Title of Project: Improving patient safety through leadership and team performance in simulations

Principal Investigator and Team Members: Fernandez, R (PI); Kozlowski, SWJ; Chao, GT; Rosenman, ER; Shandro, JR; Ilgen, JS; Grand, JA

Organization: University of Washington

Inclusive Dates of Project: 09/30/2011 – 09/29/2014

Federal Project Officer: Kerm Henriksen, PhD

Acknowledgment of Agency Support: This project was supported by grant number 1R18HS020295 from the Agency for Healthcare Research and Quality.

Grant Award Number: 1R18HS020295

A. ABSTRACT

Purpose
Team leadership is critical during patient resuscitations; yet, evidence-based team leadership training has not been systematically developed and evaluated.

Scope
The project was a three-arm, randomized, controlled trial of two different formats of leadership training on interdisciplinary team performance during simulated resuscitations. Forty-four teams (132 individuals) from the University of Washington and Detroit Receiving Hospital were recruited to participate in this project.

Methods
Participant team leaders were second- or third-year emergency medicine residents randomized to one of three conditions: control (routine clinical instruction), in-person training (routine clinical instruction plus a 4-hour training session), or web-based training (a 4-hour interactive web-based training). Training evaluation occurred at multiple levels. All team leaders were evaluated on leadership knowledge; leadership perception; and ability to impact teamwork, patient care, and team mental model formation during a simulated patient resuscitation.

Results:
A total of 44 team leaders were recruited and randomized to in-person (n=14), web-based training (n=14), and control (n=16) groups. All participants completed the study. The study site was found to significantly affect teamwork and thus was included as a covariate in analysis of training effect on primary and secondary outcomes. We found no effect of familiarity on teamwork or patient care. No effect of training was detected on the primary outcomes (teamwork and patient care) or on team mental model formation. However, both formats of leadership training did demonstrate a positive, significant effect on leadership knowledge. The project succeeded in developing an in-person and easily disseminated, evidence-based leadership training program with a comprehensive evaluation toolkit. As part of this toolkit, a novel team mental model measure was developed and evaluated. Team mental model scores positively correlated with both teamwork and patient care performance during simulated resuscitations. More work is needed to further understand mechanisms for improving and measuring leadership performance within a highly variable team setting.

Key Words: team leadership, teamwork, patient safety, mental models, randomized trial

B. PURPOSE

Over half a million people suffer cardiac arrest every year, with fewer than 10% of patients surviving to discharge with a good neurologic outcome. Organized, cross-disciplinary approaches to resuscitation have been recommended as a way to address the inherent complexities of resuscitative care. However, despite dissemination and implementation of recommended cardiac arrest management protocols, there remain significant rates of medical error and preventable adverse outcomes associated with critical resuscitation. These errors represent a significant threat to the delivery of safe patient care and contribute substantially to unnecessary medical costs.

Teamwork and communication errors are primary threats to patient safety. This is especially true in ad hoc medical emergency teams, where complex tasks are performed under uncertain and time-pressured conditions. Leadership has been identified as potential targets for improving team performance and could be leveraged to improve patient safety. However, healthcare team research is only beginning to recognize team leadership as a leverage point for improving patient safety. In nonmedical domains, leadership improves team performance by developing team-oriented goals, facilitating coordination and cooperation, and improving monitoring of tasks and environment. These functions take on higher importance as task complexity and interdependency increase and team familiarity and level of training become more variable, both of which are conditions common in resuscitation teams.

The **overall objective** of the proposed research was to (1) develop a novel, evidence-based leadership training program and (2) evaluate the impact of both leadership training and team member composition (unfamiliar vs. familiar teams) on emergency medical team performance using a simulation-based research platform. The **primary outcomes** of research were (a) teamwork and (b) patient care during simulated patient resuscitations. **Secondary outcomes** included leadership knowledge, team mental model formation, and perception of leadership skills.

This work was organized into three specific aims:

Aim 1. To assess the impact of healthcare team composition (unfamiliar vs. familiar members) on ad hoc medical emergency team performance

Aim 2a. To develop and implement a team leadership training program based on an Adaptive Learning System theoretical heuristic

Aim 2b. To assess the impact of team leadership training on medical emergency team performance

Aim 3. To assess the differential impact of team leadership training on unfamiliar versus familiar medical emergency teams

C. SCOPE (Background, Context, Settings, Participants, Incidence, Prevalence)

C.1. Background and Context

Current literature cites high-quality interdisciplinary team performance as critical to patient safety and error reduction. Emergency medical teams and code teams in particular depend upon effective team member interactions to coordinate, monitor, and adapt their collective skills to accomplish patient care activities. These teams are ad hoc interdisciplinary action teams that have been identified as particular threats to patient safety due to (1) their interdisciplinary nature; (2) the complex, dynamic, and time-pressured conditions under which they function; and (3) their high variability in membership. Although initial efforts aimed at improving healthcare emergency team performance are promising, gaps remain in knowledge and methodology that inhibit clinicians and organizational leaders from identifying and leveraging key team processes to improve patient safety.

An evidence-based leadership training intervention that improves healthcare team effectiveness would present healthcare organizations with a meaningful and feasible mechanism for patient safety improvement in the error-prone, high-risk setting of medical resuscitations. Leader training would provide teams characterized by high membership variability with the skills and adaptive behaviors necessary to counteract threats to effective teamwork implicit within their structure. Current literature in other domains supports the hypothesis that training healthcare team leaders in leadership skills will improve overall team performance and that improvement will be greatest in teams with low levels of team member and task stability (e.g., resuscitation teams).

C.2. Incidence

The overall incidence of in-hospital cardiac arrest in the U.S. is 2.85 per 1,000 hospital admissions, with survival to discharge rates of 18%-20%. The incidence and prevalence of errors during medical resuscitations are currently unknown. Cardiac arrest chart reviews and simulation-based studies suggest significant error rates in medication dosing, team communication, knowledge gaps, monitoring behaviors, and environmental issues; however, the magnitude and impact of these errors in actual clinical environments are unknown.
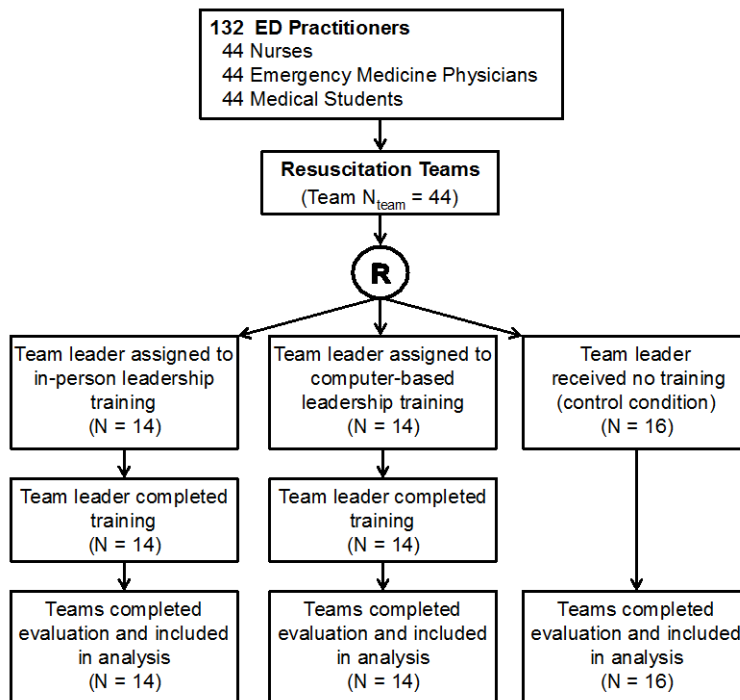
C.3. Setting and Participants

This study targets medical resuscitation team performance in an emergency department setting. We recruited 44 teams (1 MD, 1 nurse, 1 medical student) from the University of Washington (UW) Institute for Simulation and Interprofessional Studies in Seattle, WA, and the Wayne State University (WSU) Kado Family Clinical Skills Center in Detroit, MI. Team leaders were randomized to one of three conditions (control, web-based training, and in-person training). Training evaluation occurred at multiple levels. All team leaders were evaluated on leadership knowledge; leadership perception; and ability to impact teamwork, patient care, and team mental model formation during a simulated patient resuscitation.

D. METHODS

D.1. Study Design
This was a multicenter, randomized, controlled trial with three intervention arms conducted at two sites.

Figure 1. Study Design


```
                    ┌─────────────────────────────────────┐
                    │ 132  ED Practitioners                │
                    │   44 Nurses                          │
                    │   44 Emergency Medicine Physicians   │
                    │   44 Medical Students                │
                    └─────────────────────────────────────┘
                                    │
                    ┌─────────────────────────────┐
                    │  Resuscitation Teams         │
                    │  (Team N_team = 44)          │
                    └─────────────────────────────┘
                                    │
                                  ( R )
```

Team leader assigned to in-person leadership training (N = 14)	Team leader assigned to computer-based leadership training (N = 14)	Team leader received no training (control condition) (N = 16)
Team leader completed training (N = 14)	Team leader completed training (N = 14)	
Teams completed evaluation and included in analysis (N = 14)	Teams completed evaluation and included in analysis (N = 14)	Teams completed evaluation and included in analysis (N = 16)

Control Condition: Participants randomized to the **control condition** received no training above the routine clinical instruction provided all emergency medicine residents.
In-person Training Condition: Participants randomized to the **in-person training arm** of the study received a 4-hour, instructor-led training session.
Web-based Training Condition: Participants randomized to the **web-based training arm** of the study received a 4-hour, web-based, interactive training session.

D.2. Participants
To ensure that training and assessment tools were generalizable and accessible by multiple institutions, we extended our enrollment to two sites.
Physician Participants (primary study participants): Team leader participants were emergency medicine residents in their second or third year of residency recruited from the Division of Emergency Medicine at the University of Washington, Seattle, WA, and the Department of Emergency Medicine at Wayne State University, Detroit, MI. Residents were not eligible if they were not in good academic standing within their respective residency programs, were unavailable for intervention or assessment, had not participated in a minimum of four high-fidelity simulation sessions, or did not have up-to-date Advanced Cardiovascular Life Support (ACLS) certification. In total, 44 residents were approached for the study; all were eligible and included.

Nurse Team Member Participants: Nurses participating at the University of Washington were recruited from the emergency departments at Harborview Medical Center and the University of Washington Medical Center. Nurses participating at Wayne State University were recruited from the emergency department at Detroit Receiving Hospital. Nurses were ACLS certified and employed as emergency department nurses. Nurses were excluded if they were unavailable for an assessment session. The nurses served as part of the interdisciplinary teams involved in the simulation assessment. They were not randomized to receive leadership training.

Medical Student Team Member Participants: Medical students were recruited from the University of Washington and Wayne State University. Students were eligible if they were available during assessment days and had good academic standing. Students were excluded if they were unavailable for an assessment session. The medical students served as part of the interdisciplinary teams involved in the simulation assessment. They were not randomized to receive leadership training.

D.3.   Intervention (Team Leader)

Development of Leadership Training/Aim 2a

Needs Assessment: Needs assessments are critical first steps to training design, yet a recent review of the healthcare team training literature found that only 20% of studies reported some form of a needs analysis. As part of the needs assessment, we conducted systematic reviews of leadership training and assessment in healthcare teams (*H. Publications*). These reviews helped define critical competencies and training objectives for team leaders and highlighted the importance of team leader adaptability (Team Leader Competencies listed in *G. Products*). The results from our review, combined with important conceptual and empiric work in team science, supported the adoption of the Adaptive Learning System described by Kozlowski, et al. We chose the Adaptive Learning System framework because it explicitly targets adaptability and self-regulation, two critical functions for team leaders who must constantly adapt to rapidly changing clinical, environmental, and personnel characteristics.

Training Content: Training content was developed from three main sources: (1) focus group interviews conducted with nurse and physician experts, (2) critical incident reviews of adverse events, and (3) systematic and thematic reviews of the healthcare team leadership and leadership science literature. Video-recorded performances of simulated resuscitations provided the video content for the demonstration-based learning.

Training Objectives: The overall goal of the leadership training was to develop effective team leadership skills in individuals leading patient resuscitations.

Objective 1: Learners will be able to define the role of team leaders in influencing teamwork during patient resuscitations.

Objective 2: Learners will be able to identify examples of leadership processes during patient resuscitations.

Objective 3: Learners will be able to evaluate the impact of effective and ineffective leadership on teamwork and patient care.

Objective 4: Learners will be able to demonstrate effective leadership skills.

Description of Training Intervention: The training content was developed in two formats: (1) in-person, 4-hour curriculum, and (2) web-based narrated, interactive 4-hour curriculum. The formats differed in that the in-person training offered the opportunity for participants to discuss their responses and answers to questions as a group, whereas the web-based training had participants enter their responses to open-ended questions online only. Both trainings contained situation judgment test questions with formative feedback throughout the training (Situation judgment test listed in *G. Products*).

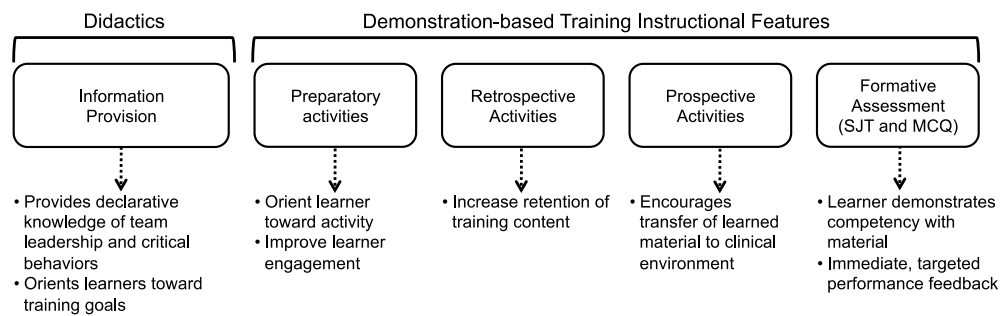|	
Training Design:	The entire training package exists as a single format that could be delivered in person or remotely (web based). Both formats have data collection capabilities to monitor learner performance on test question items and reflection responses (Both training formats listed in *G. Products*).
Didactics:	The didactic component of the training session serves to provide orientation toward critical teamwork processes and the important role team leaders play in supporting team effectiveness. This component of training is specifically geared toward meeting Objectives 1 and 2 and serves to create the foundation of knowledge needed to accomplish Objective 3.
Demonstration-based training module:	The demonstration-based training utilized actual videos of resuscitation team performance that contained examples of effective and ineffective leadership performance. Videos were transcribed and edited to provide short yet focused leadership examples that specifically tied to the leadership behaviors targeted by the training. Learners had the opportunity to (a) observe and discuss effective and ineffective examples of leadership through guided instruction, (b) reflect on their own practice and describe the impact of effective and ineffective leadership on patient care, and (c) independently identify effective and ineffective leadership skills with formative feedback. Demonstration-based training in other domains has been shown to improve transfer of behaviors to the work environment and is focused on meeting Objectives 3 and 4.
Prompted reflection:	Learners were asked to reflect on examples of leadership skills and behaviors they encountered in their own practice as well as those identified in the video training.
Situation judgment test:	Situation judgment test questions were developed to allow learners to receive real-time guided feedback throughout the training. This method was chosen as it was compatible with both in-person and web-based training.

(All training-related components are listed in *G. Products.*)

Figure 2. Team Leadership Curriculum

Didactics	Demonstration-based Training Instructional Features			
Information Provision	Preparatory activities	Retrospective Activities	Prospective Activities	Formative Assessment (SJT and MCQ)
• Provides declarative knowledge of team leadership and critical behaviors • Orients learners toward training goals	• Orient learner toward activity • Improve learner engagement	• Increase retention of training content	• Encourages transfer of learned material to clinical environment	• Learner demonstrates competency with material • Immediate, targeted performance feedback

D.4.   Data Collection

Participants were assessed 2 weeks after the training intervention.

Team leadership knowledge measure: Just prior to completing the simulation assessment, all team leader participants (control and intervention) completed a validated knowledge assessment consisting of 17 declarative-knowledge multiple-choice questions and six situation judgment test questions. A long form of the assessment with more situation judgment test questions was administered to team leaders randomized to one of the intervention groups (Team Leadership Knowledge Assessment [short form and long form] listed in *G. Products*).

Simulation: All participants were assessed on SimTEAM-L (Simulated Team Experience and Assessment Methodology for Leaders), a high-fidelity human patient simulation modified to allow participation by interdisciplinary teams directed by a physician team leader. To enable a realistic assessment environment, participants performed the simulations as members of interdisciplinary teams (a physician [team leader participants], an emergency department nurse, and a medical student). The key components of the measurement infrastructure included (1) a standardized event-based scenario that combines a "confederate" actor and functional clinical equipment (crash carts, monitors, etc.) with a human patient simulator to provide appropriate environmental and behavioral cues, (2) behavioral teamwork measures (data capture, metrics, and observational protocols to assess team process behaviors), and (3) patient care measures (data capture, metrics, and observational protocols to assess patient care behaviors). Because the simulation and associated measures were adapted from a prior study, the investigators collected additional evidence of validity to support the measurement system (Resuscitation scenario listed in *G. Products*).

Teamwork measures: The investigators adapted previously developed teamwork behavioral measures that specifically related to the simulated scenario. These measures are based on an emergency medicine teamwork taxonomy and are supported by considerable evidence of validity and reliability. The teamwork measures required adjustments, as the nature of the teams involved in this study differed slightly from our original validation study. As a result, content validity, inter-rater reliability, and response process validity evidence were collected prior to implementing the revised measures.

The final measurement tool contained 120 items (Teamwork measure listed in *G. Products*).

Patient care measures: As with the teamwork measures, we adapted previously developed patient care behavioral measures that specifically relate to the simulated scenario and are supported by evidence of validity and reliability. The patient care measures also required small modifications that required collection of additional content validity, inter-rater reliability, and response process validity evidence. The final measurement tool contained 84 items (Patient care measure listed in *G. Products*).

Leadership perception measure: All participants (team leader, nurse, medical student) completed perception measures of leadership performance. For team leaders, this was a self-perception measure. For team members, this was a peer evaluation. Validity evidence was collected to determine the correlation between team member and team leader (self-evaluation) perceptions as well as team-level outcomes (teamwork and team performance) (Leadership perception measure listed in *G. Products*).

Team mental model measure: All participants (team leader, nurse, medical student) completed a post-simulation measure to assess the degree to which the teams had developed a team mental model, or a shared understanding of the team's task. Validity evidence was collected to determine the correlation between the team mental model measure and team-level outcomes (teamwork and team performance) (Team mental model measure listed in *G. Products*).

Familiarity measure: To allow the investigators to address Aim 1, all members of the team (team leader participant, nurse, medical student) completed a familiarity measure to allow calculation of the overall level of team member familiarity within the team.

Experience: All members of the team completed an experience questionnaire specific to their team role (team leader, nurse, medical student). Items included the resident months spent in emergency department or intensive care unit, nurse experience (years) in emergency or critical care nursing, medical student year, simulation experience (number of sessions), and the number of resuscitations led.

Demographics: All members of the team completed demographic questionnaire (gender, race, ethnicity).

D.5.   Data Collection
Training-related data
Both training interventions (web based and in person) were designed to facilitate wide distribution and wide training-related data collection from a large number of sites. As it is currently structured, all participant data, including (1) reflection question data (open ended, free response), (2) formative situation judgment test, and (3) post-training evaluations (assessment situation judgment test and declarative knowledge) are collected and stored in a database that can be analyzed. This allows for both quantitative (situation judgment test, declarative knowledge tests) and qualitative (reflection data) analyses on a large number of trainees from different sites.

Assessment-related data

Demographic, experience, and familiarity questionnaires were completed using Qualtrics Software (Provo, UT). Team leader participants also completed trainee perception questionnaires and knowledge exams using Qualtrics. All team simulations were video recorded and later coded for teamwork and patient care using Noldus Observer XT (Leesburg, VA) software. Teamwork and patient care measures were coded independently by two different sets of raters to avoid introducing bias. Prior to initiating coding, all raters received approximately 10 hours of hands-on training using either teamwork or patient care measures as appropriate (Rater training program listed in *G. Product*). Following instruction on the content and nature of the rating scale, raters scored scenario performances that represented a range of performance quality (poor, average, and outstanding), and their scores were compared with those of the investigators (gold standard) to ensure accurate behavior identification and discrimination of performance levels. Both sets of coders trained until they reached an inter-rater reliability criterion of greater than 90% agreement (Cohen's Kappa>0.8).

D.6.   Data Analysis

For teamwork measures, we calculated inter-rater reliability using percent agreement across the three behavioral coders. We calculated percent agreement across three coders for each process behavior and then averaged these scores across behaviors such that each video received a score. We then averaged the scores across videos to come up with an overall percent agreement score. We verified this score by using both SPSS and formulas created in Excel. For patient care measures, a pair of raters coded each recorded simulation. We calculated Cohen's Kappa for each recording and averaged the values to give an overall measure of reliability.

First, to examine the primary aims of the project, which focused on the effectiveness of team leader training, we conducted an overall multivariate analysis of covariance (MANCOVA) that examined the effect of the experimental conditions (in-person training vs. web-based training vs. control) on the teamwork and patient care outcome measures, controlling for covariates of team familiarity, experience, leader gender, and work site (UW Medicine vs. Detroit Medical Center). Follow-up univariate analyses that examined the training and covariate effects for each outcome separately were also conducted. Second, in an effort to more fully decompose the primary findings, several additional analyses were conducted. We used a multivariate analysis of covariance to assess potential differences in the effectiveness of the two forms of training (in-person training vs. web-based training vs. control) on a variety of secondary outcome variables, including leadership knowledge scores, team mental model scores, and leadership effectiveness perceptions. Finally, to investigate relationships among team perceptual outcomes, we computed correlations among team mental models scores, team perceptions of leader effectiveness, and the individual team leader self-perception of their leadership skills.

E. RESULTS

E.1. Recruitment

Participants for this study included emergency medicine senior residents, emergency department nurses, and fourth-year medical students. Participants were recruited from two urban emergency departments: (1) UW Medicine - Harborview Medical Center and (2) Detroit Medical Center - Detroit Receiving Hospital. Recruitment was expanded to two centers to increase the number of participants to maximize the applicability of the training and proposed simulation-based outcome measures across multiple physician specialties and nursing practice sites.

Tables 1a and 1b present the demographic characteristics and self-reported experience of team leader participants (Table 1a) and across all teams (Table 1b). All recruited participants completed the study.

Table 1a: Team Leader Characteristics (n=44)

	In-person Leadership Training (n=14)	Web-based Leadership Training (n=14)	No Training (Control) (n=16)
Team Leader (participant) Characteristics			
Male, n (%)	5 (35.7%)	10 (71.4%)	10 (62.5%)
Age, mean (SD), y	29.9 (2.7)	29.0 (1.8)	30.8 (3.8)
Race, n (%)			
White	10 (71.4%)	10 (71.4%)	14 (87.5%)
Asian	3 (21.4%)	2 (14.3%)	2 (12.5%)
Black or African American	0 (0%)	1 (7.1%)	0 (0%)
Native Hawaiian or Other Pacific Islander	0 (0%)	1 (7.1%)	0 (0%)
Not reported	1 (7.1%)	0 (0%)	0 (0%)
Ethnicity, n (%)			
Not Hispanic or Latino	14 (100%)	14 (100%)	16 (100%)
Level of Training, n (%)			
PGY 2	7 (50.0%)	7 (50.0%)	11 (68.8%)
PGY 3	5 (35.7%)	7 (50.0%)	4 (25.0%)
PGY 4 or higher	2 (14.3%)	0 (0.0%)	1 (6.3%
Number of resuscitations led, mean (SD)	2.29 (1.7)	1.93 (1.9)	1.38 (.7)

Table 1b: Team-level Characteristics (n=132)

	In-person Leadership Training (n=42)	Web-based Leadership Training (n=42)	No Training (Control) (n=48)
Team-level Characteristics			
Male, n (%)	13 (31.0%)	18 (42.9%)	18 (37.5%)
Age, mean (SD), y	34.12 (9.7)	30.52 (5.8)	33.21 (9.5)
Race			
White	32 (76.2%)	27 (64.3%)	33 (68.8%)
Asian	6 (14.3%)	8 (19.0%)	8 (16.7%)
Black or African American	2 (4.8%)	6 (14.3%)	3 (6.3%)
Native Hawaiian or Other Pacific Islander	1 (2.4%)	1 (2.4%)	1 (2.1%)
American Indian or Alaskan Native	0 (0%)	0 (0%)	1 (2.1%)
Not reported	1 (2.4%)	0 (0%)	2 (4.2%)
Ethnicity, n (%)			
Not Hispanic or Latino	40 (95.2%)	41 (97.6%)	45 (93.8%)
Not reported	1 (2.4%)	0 (0%)	1 (2.1%)
Experience Composite *	1.01 (4.23) Min = -3.34 Max = 12.79	-.57 (2.53) Min = -3.52 Max = 6.60	-.38 (2.94) Min = -4.29 Max = 5.03

*calculated at the team level, therefore N=14 (in-person leadership training), N=14 (web-based training, N=16 (control). The experience composite for each team was calculated by summing the products of each of the following variables by its component score: the resident months spent in emergency department or ICU, nurse experience (years) in emergency or critical care nursing, medical student year, simulation experience (number of sessions), and the number of resuscitations led. Component scores were determined from the 7 – component solution from a principal components analysis using varimax rotation.

E.2.   Characteristics of Novel and Adapted Measures

Table 2 presents means, standard deviations, and correlations among all measures used in this study.

Teamwork and patient care measures
As noted in the *Methods*, these measures were adapted from previously developed measures that we demonstrated to have significant evidence of validity. In addition, it was observed that three of the 49 patient care items and 20 of the 120 teamwork items showed zero variance; thus, they were removed from additional analyses. Consequently, the primary outcomes of patient care and teamwork were based on 46 and 100 items, respectively.

Because these measures were altered for the purpose of this study, we established that both teamwork and patient care outcomes correlated (r=0.336, p=.026, n=44), as predicted by team effectiveness models. We also re-established that we could achieve high levels of inter-rater reliability for both teamwork (rater agreement=90.6%) and patient care (Cohen's Kappa=0.86).

Table 2: Means, standard deviations, and correlations of study measures

	mean	SD	1	2	3	4	5	6	7	8	9
1. Patient care	.464	.084	1								
2. Teamwork	.409	.052	.336 *	1							
3. Team familiarity	2.357	.836	.060	-.054	1						
4. Experience	0	3.294	-.106	.002	-.060	1					
5. Leader gender[a]	1.43	.501	.180	.165	-.138	.205	1				
6. Work site[b]	.636	.487	.093	.263†	.171	-.197	-.009	1			
7. Team leadership knowledge	9.405	2.803	-.197	-.180	.074	.016	-.167	.067	1		
8. Team mental model	73.85	14.797	.250†	.063	.042	.155	-.071	.356 *	.110	1	
9. Leader effectiveness – team member ratings	3.803	.399	-.105	.124	.137	.233	-.140	-.005	.045	.103	1
10. Leader effectiveness –leader self-ratings	3.263	.504	.143	.127	.129	-.051	-.061	.281†	.084	.165	.244

[a]Gender: 1=male, 2=female
[b]Work site: 0=WSU (Detroit), 1=UW (Seattle)
n's vary from 42 to 44 teams
*$p<.05$, †$p\leq.10$

Team leadership knowledge measure
As described in the data collection section (D.4.), 17 items were used to assess team leadership knowledge. These items described specific behaviors that tapped six team processes. Reliability of this measure was estimated with Cronbach's alpha and was found to be marginally acceptable (alpha=.67). This measure was not significantly correlated with other measures in the study.

Team mental model measure
Team mental models are thought to facilitate teamwork and team performance in highly dynamic environments. Team leaders support the development of mental models by establishing and communicating team plans and priorities. We evaluated relationships between team mental model scores and all other measures. Team mental models significantly correlated with work site (r=.356, p=.018, n=44) and marginally correlated with patient care (r=.250, p=.101, n=44).

Leader effectiveness measures
Team members rated their leaders on how effective the leader was on six team processes as well as how comfortable team members felt in speaking up and how well the leader communicated with the team. Similarly, leaders rated themselves on the same six team processes as well as items related to communications. Reliabilities for these scales were estimated by Cronbach's alpha and were .844 and .869 for team member ratings and leader self-ratings, respectively. Neither scale was found to be significantly correlated with other measures in this study. However, leader self-rating was marginally correlated with work site (r=.281, p=.064, n=44).

E.3. Analysis of Primary Outcomes: Patient Care and Teamwork

Prior to any data analyses, it was observed that the leader's gender was not balanced across training conditions; thus, it was used as a covariate in the analysis of primary outcomes to control for any gender effects.

To address the primary aims of the project, an overall multivariate analysis of covariance (MANCOVA) was conducted that examined the effects of leadership training on teamwork and patient care outcome measures, controlling for covariates of team familiarity, experience, leader gender, and work site. Aim 1 for this project focused on the impact of team familiarity on team performance. Team familiarity, which was an aggregate (i.e., mean) of the familiarity ratings among individuals on the team, did not have a significant effect on primary outcomes (F(2,34)=.216, p=.807, Wilks' Lambda=.987). It should be noted that most ratings on team familiarity were low, indicating that many team members were unfamiliar with each other (mode=1.00, mean=2.00, SD=.895, n=129). The distribution and variance of this scale restricted the opportunity to detect significant relationships between team familiarity and team outcomes. These results precluded any data analyses comparing familiar and unfamiliar teams (Aim 3).

Aim 2a pertained to the development and implementation of the leadership training program; thus, it did not involve any data analysis. Aim 2b focused on evaluating the

training program effects on team performance. Results from the MANCOVA did not support a significant effect for leadership training on the primary outcomes ($F(4,68)=1.025$, $p=.401$, Wilks' Lambda=.889). Thus, the leadership training program did not significantly improve teamwork or patient care outcomes.

Although a team's experience may have influenced primary team outcomes, the experience covariate was not significant ($F(2,34)=.885$, $p=.422$, Wilks' Lambda=.951). Similarly, leader gender did not significantly impact primary outcomes ($F(2,34)=2.360$, $p=.110$, Wilks' Lambda=.878). However, work site was found to be a significant covariate ($F(2,34)=3.911$, $p=.030$, Wilks' Lambda=.813). Follow-up analyses showed a significant difference between the Seattle and Detroit work sites for teamwork ($F(1,35)=7.909$, $p=.008$); however, this difference was not significant when work site was examined independent of other covariates. No significant work site difference was found for patient care ($F(1,35)=.172$, $p=.681$).

E.2. Analysis of Secondary Outcomes: Leader Knowledge, Leader Effectiveness, and Team Mental Model

Similar to the MANCOVA examining training effects on the primary outcomes, a MANCOVA was conducted on four secondary outcomes. We assessed team leadership knowledge scores for all team leaders just prior to the simulation assessment. Following the simulation assessment, we assessed leader self-ratings on leadership, team member ratings of the leader's effectiveness, and the team's mental model (or shared awareness of team objectives). Because work site was found to be a significant covariate in the first MANCOVA, it was specified as a covariate in this analysis as well. Results showed a significant effect for training condition for this set of outcomes ($F(8,70)=3.288$, $p=.003$, Wilks' Lambda=.528). Follow-up tests showed significant differences across training conditions for the leader knowledge test ($F(2,38)=14.919$, $p=.000$), but no significant differences were found for the other three secondary outcomes. For leader knowledge, Tukey HSD post-hoc tests showed that both the in-person and web-based training conditions were significantly higher than the control condition at the .01 level of significance. There was no significant difference between the two training conditions for the leader knowledge outcome. Work site was not found to be a significant covariate for the set of secondary outcomes ($F(4,35)=1.483$, $p=.228$, Wilks' Lambda=.855). In summary, leaders in both training conditions were significantly higher than their control group counterparts in a leadership knowledge test; however, no other training effects were found in measures of the leaders' effectiveness or on a team mental model.

F. Discussion

Team leaders are critical to effective resuscitation team performance, and failures in leadership are responsible for a significant percentage of adverse events. Despite this fact, team leadership has not been widely leveraged to improve patient safety during high-risk patient resuscitations. Our study is the first multicenter, randomized, controlled trial to systematically design and evaluate a team leadership intervention that is easily disseminated.

Although we did not detect a measurable effect of our training on team behavior, we did show an improvement in applied knowledge as a result of the training.

Prior to this study, a comprehensive description and comparison of healthcare leadership training and assessments in healthcare did not exist. We executed two large systematic reviews to ensure that our training is both comprehensive and evidence based. This novel work highlighted the lack of a commonly accepted team leadership taxonomy and conceptual framework. We addressed this issue and adapted a team leadership taxonomy from the team science literature to healthcare teams. This work was done in conjunction with the research team as well as external team and clinical scientists to ensure that the product is both rigorous and generalizable.

Healthcare institutions struggle to provide meaningful, evidence-based training to large numbers of providers. Through this project, we developed the products necessary to support delivery of such training across a large number of trainees without the need for significant resources. We built a training toolkit that can be delivered both in-person and via computer, allowing a wide distribution of training regardless of resource availability and training platform. The training uses multiple training strategies to target different learning styles and training approaches. Additionally, training can be delivered asynchronously and has built-in feedback mechanisms and assessments, minimizing the human capital needed to implement the program. The training is packaged as a toolkit that includes didactic material, demonstration-based training with video vignettes, and interactive situation judgment tests with automated feedback.

Along with the training, we developed a comprehensive assessment package that targets multiple different levels (declarative knowledge, applied knowledge, team behavior, and team cognition [mental model]). The assessment module contains declarative and applied knowledge questions that are independent of the simulation and capable of assessing individual knowledge. We collected evidence of validity for all measures. For the simulation, we created a simulation guide with a standardized patient role, mannequin cues, a debriefing guide, a teamwork checklist, a patient care checklist, and a coder training guide. Our results demonstrated a correlation between teamwork and team performance, providing evidence of validity supporting our team measurement platform. Finally, our team mental model measure is novel and correlates to teamwork as predicted by team effectiveness models. We therefore deliver a novel, multilevel assessment that can be replicated in part or in whole at multiple sites.

Our study represents the first multicenter study to systematically evaluate the impact of leadership training. We found a significant effect of training site on team process, which was unexpected. Although this difference was only detected when other covariates were considered, it did highlight the importance of multicenter trials for team-based interventions. In our study population, one of the sites is a TeamSTEPPS training site, whereas the second site does not have a formally instituted team training program. The baseline level of team training skills could skew results in a study such as ours. Expanding to multiple sites presents a challenge; however, it also provides insight into

how institutional and regional factors can influence outcomes even in a tightly controlled study such as ours.

We were not able to detect an impact of the training on teamwork or team performance during simulations. Several factors may have contributed to this finding, including the differential effect of site on our primary outcomes mentioned above. First, we purposely designed the training to be easily disseminated; however, learners were not able to practice leadership behaviors in a complex clinical environment such as the one simulated in the assessment. It is possible that performing a simulation at the time of the training and again at the time of assessment would have reinforced learned behaviors. We did demonstrate a positive relationship between training and team leadership knowledge, suggesting that our training provided a strong foundational platform upon which more experiential learning can be built. We were also limited to a single simulation for the behavioral assessment. It is possible that a broader assessment platform would have uncovered differences not detected in our sample. Finally, we used actual interprofessional teams during the simulations but were unable to control for variability in team expertise and performance. For example, the University of Washington has a rigorous TeamSTEPPS training program. It is possible that this baseline training would result in teams with stronger baseline teamwork skills and thus mask the impact of our leadership training. **Our follow-up study addresses all of these limitations by (1) employing rigorous, behaviorally based (simulation) training, (2) assessing performance over a number of different clinical events, (3) enrolling a larger number of participants, and (3) accounting for team member skill.**

The work from this project is being disseminated in a number of different venues. Thus far, we have seven peer-reviewed publications in journals such as *Critical Care Medicine*, *Academic Medicine*, and *BMJ Quality and Safety*, with additional publications planned or currently in review. Our work has also been disseminated at several national conferences, including the annual meetings of the *American Heart Association* and the *International Symposium on Human Factors and Ergonomics in Health Care*. Our ongoing follow-up study addresses the limitations in this study, further advances the assessment component of work by measuring leader performance during actual patient events, and will extend knowledge in multilevel analysis in complex healthcare teams.

G. **Products**

Team leadership competencies

In-person team leadership training materials
- Didactic materials
- Demonstration-based training with positive and negative examples
- Prompted reflection (video vignettes with self-reflection opportunity)
- Situation judgment testing with embedded, targeted feedback

Web-based team leadership training materials (automated, narrated version of in-person training with same content)

Team leadership knowledge assessment (long form and short form)
- Declarative knowledge
- Situation judgment test

Simulation-based assessment (SimTEAM-L)
- Scenario
- Confederate script
- Rater training program
- Validated teamwork measure
- Validated patient care measure
- Post-simulation team leadership perception measure
- Team mental model measure

H. Publications

1. Grand JA, Pearce M, Rench TA, et al. Going DEEP: Guidelines for building simulation-based team assessments. *BMJ Qual Saf.* 2013; 22(5):436-448. PubMed PMID: 23355693.

2. Fernandez R, Pearce M, Grand JA, et al. Evaluation of a computer-based educational intervention to improve medical teamwork and performance during simulated patient resuscitations. *Crit Care Med.* 2013;41(11):2551-2562. PubMed PMID: 23949473.

3. Fialkow M, Adams CA, Carranza, et al. An *in situ* standardized patient-based simulation to train postpartum hemorrhage and team skills on a labor & delivery unit. *Simul Healthc.* 2014;9(1):65-71. PubMed PMID: 24401919.

4. Rosenman E, Shandro JR, Ilgen, JS, et al. Leadership training for healthcare practitioners: A systematic review. *Acad Med,* 2014; 89(9). PubMed PMID: 25006708

5. Rosenman ED, Ilgen JS, Shandro JA, et al. Team leadership assessment tools in health care action teams: A systematic review. *Acad Med.* 2015;90(10):1408-1422. PubMed PMID: 26200585

6. Fernandez R, Grand JA. Leveraging social science-healthcare collaborations to improve teamwork and patient safety. *Curr. Probl Pediatr Adolesc. Health Care.* 2015;45(12):370-377. PMID: 26573242

7. Rosenman ED, Branzetti JB, Fernandez R. Assessing Team Leadership in Emergency Medicine: The Milestones and Beyond. *J Grad Med Educ.* (in press).

8. Fernandez R, Shah S, Rosenman ED, et al. Developing team cognition: A role for simulation. *Simul Healthc.* (in review)

9. Gittinger M, Brolliar S, Grand JA, et al. Evaluation of the effect of an automated chest compression device on team communication and patient management during a simulated cardiac arrest: A pilot study. *Simul Healthc.* (in review).

Book Chapters

1. Kozlowski SWJ, Chao GT, Chang C, Fernandez R. Team dynamics: Using "big data" to advance the science of team effectiveness. In Tonidandel, King, & Cortina (Eds.), *Big Data at Work: The Data Science Revolution and Organizational Psychology.* Mahwah, NJ: Lawrence Erlbaum; 2016: p. 272-309.

2. **Fernandez R**, Adedipe AA, Rosenman ED, Compton S, Kozlowski SWJ. Simulation-based measurement and program evaluation: Demonstrating effectiveness. In Strother, C, Okuda, H, McLaughlin S (Eds.), *Comprehensive Health Simulation Series: Emergency Medicine Edition.* Springer (in press).

Presentations

1. A randomized comparison study to evaluate the effectiveness of a computer-based teamwork training intervention on medical teamwork and patient care performance. Annual Meeting, Society for Academic Emergency Medicine, Atlanta, GA. 2013.

2. The role of simulation and team training in resuscitation. American Heart Association Annual Meeting. Dallas, TX. 2013.

3. Simulation-based healthcare team leadership training: Identifying gaps and opportunities. International Meeting for Simulation in Healthcare. San Francisco, CA. 2014.

4. Understanding and evaluating effectiveness of teams in emergency simulation scenarios. International Meeting for Simulation in Healthcare. San Francisco, CA. 2014.

5. Description of a Novel Team Leadership Training Program. International Symposium on Human Factors and Ergonomics in Health Care: Improving the Outcomes. Baltimore, MD. 2015.