., & Sosik, J. (2003). Group potency and collective efficacy: Examining their predictive validity, level of analysis, and effects of performance feedback on future group performance. *Group and Organization Management, 28*, 366-391. doi:10.1177/1059601102250821

Kozlowski, S. W., & Bell, B. S. (2003). Work groups and teams in organizations. In W. Borman, D. Ilgen, & R. Klimoski (Eds.) *Handbook of psychology: Volume 12 industrial and organizational psychology* (pp. 333-375). London: Wiley. Retrieved from EBSCOhost.

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest, 7,* 77-124. doi:10.1111/j.1529-1006.2006.00030.x

Krokos, K. J., Baker, D. P., Alonso, A., & Day, R. (2009). Assessing team processes in complex environments: Challenges in transitioning research to practice. In E. Salas, G. F., Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 383-408). New York: Routledge.

LePine, J. A., Hollenbeck, J. R., Ilgen, D. R., & Hedlund, J. (1997). Effects of individual differences on the performance of hierarchical decision-making teams: Much more than g. *Journal of Applied Psychology, 82*, 803-811.

Lindsley, D. H., Brass, D. J., & Thomas, J. B. (1995). Efficacy-performance spirals: A multilevel perspective. *Academy of Management Journal, 20*, 645-678. doi:10.2307/258790

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *The Academy of Management Review, 26*, 356-375. doi:10.2307/259182

Mathieu, J., Maynard, T., Rapp, T., & Gibson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management, 34*, 410-476. doi: 10.1177/0149206308316061

Mathieu, J. E., Rapp, T. L., Maynard, M. T., & Mangos, P. M. (2010). Interactive effects of team and task shared mental models as related to air traffic controllers' collective efficacy and effectiveness. *Human Performance, 23,* 22-40. doi:10.1080/08959280903400150

NASA FOCUS Lab. (2013). *Universal E-Lines Introduction* [PowerPoint slides.] Retrieved from http://mtsu.edu/nasafocuslab/Onboarding%20SPR%202013.pdf

NEXTOR. (2010). *Total delay impact study: A comprehensive assessment of the costs and impacts of flight delay in the United States*. Retrieved January 28, 2013 from http://nextor.org/pubs/TDI_Report_Final_11_03_1-.pdf

Pulakos, E. P. (2009). *Performance management: A new approach for driving business results*. Malden, MA: Wiley-Blackwell.

Quiñones, M. A. (1995). Pretraining context effects: Training assignment as feedback. *Journal of Applied Psychology, 80*, 226-238. doi: 10.1037/0021-9010.80.2.226

Sanders, E. (2013). *Using social network analysis to measure shared mental models of communication and interdependence in a simulated flight operations command center.* (Unpublished master's thesis.) Middle Tennessee State University, Murfreesboro, TN.

Shea, C. M., & Howell, J. M. (2000). Efficacy-performance spirals: An empirical test. *Journal of Management, 26*, 791-812. doi:10.1177/014920630002600409

Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology, 47*, 149-155. doi:10.1037/h0047060

Stajkovic, A. D., Lee, D., Nyberg, A. J. (2009). Collective efficacy, group potency, and group performance: Meta-analyses of their relationships, and test of a mediation model. *Journal of Applied Psychology, 94*, 814-828. doi:10.1037/a0015659

Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.

Stout, R. J., Cannon-Bowers, J. A., & Salas, E. (1996). The role of shared mental models in developing team situational awareness: Implications for training. *Training Research Journal, 2,* 85-116.

Switzer, F. S., III, & Roth, P. L. (2002). Coping with missing data. In S. G. Rogelberg (Ed.), *Handbook of research methods in industrial and organizational psychology* (pp. 311-323). Malden, MA: Blackwell.

Tannenbaum, S. I., & Cersaoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 55*, 231-245. doi:10.1177/0018720812448394

Tasa, K., Sears, G. J., & Schat, A. C. H. (2011). Personality and teamwork behavior in context: The cross-level moderating role of collective efficacy. *Journal of Organizational Behavior, 32*, 65-85. doi:1 0.1002/job.680

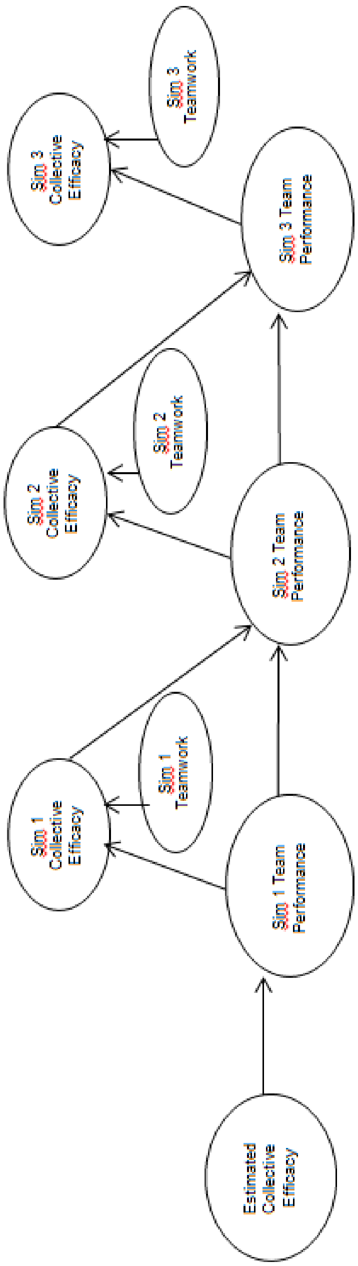
Tasa, K., Taggar, S., & Seijts, G. H. (2007). The development of collective efficacy in teams: A multilevel and longitudinal perspective. *Journal of Applied Psychology, 92*, 17-27. doi:10.1037/0021-9010.92.1.17

Villado, A. J., & Arthur, W., Jr. (2013). The comparative effect of subjective and objective after-action reviews on team performance on a complex task. *Journal of Applied Psychology.* Advance online publication. doi:10.1037/a0031510

APPENDICES

APPENDIX A

PROPOSED MODEL WITH TIMELINE



	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
Teams 1,4	Onboarding	Training	Sim 1	AAR 1	Off	Sim 2	AAR 2	Off	Sim 3	AAR 3	Off	Off
Teams 2, 5	Onboarding	Training	Off	Sim 1	AAR 1	Off	Sim 2	AAR 2	Off	Sim 3	AAR 3	Off
Teams 3, 6	Onboarding	Training	Off	Off	Sim 1	AAR 1	Off	Sim 2	AAR 2	Off	Sim 3	AAR 3

Note. Sim = Simulation.

APPENDIX B

ROLE JOB DESCRIPTIONS

FLIGHT OPS COORDINATOR

This is one of the most important jobs at flight operations. The Flight Operations Coordinator (FOC) is responsible for making sure flights depart and arrive on time and that the airline operates as smoothly and efficiently as possible. He or she also burdened with ensuring that all flight operations conform to FAA Part 121 Federal Aviation Regulations, as well as FAA approved operations specifications. FOC is the only person at flight operations who has the authority to release an aircraft for dispatch. FOC is also the final authority concerning cancellation, rerouting or delay of a flight. This position will make decisions based on information given to you by other team members within flight operations. For example, weather & forecasting personnel should review the weather for each flight and report whether the flight is legal to depart. They may also inform this position of unusual conditions that may impact on time performance. In the same way, crew scheduling will keep this position informed of pilot duty times. Maintenance personnel will provide information regarding aircraft discrepancies and other issues. This position may release a flight for dispatch only after it has been determine that all requirements have been met.

Remember, as the Flight Ops Coordinator, it is expected that this person will lead the team members and encourage them to share important information that will help make better decisions. It is also important to keep in mind that the scope of decisions that impact on-time performance and overall operational efficiency doesn’t stop when the shift ends. A schedule disruption could easily impact subsequent flights and passengers for days.

FLIGHT OPS DATA – Flight Tracking & Schedule Management

The Flight Ops Data (FOD) Specialist is responsible for updating the status of the flights at Universal E-Lines. Other team members at flight operations will need up-to-date status information so that they can better plan for and make decisions. This position is directly across from the Flight Operations Coordinator, who ultimately makes all critical decisions for the airline. When a flight has been released for dispatch, he/she will verbally inform FOD of that change in order to update the system. FOD will also be monitoring the progress of each flight on in-house radar displays. The Flight Ops Coordinator may delegate some decision-making responsibility to this position if his/her workload is high. This position will be working alongside another data specialist who will be responsible for flight planning, fuel management and some financial data. It is both data specialists' jobs to share information and update one another throughout the shift. During that time, our company will monitor 140 departures and arrivals at the sixteen cities we serve.

FLIGHT OPS DATA – Flight Planning, Fuel Management, Financial Reporting

The Flight Ops Data (FOD) Specialist is responsible for collecting and processing information about flights that are scheduled to depart. Much of this involves final flight route planning, fuel management, passenger manifests, aircraft weight & balance, takeoff and cruise performance and, at times, even some financial data. Most flights will already have their flight plan information calculated, but last minute changes in weather, passengers, and routing can change flight characteristics. Sometimes, this position may be presented with a situation where a flight cannot accept more passengers or must take reduced fuel to remain within limitations. FOD will consult with the Flight Ops Coordinator about these decisions, as they have the potential to impact flights later in the day. Finally, FOD will review all flight plans before each flight is released for dispatch. Most dispatch releases will be sent to the flight crew automatically, while others will require FOD to send them manually.

CREW SCHEDULING

The Crew Scheduling starts during the afternoon shift, so they will not have to assign crews to each flight. This person will, however, be responsible for checking each crew’s duty times to ensure that they are legally able to fly a given leg of a flight. The scheduler will also check other FAA required currencies and qualifications. The flight operations coordinator will need this information before he/she can release a flight for dispatch. Along with these normal duties, the scheduler may be asked to find a replacement crew and work with others on your team to resolve personnel issues.

MAINTENANCE COORDINATOR

The Maintenance Coordinator is responsible for managing all maintenance aspects of Universal E-Lines’ fleet. These roles include informing the flight operations coordinator of the airworthiness of each aircraft before it can be released for dispatch. This involves ensuring an aircraft will not violate inspection requirements. The maintenance coordinator is also responsible for managing minimum equipment lists for each aircraft. In some cases, the flight operations coordinator may ask if there is an alternate aircraft available for a leg of a flight. The maintenance controller may also need to help flight crews with aircraft problems as they arise.

WEATHER & FORECASTING

The Staff Meteorologist at the Weather & Forecasting station is responsible for monitoring weather conditions in the Southeastern United States. This position will be asked by the Flight Operations Coordinator whether each flight can legally depart in the given weather conditions. A flight cannot be released for dispatch until it is reported that weather conditions permit the flight to take place. This process involves reviewing airport terminal approach procedures and takeoff/landing requirements. The forecaster must keep the flight operations coordinator informed of all relevant weather information that you feel he/she may need to make good decisions. This person may also be asked to speak with pilots about weather conditions and strategize aircraft speeds, altitudes and routes that have the potential to improve efficiency. A pilot may also contact you if he/she experiences unexpected weather conditions. Finally, periodically the forecaster may be asked to provide current conditions to a flight planner at the Flight Ops Data position so he/she can better calculate aircraft performance.

APPENDIX C

LAB WORKSPACES

Figures 5 and 6. Lab layout.

Figure 7. Lab layout.

Figure 8. Flight Operations Coordinator.

Figure 9. Flight Operations Data 1 and 2.

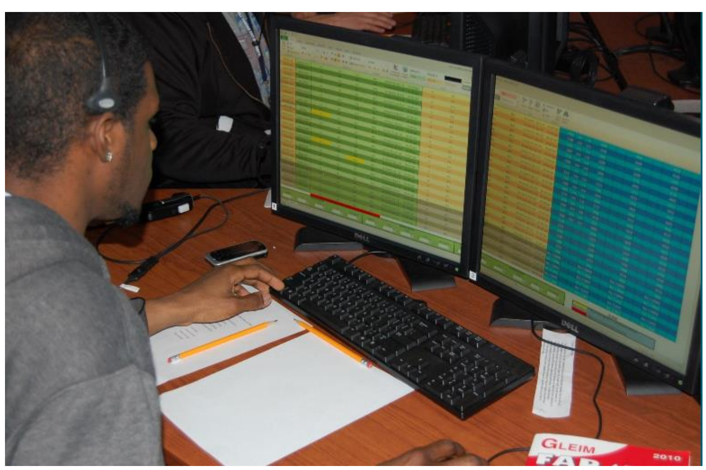


Figure 10. Crew Scheduling.

Figure 11. Maintenance Control.

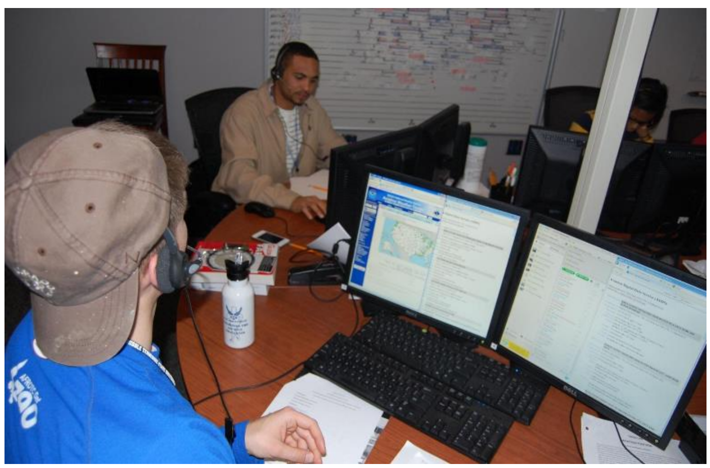


Figure 12. Weather and Forecasting.

APPENDIX D

TEAM PERFORMANCE: OBSERVER RUBRIC

	1 Trainee Level	**3 Developing Level**	**5 Experienced Level**	**7 Professional Level**	**Score**
Problem Solving Objective or Performance	Problems go unseen until a point where they become critical. Once the problem is recognized corrective actions are slow or non-existent.	Reacts to problems after they occur. Implementing corrective actions is non-assertive.	Resolves problems as they arise quickly and efficiently.	Team members work together to pre-empt problems. The team sees a problem in the making and implements corrective measures in advance.	
Problem Solving Objective or Performance	Decisions go un-made because team members are unaware that a decision is called for.	Decisions are non-assertive and tentatively made. Little or no confidence in decisions is displayed.	Decisions are made in a timely manner.	Decisions are made assertively and with confidence.	
Problem Solving Objective or Performance	Events of the scenario go undetected.	Ramifications of the events in the scenario are not completely understood. Decisions are made without understanding the full consequences of the choices that are made.	Consequences of the events of the scenario are understood, and actions are taken to reduce adverse effects.	Consequences of the events of the scenario are understood, anticipated, and mitigated to the greatest possible extent.	

	1 **Trainee Level**	3 **Developing Level**	5 **Experienced Level**	7 **Professional Level**	**Score**
Problem Solving Objective or Performance	The events of the scenario seem to overwhelm some or all team members. Increased workload events come as a surprise - team members never anticipate events to come.	Team members use downtime when the action of the scenario is reduced to “catch their breath.” There is no preparation in the lulls for the peek workload times.	Team members do not waste downtime opportunities, but no time-use strategy is in place.	Team members use downtime or lulls in the action to prepare for the busy times. Team members anticipate what is to come. Time-use strategies are used by individuals as well as the group as a whole.	

APPENDIX E

TEAMWORK MEASURE

Action Phase Teamwork Performance

Please use the following scale to describe your team that just completed the NASA lab simulation exercise.

To what extent does our team actively work to …..

1	**2**	**3**	**4**	**5**
Not at all	**Very Little**	**To Some Extent**	**To a Great Deal**	**To a Very Great Extent**

Monitoring Progress Toward Goals

_____1. Regularly monitor how well we are meeting our team goals?

_____2. Use clearly defined metrics to assess our progress?

_____3. Seek timely feedback from stakeholders (e.g., customers, top management, other organizational units) about how well we are meeting our goals?

Resource and Systems Monitoring

_____4. Monitor and manage our resources (e.g., financial, equipment, etc.)?

_____5. Monitor important aspects of our work environment (e.g., inventories, equipment and process operations, information flows)?

_____6. Monitor events and conditions outside the team that influence our operations?

Team Monitoring and Backup

_____7. Develop standards for acceptable team member performance?

_____8. Balance the workload among our team members?

_____9. Assist each other when help is needed?

Coordination

_____10. Communicate well with each other?

_____11. Smoothly integrate our work efforts?

_____12. Coordinate our activities with one another?

Transition Processes Teamwork Performance

Please use the following scale to describe your team that just completed the NASA lab simulation exercise.

To what extent does our team actively work to …..

1	**2**	**3**	**4**	**5**
Not at all	**Very Little**	**To Some Extent**	**To a Great Deal**	**To a Very Great Extent**

Mission Analysis

_____1. Identify our main tasks?

_____2. Identify the key challenges that we expect to face?

_____3. Determine the resources that we need to be successful?

Goal Specification

_____4. Set goals for the team?

_____5. Ensure that everyone on our team clearly understands our goals?

_____6. Link our goals with the strategic direction of the organization?

Strategy Formulation & Planning

_____7. Develop an overall strategy to guide our team activities?

_____8. Prepare contingency ("if-then") plans to deal with uncertain situations?

_____9. Know when to stick with a given working plan, and when to adopt a different one?

Interpersonal Teamwork Performance

Please use the following scale to describe your team that just completed the NASA lab simulation exercise.

To what extent does our team actively work to …..

1	**2**	**3**	**4**	**5**
Not at all	**Very Little**	**To Some Extent**	**To a Great Deal**	**To a Very Great Extent**

Conflict Management

_____1. Deal with personal conflicts in fair and equitable ways?

_____2. Show respect for one another?

_____3. Maintain group harmony?

Motivating and Confidence Building

_____4. Take pride in our accomplishments?

_____5. Develop confidence in our team's ability to perform well?

_____6. Encourage each other to perform our very best?

Affect Management

_____7. Share a sense of togetherness and cohesion?

_____8. Manage stress?

_____9. Keep a good emotional balance in the team?

APPENDIX F

COLLECTIVE EFFICACY MEASURE

The following questions ask about your perceptions of your team's expected performance in the NASA lab simulation exercise. Please respond to each question using the scale below (for each question, CIRCLE the number that best reflects your response). Please answer openly and honestly, there are no right or wrong answers.

ITEMS		**Strongly Disagree**	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
1.	I feel confident in my team's ability to perform the simulation.	1	2	3	4	5
2.	I think my team can eventually reach a high level of performance on the simulation.	1	2	3	4	5
3.	I am sure my team can learn how to perform this simulation effectively in a relatively short period of time.	1	2	3	4	5
4.	I don't feel that my team is as capable of performing the simulation.	1	2	3	4	5
5.	On the average, other teams are probably much more capable of performing this simulation than my team.	1	2	3	4	5
6.	My team will learn the simulation quickly, in comparison to other teams.	1	2	3	4	5
7.	I am not sure my team can ever reach a high level of performance on this simulation, no matter how much practice and training we get.	1	2	3	4	5
8.	It would take my team a long time to learn how to perform this simulation effectively.	1	2	3	4	5
9.	I am not confident that my team can perform this task effectively.	1	2	3	4	5
10.	I doubt that my team's performance will be very adequate on the simulation.	1	2	3	4	5

APPENDIX G

MTSU IRB APPROVAL LETTER

MIDDLE TENNESSEE STATE UNIVERSITY

September 9, 2013

Shanna Jinkerson, Richard G. Moffett, III
Department of Psychology
smj4k@mtmail.mtsu.edu, Richard.Moffett@mtsu.edu

Protocol Title: "Collective Efficacy in a High-Fidelity Airline Operations Center Simulation"

Protocol Number: 14-067

Dear Investigator(s),

The exemption is pursuant to 45 CFR 46.101(b) (4). This is because the research being conducted involves the use existing data that is de-identified.

You will need to submit an end-of-project report to the Compliance Office upon completion of your research. Complete research means that you have finished collecting data and you are ready to submit your thesis and/or publish your findings. Should you not finish your research within the three (3) year period, you must submit a Progress Report and request a continuation prior to the expiration date. Please allow time for review and requested revisions. Your study expires on **September 9, 2016.**

Any change to the protocol must be submitted to the IRB before implementing this change. According to MTSU Policy, a researcher is defined as anyone who works with data or has contact with participants. Anyone meeting this definition needs to be listed on the protocol and needs to provide a certificate of training to the Office of Compliance. **If you add researchers to an approved project, please forward an updated list of researchers and their certificates of training to the Office of Compliance before they begin to work on the project.** Once your research is completed, please send us a copy of the final report questionnaire to the Office of Compliance. This form can be located at www.mtsu.edu/irb on the forms page.

Also, all research materials must be retained by the PI or **faculty advisor (if the PI is a student)** for at least three (3) years after study completion. Should you have any questions or need additional information, please do not hesitate to contact me.

Sincerely,

Andrew W. Jones

Andrew W. Jones
Compliance Office
Graduate Assistant to:
Kellie Hilker
Compliance@mtsu.edu