Journal of Experimental Psychology: Applied
2011, Vol. 17, No. 3, 257–269

1076-898X/11/$12.00 DOI: 10.1037/a0025148

Interactions of Team Mental Models and Monitoring Behaviors Predict Team Performance in Simulated Anesthesia Inductions

Michael J. Burtscher and Michaela Kolbe
ETH Zurich, Zurich, Switzerland

Johannes Wacker
University Hospital Zurich, Zurich, Switzerland

Tanja Manser
University of Aberdeen

In the present study, we investigated how two team mental model properties (similarity vs. accuracy) and two forms of monitoring behavior (team vs. systems) interacted to predict team performance in anesthesia. In particular, we were interested in whether the relationship between monitoring behavior and team performance was moderated by team mental model properties. Thirty-one two-person teams consisting of anesthesia resident and anesthesia nurse were videotaped during a simulated anesthesia induction of general anesthesia. Team mental models were assessed with a newly developed measurement tool based on the concept-mapping technique. Monitoring behavior was coded by two organizational psychologists using a structured observation system. Team performance was rated by two expert anesthetists using a performance-checklist. Moderated multiple regression analysis revealed that team mental model similarity moderated the relationship between team monitoring and performance; a higher level of team monitoring in the absence of a similar team mental model had a negative effect on performance. Furthermore, team mental model similarity and accuracy interacted to predict team performance. Our findings provide new insights on factors influencing the relationship between team processes and team performance in health care. When investigating the effectiveness of a specific team coordination behavior, team cognition has to be taken into account. This represents a necessary and compelling extension of the popular process-outcome relationship on which previous teamwork research in health care has focused. Moreover, the current study adds further external validity to the concept of team mental models by highlighting its usefulness in health care.

Keywords: teamwork, monitoring behavior, team mental model, team performance, teamwork in health care

The influence of team processes on human performance in health care, and thus safe patient care, has recently become a major focus of research (Fletcher, McGeorge, Flin, Glavin, & Maran, 2002; Manser, 2009; Rall & Gaba, 2005; Risser et al., 1999). Team monitoring—observing the activities and performances of other team members (Dickinson & McIntyre, 1997)—is one of the team processes considered crucial for effective teamwork (LePine, Piccolo, Jackson, Mathieu, & Saul, 2008; Marks, Mathieu, & Zaccaro, 2001; Rousseau, Aube, & Savoie, 2006; Salas, Sims, & Burke, 2005). Team monitoring is very relevant in dynamic health care settings in which teams face unpredictable circumstances and have to cope with unexpected events (Kolbe, Burtscher, Manser, Künzle, & Grote, 2011). However, empirical studies on the functionality of team monitoring are sparse (Kolbe et al., 2010; Marks & Panzer, 2004). This is particularly true regarding the question whether the relationship between team monitoring and performance is influenced by properties of the team mental model (TMM); the organized understanding of relevant knowledge that is shared across team members (Cannon-Bowers, Salas, & Converse, 1993; Klimoski & Mohammed, 1994). In other words: Does effective team monitoring require that team members have a shared understanding of their common task? The current work aims to fill this gap. We investigated the relationship between team monitoring and performance in anesthesia teams focusing on the potential moderating role of TMM. Thereby, we attempted to provide insights on how team monitoring—a specific and trainable behavior—was related to team performance and patient safety.

During the past decade, a series of studies have focused on the role of teamwork for providing safe patient care. Researchers were able to establish a link between team process characteristics and outcome measures. For example, communication failures were

Michael J. Burtscher, Center for Organizational and Occupational Sciences, ETH Zurich, Zurich, Switzerland; Michaela Kolbe, Organization, Work, and Technology Group, ETH Zurich; Johannes Wacker, Institute of Anaesthesiology, University Hospital Zurich, Zurich, Switzerland; Tanja Manser, Industrial Psychology Research Centre, University of Aberdeen, Aberdeen, UK.

Michael J. Burtscher is now at Department of Psychology, University of Zurich; Johannes Wacker is now at Institute of Physicians for Anaesthesia and Intensive Care, Klinik Hirslanden Zurich; Tanja Manser is now at Department of Psychology, University of Fribourg.

This research was supported by an ETH Independent Investigators' Research Award (ETHIIRA grant ETH-26 09-3). We thank Bastian Grande and Renata Nohynkova for their help with team performance assessment, Silja S. Sollberger and Raphael Agosti for behavioral coding, and Kimberly A. Smith-Jentsch and Bertolt Meyer for their helpful comments on an earlier version of this paper. Furthermore, we thank Gudela Grote and Donat R. Spahn for the opportunity to collect data within a previous research project (Swiss National Science Foundation Grant 100013-116673).

Correspondence concerning this article should be addressed to Michael J. Burtscher, University of Zurich, Department of Psychology-Social Psychology, Binzmuehlestrasse 14/15, 8050 Zurich, Switzerland. E-mail: m.burtscher@psychologie.uzh.ch

shown to contribute to surgical injuries (Greenberg et al., 2007). Additionally, studies on anesthesia teams have related coordination processes to clinical performance (Burtscher et al., 2011; Burtscher, Wacker, Grote, & Manser, 2010; Manser, Harrison, Gaba, & Howard, 2009). Similar results from intensive care, neonatology, and surgery support the claim that teamwork is vital for effective and safe patient treatment (Catchpole, Giddings, et al., 2007; Reader, Flin, & Cuthbertson, 2007; Williams, Lasky, Dannemiller, Andrei, & Thomas, 2010).

Teamwork in health care is often discussed in the context of other high-risk industries, such as aviation, military, or the oil industry (Catchpole, de Leval, et al., 2007; Hudson, 2003; Thomas, Sexton, & Helmreich, 2004). These industries have recognized effective teamwork as a major factor to reduce errors and enhance safety (Shapiro et al., 2004; Wilson, Burke, Priest, & Salas, 2005). Teams in high-risk industries have been classified as action teams: teams of specialists performing tasks that require much coordinative effort and that are repeated under different conditions (Sundstrom, McIntyre, Halfhill, & Richards, 2000). In health care, anesthesia teams can be regarded as a classic example of an action team; they consist of specialists from different professions (physicians and nurses) who are often assembled ad hoc and work under changing conditions where hazardous nonroutine events are likely to occur at any time (Manser, 2009; Oken et al., 2007; Weinger & Slagle, 2002). Furthermore, anesthesia teams frequently engage in complex tasks that allow several solutions, which makes effective teamwork even more difficult.

In view of their similarities, it is not surprising that the transfer of concepts and methods from other high-risk industries—most notably from aviation—has helped to improve teamwork in health care (Gaba, 2011; Gaba, Howard, Fish, Smith, & Sowb, 2001; Salas, Rosen, & King, 2007; Thomas et al., 2004). In aviation, team monitoring and TMM are two of the most prominent concepts that have been successfully linked to team performance (Marks & Panzer, 2004; Smith-Jentsch, Mathieu, & Kraiger, 2005). To increase our understanding of team effectiveness in health care we used both concepts—team monitoring and TMM—and investigated their relationship to the performance of anesthesia teams.

Team monitoring—also termed mutual performance monitoring—comprises teammates assessing each other's work behaviors (Rousseau et al., 2006). Thereby, team members can ensure that procedures are being followed correctly and that everything is running as expected (McIntyre & Salas, 1995). Team monitoring has been considered particularly relevant for teams engaging in highly interdependent tasks in dynamic work environments (Marks & Panzer, 2004; Rousseau et al., 2006). Marks and Panzer's work (2004) supports this claim by linking team monitoring to the performance of simulated flight combat teams. Team monitoring is an example of a team coordination behavior. More specifically, team monitoring is a mechanism of *implicit* team coordination; it allows for anticipation of actions and needs of other team members and the subsequent adjustment of one's own behavior (Kolbe et al., 2011; Rico, Sánchez-Manzanares, Gil, & Gibson, 2008). In anesthesia, for example, team monitoring includes watching a teammate while he or she is inserting the respiratory tube into the patient's trachea. Team members who engage in team monitoring are more likely to identify mistakes in other members' actions and can react properly—for example, by providing backup (Marks et al., 2001; Salas et al., 2005).

Team members can also gain performance-relevant information through systems monitoring. Systems monitoring includes tracking team resources and environmental conditions as they are related to task accomplishment (Marks et al., 2001). Systems monitoring can help teams to more quickly identify nonroutine events, and thus, prevent them from turning into critical situations. In anesthesia, for example, systems monitoring includes keeping track of patients' vital signs to detect deviations (e.g., sudden drop in heart rate).

The relevance of monitoring behaviors has been highlighted for teams in dynamic domains of health care (Burke, Salas, Wilson-Donnelly, & Priest, 2004; Fernandez, Kozlowski, Shapiro, & Salas, 2008; Wilson et al., 2005). In an interview study with Advanced Life Support instructors, team monitoring was identified as a major nontechnical skill to improve team performance in multiprofessional cardiac arrest teams (Andersen, Jensen, Lippert, & Ostergaard, 2010). Moreover, teams in anesthesia were found to spend a significant amount of time on monitoring behaviors—especially during performance critical phases of the task (Manser, Howard, & Gaba, 2008). However, empirical evidence on the effectiveness of team monitoring in health care is still limited. In an observational study, team monitoring could not be directly related to the performance of anesthesia teams (Kolbe et al., 2010). However, higher-performing teams were found to exhibit different behaviors immediately after team monitoring than lower performing teams; for example, they were more likely to provide assistance to their teammates. One likely explanation is that they were able to do so because they had a similar mental model of their task allowing them to draw the right conclusions and react accordingly. For that reason, it appears that effective team monitoring is not simply achieved by watching teammates doing their work. To understand the meaning of team members' actions, to pace one's own task performance and to effectively provide others with support, a shared understanding of the joint task execution is needed; that is, teams need to have similar mental models of their task (Salas et al., 2005).

The following example illustrates how the interplay between TMM similarity and team monitoring can affect team performance in anesthesia. If team members agree on the order in which certain tasks are to be carried out, team monitoring helps coordinating their joint task execution. If one member notices her teammate looking about, she can conclude that he is looking for tape to fix the tube because this is the next procedural step according to her mental model. If both have a shared understanding of their task—that is, they agree on which subtask should be carried out next—her teammate will be happy to be handed the tape. If he has a different task model, he will disapprove of her action. His disapproval would possibly require further clarification and thus, interrupt the workflow. Hence, one should consider TMM when investigating the effectiveness of monitoring behaviors.

The important role of TMM for effective teamwork is emphasized in current theoretical team performance models (Burke et al., 2006; Salas et al., 2005). TMM similarity is considered to be positively related to team performance; that is, the greater the degree of similarity among team members, the better the team performance. This relationship can be mediated by team processes, such as communication (Marks, Zaccaro, & Mathieu, 2000), and

coordination (Marks, Sabella, Burke, & Zaccaro, 2002). Additionally, TMM have been shown to moderate the relationship between team processes and outcomes (Bonito, 2004).

In the literature, two distinct properties of TMM are discussed: similarity and accuracy (Mohammed, Ferzandi, & Hamilton, 2010). *Similarity* refers to the degree of sharedness between team members' mental models; *accuracy* can be defined as the average degree of sharedness between the team members' mental models and the mental model of an expert in the respective area. The existence of these two properties reflects the fact that having team members with a shared understanding of their work does not necessarily result in improved performance. To improve performance, a mental model should not only be shared but also has to be correct (Smith-Jentsch, 2009). Indeed, both properties interacted to predict team performance in a flight simulator task (Mathieu, Heffner, Goodwin, Cannon-Bowers, & Salas, 2005).

TMM research has been carried out in many domains, particularly in high-risk industries such as military and aviation (Lim & Klein, 2006; Marks et al., 2002; Smith-Jentsch, Cannon-Bowers, Tannenbaum, & Salas, 2008; Smith-Jentsch et al., 2005). Several authors have strongly advocated considering TMM when investigating teamwork in health care (Catchpole et al., 2006; Flin & Maran, 2004; Healey, Undre, & Vincent, 2004; Undre, Healey, Darzi, & Vincent, 2006). So far, however, no empirical study on TMM in health care has been published (DeChurch & Mesmer-Magnus, 2010; Mohammed, Ferzandi, & Hamilton, 2010). One of the main reasons for this lack of empirical studies might be the fact that TMM measurement is a difficult matter, especially in complex dynamic settings (Ellwart, Biemann, & Rack, 2011).

We addressed these research gaps by investigating the interaction of TMM and monitoring behaviors in predicting the performance of teams of anesthesia professionals during a simulated induction of general anesthesia. Thereby, we aim to provide new insights into the preconditions of effective teamwork in health care action teams. This would improve our understanding of how these teams deal with their ad hoc composition and dynamic work environment, potentially offering new directions for team design and team training.

The Present Study

We were interested in how the two properties of TMM (similarity vs. accuracy) and the two forms of monitoring (team vs. systems) would interact to predict team performance. Based on the research outlined above, we formulated four hypotheses. *TMM similarity* would moderate the relationship between *team monitoring* and *team performance* (H1), as efficient team monitoring requires team members to have a shared understanding of their common task (Salas et al., 2005). We also hypothesized that *TMM accuracy* would moderate the relationship between *systems monitoring* and *team performance* (H2), as successful detection of irregularities requires a precise picture of the task; particularly, in terms of which environmental conditions are related to task accomplishment. Furthermore, we hypothesized that *both TMM properties* would interact to predict *team performance* (H3)—as was shown in comparable high-risk settings (Mathieu et al., 2005). Finally, we hypothesized that *team monitoring* and *systems monitoring* would interact with each other to predict *team performance* (H4), as they were both thought to be crucial for action teams in dynamic environments (Rousseau et al., 2006).

Method

The study was approved by the local ethics committee (KEK-StV-Nr. 09/05) and registered at ClinicalTrial.gov (NCT00706108). Written consent was obtained from all participants before data collection. All participants were explicitly informed that they could withdraw from the simulation at any time. One team withdrew after the simulation, so the final sample included 31 teams.

Participants

Participants were a convenience sample comprising 31 anesthesia residents and 31 anesthesia nurses from a large Swiss teaching hospital working in teams of two. Residents (14 women) were between the ages of 26 and 37 years ($M = 30.35$, $SD = 2.11$); their professional experience ranged between 3 months and 5 years ($M = 2.10$, $SD = 1.28$). Nurses (23 women) were between the ages of 26 and 55 years ($M = 33.4$, $SD = 8.13$); their professional experience ranged between 4 months and 20 years ($M = 3.61$, $SD = 5.12$). All participants had been working in the hospital's department of anesthesiology for at least 3 months guaranteeing sufficient acquaintance with equipment and procedures. We ensured that teams were assembled ad hoc—participants did not know beforehand with whom they were to form a team.

Apparatus

Simulations were conducted in a genuine operating room to provide a representative research setting. Moreover, the hospital's regular anesthesia equipment including syringes, stethoscope, and standard airway devices was provided. We used a simulator mannequin for advanced life support training (Laerdal mannequin allowing mask ventilation, intubation, and rhythm simulation using Megacode heart rhythm simulator), which was in all cases operated by the same experienced anesthesia attending.

Participants were videotaped during the entire scenario using two cameras: one positioned to give a full view; one to focus on the work area of the anesthesia team during induction. Audio was recorded using a ceiling microphone. Additionally, we recorded data from the vital signs monitor (e.g., heart-rate) and the ventilator (e.g., airway pressures). All data channels were synchronized via a master-slave recording system to assure simultaneity during playback.

Task

The current study focuses on anesthesia induction—the first stage of general anesthesia. During this stage, certain drugs are administered to the patient to ensure unconsciousness, analgesia (pain suppression), and muscle relaxation. Furthermore, an endotracheal tube is inserted to enable mechanical ventilation.

Anesthesia induction with tracheal intubation was chosen for three reasons. First, it requires a high level of teamwork and coordination making it more suitable to investigate team processes (Phipps, Meakin, Beatty, Nsoedo, & Parker, 2008). Second, during induction of anesthesia, teams operate for the most part with little

interference from outside. By contrast, during other periods, teams have to cooperate closely with surgeons that complicates the investigation of genuine anesthesia team processes, that is, distinguishing intrateam from interteam-coordination (Helmreich & Schaefer, 1994). Finally, intubation can be regarded as a highly representative task in anesthesia and emergency medicine enhancing the potential generalizability of the current study's findings.

The simulation scenario included a 25-year old male "patient" scheduled for right knee arthroscopy. He was described as being healthy, his intubation anatomy looked normal, and his teeth were normal. Participants were informed that, based on the preoperative anesthesia meeting the evening before, a general anesthetic with tracheal intubation using an intravenous technique had been planned. The patient had already been brought to the operating room and drugs and equipment had been prepared by another nurse who had her break now. Participants were asked to proceed with the anesthetic as quickly as possible. This included anesthetizing the patient and inserting an endotracheal tube in his airway. During the scenario, several common nonroutine events—deviations from optimal care (Weinger, Slagle, Jain, & Ordonez, 2003)—such as a sudden drop in the patient's blood pressure were simulated. This was to reflect the daily practice of anesthesia and to make the scenario more challenging. A staff anesthetist was available on call.

Measures

Team mental models. As mentioned above, a variety of different approaches have been used to measure TMM. To obtain an overview of the settings and methods where TMM have been assessed, we reviewed the relevant literature particularly focusing on the TMM measurement procedures and the effects of TMM on team performance (Cannon-Bowers & Salas, 2001; DeChurch & Mesmer-Magnus, 2010; Mohammed et al., 2010; Mohammed, Klimoski, & Rentsch, 2000).

Our goal was to identify a method that was meaningful to health care practitioners, expeditious for the use in simulation-based research, and met scientific standards for TMM assessment. We discussed potential measurement methods with anesthetists, experts in medical simulation, and psychologists with experience in TMM measurement. After thorough consideration, we decided to use a concept-mapping approach. This technique has been successfully used in previous experimental studies with action teams (Ellis, 2006; Marks et al., 2002; Marks et al., 2000). It requires participants to sort different concepts into a given structure (map), for example, different subtasks of a process have to be brought in the right chronological order. The resulting map represents the participant's mental model.

To identify relevant concepts (i.e., subtasks) for anesthesia induction, we interviewed subject matter experts—five senior anesthetists and five experienced anesthesia nurses—and reviewed relevant protocols and guidelines. The resulting concepts were then discussed with three other senior anesthetists. Finally, we agreed on 30 concepts relevant to the team task of induction of general anesthesia.

Each participant was given the 30 concepts printed on small magnets and a magnet board with the map (see Figure 1). Then they were asked to fill in the map with the subtasks thereby indicating which team member should perform which task and putting them in chronological order.

TMM similarity. To obtain a measure of similarity, individual maps were compared within each team. One point for similarity was given if a concept was assigned to the same person; one additional point was given if (a) concepts were placed in the same position and (b) had the same previous concept (i.e., concept b was placed after concept a regardless whether they had different absolute positions).

TMM accuracy. To assess the accuracy of the TMM, we asked two subject matter experts to do the mapping task. Participants' maps were then compared with both experts' models using the scoring algorithm described above. The results of both comparisons were significantly correlated ($r = .51$, $p = .003$); we collapsed them into a single rating.[1] These individual accuracy ratings were in turn averaged within the team to obtain a team-level measure of accuracy. Both measures' possible range was from 0 to 88 points with higher scores indicating larger degrees of similarity and accuracy.

Monitoring behavior. Team monitoring and systems monitoring were assessed on the basis of video observations using Interact—a specialized software for behavioral observation (Mangold International GmbH, Arnstorf, Germany). Team monitoring was coded each time a team member was observing the actions of his or her teammate, that is, one team member watched what the other team member was doing. Systems monitoring was coded each time a team member watched the patient, the vital signs monitor, or the display of the ventilator. Two trained observers—graduate students with experience in behavioral coding—rated all video recordings. To assess interrater reliability, three randomly selected videos representing 10% of the data were coded by both observers. Subsequent calculation of Cohen's Kappa revealed good to excellent agreement of .76 for team monitoring and of .88 for systems monitoring (Cohen, 1960; Landis & Koch, 1977).

Team and systems monitoring scores were calculated for both team members. They were defined as the time spent on the respective monitoring behavior in relation to the total duration of the simulation. For example, if an anesthesia nurse spent 2 min of a 20-min simulation on systems monitoring his or her systems monitoring score was 10%. To allow for analyses at the team-level, individual scores were aggregated for each team. For example, if a nurse had a score of 10% and a resident one of 5%, the team score was 15%. Because percentages are in fact nominal data, we used arcsine transformation as has been done in similar studies (Tschan et al., 2006).

Team performance. Two experienced staff anesthetists rated team performance using a checklist comprising 83 items. The checklist had been validated using Delphi-technique and was used in a previous study (Burtscher et al., 2010). It constitutes an institutional standard for performing an induction of general an-

[1] Previous studies aggregated accuracy scores at higher levels of agreement (e.g., between .70 and .80; Marks, Zaccaro, & Mathieu, 2000). These studies, however, involved more artificial settings (e.g., computer simulations, student participants) and less complex tasks. By contrast, the current study uses medical professionals performing a realistic task in a representative environment. In view of these circumstances, we feel justified in collapsing over the accuracy scores at a moderate level of agreement.

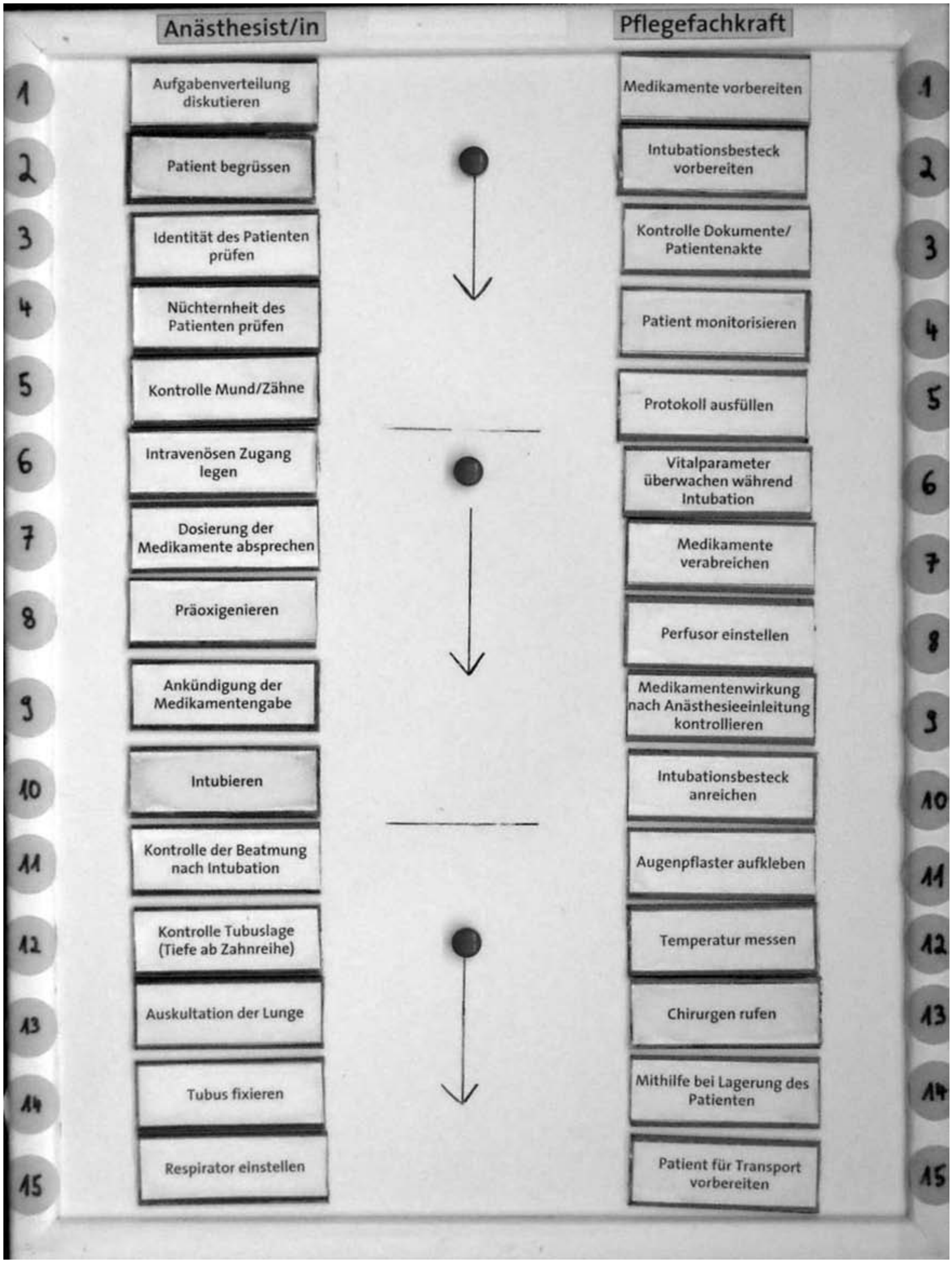

Figure 1. Photo of a completed concept map. Please note that the concepts are labeled in German. Examples include *ventilate patient*, *adjust head position*, *hand intubation set*, and *announce drug administration*. A complete translation can be obtained from the first author.

esthesia. In the current study, the checklist served as a measure for the degree to which the teams were able to deal with the simulated nonroutine events and meet a performance standard. A perfect performance—resulting in a score of 100%—means that the respective team conducted the induction perfectly according to the standard despite the occurrence of several nonroutine events. Again, to establish interrater reliability, three randomly selected videotapes were coded by both raters. Interrater reliability was subsequently calculated using the ICC coefficient (Shrout & Fleiss, 1979), which was 0.98, indicating high reliability.

Procedure

The study represents an observational study with two measurement points (before and during the simulation). Teams were assembled ad hoc based on individual availability. Upon arrival, participants were told about the purpose and procedure of the study and were handed the consent forms. They discussed questions and potential concerns with one of the experimenters. After having consented, participants were separated. They completed a personal background questionnaire and the concept-mapping task used to assess their mental models. Next, they were brought together again and received a detailed technical briefing about the simulation scenario and the features of the simulator mannequin. The simulation started the moment the team entered the operating room. After the simulation, participants received video-based de-briefing on their teamwork and performance.

Results

The duration of the 31 cases ranged from 17.74 (17 min 44 sec) to 33.09 (33 min 5 sec) minutes ($M = 21.45$, $SD = 4.21$). Table 1 contains descriptive statistics and correlations between all study variables. Neither team monitoring nor systems monitoring was related to team performance, nor were both measures significantly correlated. Similarly, neither accuracy nor similarity of TMM was directly related to team performance. However, both measures were positively correlated. In addition, participants' professional experience did not show a significant correlation with any of the study variables.

We used moderated multiple regression analysis to test our hypotheses (Cohen, Cohen, West, & Aiken, 2003). Variables were mean centered before the interaction term was created to reduce potential multicollinearity. Participants' professional experience was entered as a control variable. When we detected a significant interaction, we further analyzed the pattern that caused this interaction. In particular, we performed simple slope analyses using an online tool (Preacher, Curran, & Bauer, 2006).

H1: Team Monitoring and Team Mental Model Similarity Interact to Predict Team Performance

To test this hypothesis, we regressed team performance on both variables and added their interaction in a second step. The regression yielded nonsignificant beta weights for both variables (see Table 2), whereas adding the interaction term produced a significant $\Delta R^2 = .13$, $p = .042$. Furthermore, the full regression model including all three predictors and the control variable professional experience accounted for about 28% of the variance in the performance rating ($R^2 = .28$, $p = .065$). Thus, hypothesis 1 was supported. The interaction between team monitoring and TMM is depicted in Figure 2; values of 1 *SD* above and below the means were used to indicate high and low values. As the plot shows, in the absence of a sufficiently similar TMM of the common task, a higher level of team monitoring had a negative effect on performance. Additional simple slope analyses confirmed that when TMM similarity was low (-1 *SD*), team monitoring was negatively related to team performance (slope $= -2.22$), $t(26) = -2.77$, $p = .010$. By contrast, when TMM similarity was high (1 *SD*), team monitoring had no effect on team performance (slope $= 0.20$), $t(26) = 0.23$, $p = .82$. The region of significance for TMM similarity ranges form -1.31 to 148.69, indicating that any given slope outside this region is significant. Centered TMM similarity ranges from -12.65 to 12.35. Therefore, we can conclude that for teams with a less than average degree of TMM similarity, team monitoring is negatively related to team performance. For teams with an average or better degree of TMM similarity, there was no relationship between both variables. Even if a team had a perfectly similar TMM, it would not benefit from team monitoring since the upper bound of 148.69 exceeds the possible range of the similarity rating.

Table 1
Summary of Intercorrelations, Means, and SD for All Study Variables

Variables	Professional experience	TMM similarity	TMM accuracy	Team monitoring	Systems monitoring	Team performance
Professional experience	1.0					
TMM similarity	.17	1.0				
TMM accuracy	.24	.65**	1.0			
Team monitoring	−.17	−.07	.14	1.0		
Systems monitoring	−.08	−.02	.22	.29	1.0	
Team performance	−.20	.12	−.08	−.27	.16	1.0
M	2.92	34.65	36.40	5.62	24.62	66.36
SD	2.56	6.44	3.16	2.77	7.95	9.99

Note. $N = 31$ teams. All variables are at the team-level. Professional experience refers to the averaged experience in anesthesia per team measured in years.
** $p < .01$.

Table 2
Results of the Four Moderated Multiple Regression Analyses Predicting Team-Performance

	Step 1		Step 2		Step 3	
	β	*p*	β	*p*	β	*p*
Predictor						
Professional experience	−.20	.29	−.27	.15	−.20	.27
TMM similarity (centered)			.14	.43	.14	.42
Team monitoring (centered)			−.31	.10	−.28	.11
Similarity × Team monitoring					.36	.04
R^2	.04		.15		.28†	
ΔR^2			.11		.13*	
Predictor						
Professional experience	−.20	.29	−.17	.40	−.20	.32
TMM accuracy (centered)			−.08	.69	−.06	.76
Systems monitoring (centered)			.16	.41	.18	.36
Accuracy × Systems monitoring					−.16	.42
R^2	.04		.07		.09	
ΔR^2			.03		.02	
Predictor						
Professional experience	−.20	.29	−.19	.32	−.18	.30
TMM accuracy (centered)			−.23	.35	−19	.41
TMM similarity (centered)			.30	.22	.20	.37
Accuracy × Similarity					.42	.02
R^2	.04		.09		.26†	
ΔR^2			.05		.17*	
Predictor						
Professional experience	−.20	.29	−.24	.18	−.20	.25
Team monitoring (centered)			−.38	.05	−.44	.02
Systems monitoring (centered)			.25	.18	.21	.23
Team × Systems monitoring					−.36	.04
R^2	.04		.19		.31*	
ΔR^2			.15		.12*	

Note. $N = 31$ teams. All variables are at the team-level. Professional experience refers to the averaged experience in anesthesia per team measured in years.
† $p < .10$. * $p < .05$.

H2: Systems Monitoring and Team Mental Model Accuracy Interact to Predict Team Performance

Our second hypothesis was not supported. The regression yielded nonsignificant beta weights for both variables as well as for their interaction (see Table 2). This indicates that, contrary to our assumption, having versus not having an accurate mental model of the common task does not influence the effectiveness of systems monitoring.

H3: Team Mental Model Accuracy and Similarity Interact to Predict Team Performance

Hypothesis 3 was supported. As expected, the interaction between TMM accuracy and similarity was significant (see Table 2), it accounted for about 17% of the variance in the performance rating. To further analyze this interaction we performed simple slope analyses. When TMM accuracy was low (−1 *SD*), TMM similarity was not related to team performance (slope = −0.004), $t(26) = -0.96$, $p = .34$. By contrast, when TMM accuracy was high (1 *SD*), TMM similarity was positively related to team performance (slope = 0.01), $t(26) = 2.08$, $p = .048$. The region of significance for TMM accuracy ranges form −9.58 to 3.03, indicating that any given slope outside this region is significant. Given that centered TMM accuracy ranges from −7.15 to 6.60, this indicates that the effect of TMM similarity on team performance is significant only for relatively high values—almost 1 *SD* above the mean—of TMM accuracy. Having a similar mental model of the task is only beneficial if this model is also accurate.

H4: Team Monitoring and Systems Monitoring Interact to Predict Team Performance

In support of Hypothesis 4, we found a significant interaction between both types of monitoring. Interestingly, team monitoring was negatively related to team performance (see Table 2). The full regression model including all three predictors and the control variable professional experience accounted for nearly a third of the variance in the performance rating ($R^2 = .31$, $p = .043$). Figure 3 shows a plot of the interaction. Simple slope test revealed that when team monitoring was low (−1 *SD*), systems monitoring was positively related to team performance (slope = 0.79), $t(26) = 2.56$, $p = .017$; whereas for high team monitoring (1 *SD*) there was no relationship between systems monitoring and team performance (slope = −0.27), $t(26) = -0.80$, $p = .43$. The region of significance for team monitoring ranges from −0.01 to 0.46, indicating that any given slope outside this region is significant. Centered team monitoring ranges from −0.04 to 0.08. In view of this finding, it appears to have been more beneficial to focus only on

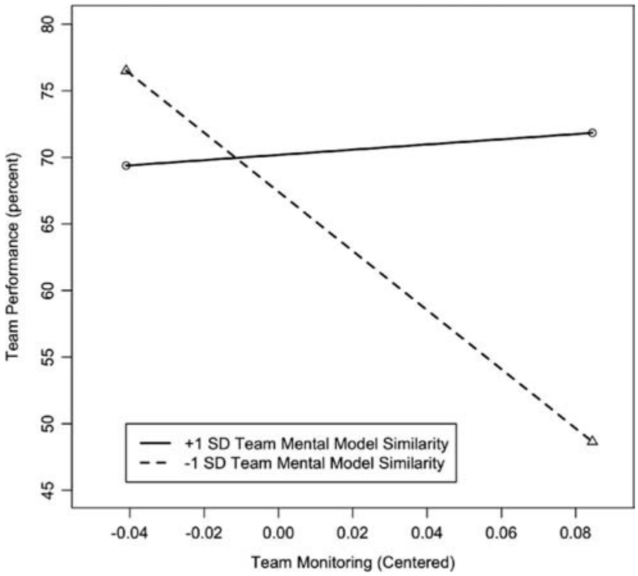

Figure 2. Team monitoring × Team mental model similarity interaction as related to the performance of the anesthesia teams.

systems monitoring. In fact, teams that tried to capitalize on both team monitoring and systems monitoring showed the worst performance (see Figure 3).

Discussion

We investigated how TMM properties (similarity vs. accuracy) and two forms of monitoring (team vs. systems) interacted to predict the performance of anesthesia teams in a simulated induction of general anesthesia. This study represents an important contribution to research on action teams in health care as it increases our understanding of how the relationship between team coordination behaviors (i.e., monitoring) and performance is moderated by team cognition. Three main findings emphasize the significance of these interactions between team cognition and observable team behaviors.

First, our findings suggest that a similar TMM is a prerequisite for effective team monitoring. Without a shared understanding of the task, more team monitoring disrupts performance. Team members need to know in advance what to monitor; otherwise team monitoring can be both misleading and a waste of resources. If team members have a shared understanding of their common task, team monitoring can facilitate coordinating the joint task execution. For example, if a team member recognizes that a teammate has started a certain procedure, he or she can begin working on a complementary task. In this case, team monitoring in combination with a similar TMM substitutes more resource-intensive explicit coordination mechanisms. Moreover, team monitoring may provide team members with an implicit update on whether their teammates are still on track. If they do not seem to be on track, one could provide assistance or instructions (Kolbe et al., 2010). Based on this finding, we conclude that health care action teams will benefit from establishing a TMM before performing as a team, for example, during preoperational briefings (Einav et al., 2010).

Second, it is not sufficient to have a shared understanding of the task. Our results confirm that to be effective, a mental model should not only be shared but also has to be correct (Mathieu et al., 2005; Smith-Jentsch, 2009). Applying this finding to preoperational briefings one may conclude that they would optimally involve subject matter experts (i.e., an attending anesthetist). These experts could ensure the accuracy of the to-be-established TMM.

Third, our findings suggest that teams should carefully balance their teamwork behaviors. In our study, anesthesia teams showing high levels of both team and systems monitoring had the lowest performance scores. Whereas the effectiveness of systems monitoring was, unexpectedly, not influenced by having versus not having an accurate TMM, it was influenced by team monitoring. Only when team monitoring was low, was systems monitoring positively related to team performance. One may conclude that spending much time on both team and systems monitoring subtracts too many resources (e.g., time) from the actual task. This finding is in line with research on the appropriate amount of backup behavior—another crucial teamwork behavior that involves providing resources and task-related effort to struggling teammates (Fernandez et al., 2008; Salas et al., 2005). Research suggests that team members risk neglecting their own taskwork, if they engage in too much backup behavior (Barnes et al., 2008; Porter, Gogus, & Yu, 2010).

In summary, our results provide new insights on factors influencing the relationship between team processes and performance in health care—a necessary and compelling extension of the popular process-outcome relationship on which previous research has focused. Relying on the input-process-outcome model (Mathieu, Maynard, Rapp, & Gilson, 2008; McGrath, 1984), researchers were able to relate specific observable teamwork behaviors (e.g., information sharing) to outcome variables (e.g., occurrence of errors or the duration of an operation) (Catchpole, Mishra, Handa,

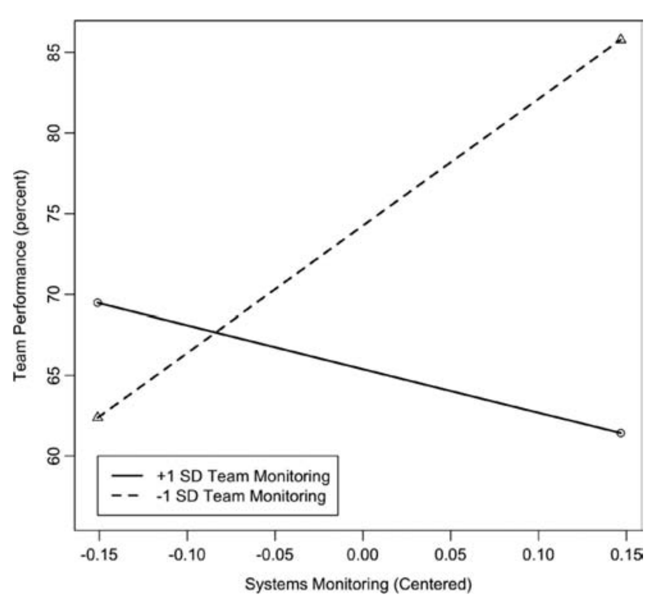

Figure 3. Team monitoring × Systems monitoring interaction as related to the performance of the anesthesia teams.

& McCulloch, 2008; Manser et al., 2009; Mazzocco et al., 2009; Williams et al., 2010). The implicit assumption of these studies seems to be that exhibiting more of a specific behavior—for example, more information sharing—would invariably lead to better performance—for example, fewer errors. This linear relationship may hold true for some settings. However, in dynamic health care domains it is very likely that the relationship between teamwork and performance is more complex and that performance cannot always be improved by increasing the level of certain behaviors (Kolbe et al., 2011; Rico et al., 2008). Our results provide support for this claim. Team monitoring was not directly related to team performance. Even more, in the absence of a similar TMM, team monitoring disrupted performance.

The following situation immediately after intubation of the patient exemplifies the relevance of a similar TMM for effective team monitoring: A team member observes her teammate and concludes that she is struggling with the fixation of the endotracheal tube. Hence, she moves over and attempts to assist. The teammate, however, is not struggling but is—while apparently busy with the tube—expecting a comment regarding the appearance of a CO_2 reading on the monitor, which is an important evidence for correct tube placement in the trachea. Both members have different mental models of the situation or the priorities, and thus, different opinions of what should be done next. As a result, the teammate may consider this assistance inappropriate. In contrast, a similar TMM in the example above would have prompted an unsolicited confirmation of the CO_2 reading by the assisting team member. In particular, if the intubating teammate is still busy with details of tube fixation and—as it is often the case at this stage—he or she has no direct view of the monitor while maintaining the visual focus on the tube.

Without an appropriate TMM, explicit coordination (e.g., clarifying questions) can compensate for the lack of a shared understanding by reducing uncertainty and getting everyone "on the same page." In fact, explicit coordination has shown to be an effective strategy for teams treating a cardiac arrest (Marsch et al., 2004). Moreover, it could be useful to establish a shared understanding of the task before the actual task execution. During briefings, for example, explicit coordination can help to clarify procedures (e.g., "before we fix the tube, we need to check whether the tube has been placed correctly"). Postoperational debriefings represent another possibility to enhance team members' shared understanding. In this context, a constructive discussion of errors and misunderstandings can foster the establishment of a similar and accurate TMM.

This leads to a more general contribution of this study: an enhanced understanding of the interplay between team cognition and coordination behavior in action teams. Some researchers considered implicit coordination particularly effective in those teams (Entin & Serfaty, 1999). By using implicit coordination behaviors, such as team monitoring, teams would be able to function more efficiently because this way of coordination is smoother and less time consuming compared to explicit coordination. Implicit coordination, however, can only be successful if team members hold a similar and accurate TMM of their task (Kolbe et al., 2009; Rico et al., 2008; Stout, Cannon-Bowers, Salas, & Milanovich, 1999). The current work represents, to the best of our knowledge, the first empirical study to identify a link between TMM, monitoring as implicit coordination behavior, and team performance. Furthermore, this study adds further external validity to the concept of TMM by transferring results from aviation—for example, on the interaction of TMM properties (Mathieu et al., 2005)—to health care. Our results also suggest, in line with previous research (Cooke, Gorman, Duran, & Taylor, 2007), that the underlying mechanisms of team cognition can be much better understood if one considers associated team processes, such as team and systems monitoring.

Our study also provides new insight into the functionality of monitoring behavior. We assessed two forms of monitoring—team and systems monitoring—simultaneously in an applied setting. Our findings indicate that both forms have to be taken into account when attempting to determine the importance of monitoring for performance. The fact that team monitoring was not positively related to team performance and that teams trying to focus on both forms of monitoring had the lowest performance scores could also be a peculiarity of the current setting. In a two-person team, one team member concentrates on the primary manual task (e.g., inserting the endotracheal tube) while the other engages in monitoring behaviors. This person has to decide whether to keep track of his or her teammate or to observe the work environment. Intubation is a relatively standardized task and the nonroutine events could be easier detected when focusing on the work environment. Therefore, it is not surprising that systems monitoring was more beneficial. Team monitoring may be more beneficial in a less standardized team-working situation and with increasing team size. Although our results suggest that trying to monitor both the teammate and the environment exceeds the capacity of a single person, the same situation may be different for a three-person team: one member could focus on team monitoring while the other focuses on systems monitoring. Another reason for our inability to detect a positive effect of team monitoring might be a potentially nonlinear relationship between team monitoring and team performance. As mentioned above, research on backup behavior suggests that too much backup behavior may be dysfunctional (Barnes et al., 2008). The same could be true for team monitoring; the relationship to team performance could be curvilinear with both too little and too much team monitoring having a negative effect on team performance.

Another aspect worth mentioning concerns the relatively low average values of both TMM similarity and accuracy. They can be attributed to the peculiarities of the current setting. Teams were assembled ad hoc and had no opportunity to discuss the scenario before the start of the simulation. This situation, however, reflects the actual working conditions in anesthesia. To some extent, the relatively low similarity and accuracy values might be caused by the different training settings of residents and nurses. Even though their training focuses on the same practical procedures, differing training priorities may still result in reduced TMM similarity. Furthermore, compared to previous studies on TMM, teams had to perform a relatively complex task; particularly, since we introduced several unexpected events. As a result, several correct solutions were possible, which could explain the relatively low accuracy scores. Interestingly, even under these circumstances, both TMM similarity and accuracy affected team performance.

Limitations of the Present Study

The limited sample size because of the inclusion of anesthesia professionals prevented us from testing all interactions simultaneously—for example, the three-way interaction between team monitoring and TMM similarity and TMM accuracy. While theoretically important to differentiate these two aspects of the TMM as potential moderators, it was difficult to do in our study—even more so as similarity and accuracy were correlated. In addition, we did not use an experimental design, which precludes us from drawing inferences about causal relationships. However, using different measurement points and methods to assess key variables in combination with regression analysis strongly suggests that differences in TMM were a cause for the observed differences in team performance. Another limitation concerns team composition. In the current study, nurses were on average more experienced than residents resulting in diverse teams with regard to professional experience. Although teams of more experienced nurses and less experienced residents are very common, the overrepresentation of this team composition may limit the generalizability of our findings. Finally, as mentioned above, teams in the current study exhibited relatively low averages of accuracy and similarity. The associated limited range of both indices may reduce the ability to generalize our findings.

Future Research

In view of the limitations of the current study, a logical next step would involve a systematic manipulation of TMM by training or briefing procedures. Cross-training, for example, has been shown to positively affect similarity and accuracy of team knowledge (Cooke et al., 2003). An experimental manipulation would help to clarify causal relationships between TMM, monitoring behavior, and team performance. In addition, future studies should investigate whether TMM also affect the influence of other implicit coordination mechanisms—for example, talking to the room (Kolbe et al., 2010; Tschan et al., 2009; Waller & Uitdewilligen, 2008)—on team performance. We would expect that TMM properties would also moderate this relationship. Another interesting point concerns the role of backup behavior. Future studies should investigate whether the influence of team monitoring on team performance is mediated by backup behavior. Furthermore, different task characteristics (e.g., standardization, interdependence) and team characteristics (e.g., team member familiarity, team size) might influence the effectiveness of monitoring behaviors. Interestingly, our findings regarding the functionality of TMM similarity for team monitoring point to a potential threshold level. Whereas at low levels of similarity more team monitoring was associated with poor performance, at high levels of similarity team monitoring did not affect team performance. It might be the case that in some situations, TMM would be more appropriately conceptualized in terms of a binary construct; that is, a team either has or has not a sufficiently similar model. Finally, the developmental aspect of monitoring and its relation to performance should be studied in more detail. On the one hand, we would assume that team monitoring varies as a function of other team characteristics. For example, teams that—unlike those in the current study—work together for a long time have the possibility to develop mutual feelings of trust. High levels of trust were associated with low levels of monitoring (Langfred, 2004). It would be interesting to investigate the influence of team-level affective states, such as feelings of mutual trust or group potency (Jung & Sosik, 2003), on the level of team monitoring. On the other hand, it is worth considering that the relationship between monitoring and team performance is nonlinear. In this sense, detecting patterns of different team processes—instead of focusing on linear process-outcome relations—should be a focus of future research.

Conclusions

Our findings highlight the importance of considering both TMM and monitoring behavior when investigating the performance of health care teams. In the absence of a similar TMM, higher levels of team monitoring were associated with lower levels of team performance. Furthermore, both TMM similarity and TMM accuracy interacted to predict team performance. Because many teams in health care are assembled ad hoc and exist only for a limited time, they do not have the possibility to develop a similar TMM during their work. Moreover, merely working together does not necessarily result in a similar and accurate TMM (Mohammed et al., 2000). Therefore, interventions that help teams of health care professionals to develop a similar and accurate mental representation of their task might be beneficial for team performance. As findings from other high-risk industries indicate, cross-training and guided team self-correction can foster TMM development (Cannon-Bowers, Salas, & Blickensderfer, 1998; Marks et al., 2002; Smith-Jentsch et al., 2008).

References

Andersen, P. O., Jensen, M. K., Lippert, A., & Ostergaard, D. (2010). Identifying non-technical skills and barriers for improvement of teamwork in cardiac arrest teams. *Resuscitation, 81,* 695–702. doi:10.1016/j.resuscitation.2010.01.024

Barnes, M. C., Hollenbeck, J. R., Wagner, D. T., DeRue, D. S., Nahrgang, J. D., & Schwind, K. M. (2008). Harmful help: The costs of backing-up behavior in teams. *Journal of Applied Psychology, 93,* 529–539. doi:10.1037/0021-9010.93.3.529

Bonito, J.-A. (2004). Shared Cognition and Participation in Small Groups: Similarity of Member Prototypes. *Communication Research, 31,* 704–730. doi:10.1177/0093650204269406

Burke, C. S., Salas, E., Wilson-Donnelly, K., & Priest, H. (2004). How to turn a team of experts into an expert medical team: Guidance from the aviation and military communities. *Quality and Safety in Health Care, 13*(Suppl), i96–i104. doi:10.1136/qshc.2004.009829

Burtscher, M. J., Manser, T., Kolbe, M., Grote, G., Grande, B., Spahn, D. R., & Wacker, J. (2011). Adaptation in anaesthesia team coordination in response to a simulated critical event and its relationship to clinical performance. *British Journal of Anaesthesia.* doi:10.1093/bja/aer039

Burtscher, M. J., Wacker, J., Grote, G., & Manser, T. (2010). Managing non-routine events in anesthesia: The role of adaptive coordination. *Human Factors, 52,* 282–294. doi:10.1177/0018720809359178

Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. [Review]. *Journal of Organizational Behavior, 22,* 195–202. doi:10.1002/job.82

Cannon-Bowers, J. A., Salas, E., & Blickensderfer, E. (1998). The impact of cross-training and workload on team functioning: A replication and extension of initial findings. *Human Factors, 40,* 92–101. doi:10.1518/001872098779480550

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Indi-*

vidual and group decision making (pp. 221–246). Hillsdale, NJ: Lawrence Erlbaum Associates.

Catchpole, K. R., de Leval, M. R., McEwan, A., Pigott, N., Elliott, M. J., McQuillan, A., & Goldman, A. J. (2007). Patient handover from surgery to intensive care: Using Formula 1 pit-stop and aviation models to improve safety and quality. *Paediatric Anaesthesia, 17,* 470–478. doi:10.1111/j.1460-9592.2006.02239.x

Catchpole, K. R., Giddings, A. E., de Leval, M. R., Peek, G. J., Godden, P. J., Utley, M., & Dale, T. (2006). Identification of systems failures in successful paediatric cardiac surgery. *Ergonomics, 49,* 567–588. doi:10.1080/00140130600568865

Catchpole, K. R., Giddings, A. E., Wilkinson, M., Hirst, G., Dale, T., & de Leval, M. R. (2007). Improving patient safety by identifying latent failures in successful operations. *Surgery, 142,* 102–110. doi:10.1016/j.surg.2007.01.033

Catchpole, K. R., Mishra, A., Handa, A., & McCulloch, P. (2008). Teamwork and error in the operating room: Analysis of skills and roles. *Annals of Surgery, 247,* 699–706. doi:10.1097/SLA.0b013e3181642ec8

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46. doi:10.1177/001316446002000104

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Cooke, N. J., Gorman, J. C., Duran, J. L., & Taylor, A. R. (2007). Team cognition in experienced command-and-control teams. *Journal of Experimental Psychology: Applied, 13,* 146–157. doi:10.1037/1076-898X.13.3.146

Cooke, N. J., Kiekel, P. A., Salas, E., Stout, R., Bowers, C., & Cannon-Bowers, J. (2003). Measuring team knowledge: A window to the cognitive underpinnings of team performance. *Group Dynamics: Theory, Research, and Practice, 7,* 179–199. doi:10.1037/1089-2699.7.3.179

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology, 95,* 32–53. doi:10.1037/a0017328

Dickinson, T. L., & McIntyre, R. M. (1997). A conceptual framework for teamwork measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement* (pp. 19–43). Mahwah, NJ: Lawrence Erlbaum

Einav, Y., Gopher, D., Kara, I., Ben-Yosef, O., Lawn, M., Laufer, N., & Donchin, Y. (2010). Preoperative briefing in the operating room: Shared cognition, teamwork, and patient safety. *Chest, 137,* 443–449. doi:10.1378/chest.08-1732

Ellis, A-P. J. (2006). System breakdown: The role of mental models and transactive memory in the relationship between acute stress and team performance. *Academy of Management Journal, 49,* 576–589. doi:10.5465/AMJ.2006.21794674

Ellwart, T., Biemann, T., & Rack, O. (2011). Measurement of team knowledge in the field: Methodological advantages and limitations. In M. Boos, M. Kolbe, P. Kappeler, & T. Ellwart (Eds.), *Coordination in human and primate groups* (pp. 155–176). Heidelberg: Springer. doi:10.1007/978-3-642-15355-6_9

Entin, E. E., & Serfaty, D. (1999). Adaptive Team Coordination. *Human Factors, 41,* 312–325. doi:10.1518/001872099779591196

Fernandez, R., Kozlowski, S. W. J., Shapiro, M. J., & Salas, E. (2008). Toward a Definition of Teamwork in Emergency Medicine. *Academic Emergency Medicine, 15,* 1104–1112. doi:10.1111/j.1553-2712.2008.00250.x

Fletcher, G. C. L., McGeorge, P., Flin, R. H., Glavin, R. J., & Maran, N. J. (2002). The role of non-technical-skills in anaesthesia: A review of current literature. [Review]. *British Journal of Anaesthesia, 88,* 418–429. doi:10.1093/bja/88.3.418

Flin, R., & Maran, N. (2004). Identifying and training non-technical skills for teams in acute medicine. *Quality and Safety in Health Care, 13*(Suppl), i80–i84. doi:10.1136/qshc.2004.009993

Gaba, D. M. (2011). Have we gone too far in translating ideas from aviation to patient safety? No. *British Medical Journal, 342.* doi:10.1136/bmj.c7310

Gaba, D. M., Howard, S. K., Fish, K. J., Smith, B. E., & Sowb, Y. A. (2001). Simulation-based training in Anesthesia Crisis Resource Management (ACRM): A decade of experience. *Simulation Gaming, 32,* 175–193. doi:10.1177/104687810103200206

Greenberg, C. C., Regenbogen, S. E., Studdert, D. M., Lipsitz, S. R., Rogers, S. O., Zinner, M. J., & Gawande, A. A. (2007). Patterns of communication breakdowns resulting in injury to surgical patients. *Journal of the American College of Surgeons, 204,* 533–540. doi:10.1016/j.jamcollsurg.2007.01.010

Healey, A. N., Undre, S., & Vincent, C. A. (2004). Developing observational measures of performance in surgical teams. *Quality and Safety in Health Care, 13,* 33–40. doi:10.1136/qshc.2004.009936

Helmreich, R. L., & Schaefer, H-G. (1994). Team performance in the operating room. In M. S. Bogner (Ed.), *Human error in medicine* (pp. 225–253). Hillsdale, NJ: Lawrence Erlbaum.

Hudson, P. (2003). Applying the lessons of high risk industries to health care. *Quality and Safety in Health Care, 12,* i7–i12. doi:10.1136/qhc.12.suppl_1.i7

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

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management, 20,* 403–437. doi:10.1177/014920639402000206

Kolbe, M., Burtscher, M. J., Manser, T., Künzle, B., & Grote, G. (2011). The role of coordination for preventing harm in healthcare groups: Research examples from anaesthesia and an integrated model of coordination for action teams in healthcare. In M. Boos, M. Kolbe, P. Kappeler, & T. Ellwart (Eds.), *Coordination in human and primate groups* (pp. 75–92). Heidelberg: Springer. doi:10.1007/978-3-642-15355-6_5

Kolbe, M., Künzle, B., Manser, T., Zala-Mezö, E., Wacker, J., & Grote, G. (2009). Measuring coordination behaviour in anaesthesia teams during induction of general anaesthetics. In R. Flin & L. Mitchell (Eds.), *Safer surgery: Analysing behaviour in the operating theatre* (pp. 202–220). Aldershot: Ashgate.

Kolbe, M., Künzle, B., Zala-Mezö, E., Burtscher, M. J., Wacker, J., Spahn, D. R., & Grote, G. (2010). *The functions of team monitoring and 'talking to the room' for performance in anesthesia teams.* Paper presented at the 54th Annual Meeting of the Human Factors and Ergonomics Society, San Francisco, CA.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174. doi:10.2307/2529310

Langfred, C. W. (2004). Too much of a good thing? Negative effects of high trust and individual autonomy in self-managing teams. *The Academy of Management Journal, 47,* 385–399. doi:10.2307/20159588

LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., & Saul, J. R. (2008). A meta-analysis of teamwork processes: Tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology, 61* 273–307. doi:10.1111/j.1744-6570.2008.00114.x

Lim, B-C., & Klein, K.-J. (2006). Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior, 27* 403–418. doi:10.1002/job.387

Manser, T. (2009). Teamwork and patient safety in dynamic domains of healthcare: A review of the literature. *Acta Anaesthesiologica Scandinavica, 53,* 143–151. doi:10.1111/j.1399-6576.2008.01717.x

Manser, T., Harrison, T. K., Gaba, D. M., & Howard, S. K. (2009). Coordination patterns related to high clinical performance in a simulated anesthetic crisis. *Anesthesia and Analgesia, 108,* 1606–1615. doi:10.1213/ane.0b013e3181981d36

Manser, T., Howard, S. K., & Gaba, D. M. (2008). Adaptive coordination in cardiac anaesthesia: A study of situational changes in coordination patterns using a new observation system. *Ergonomics, 51,* 1153–1178. doi:10.1080/00140130801961919

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,* 356–376. doi:10.2307/259182

Marks, M. A., & Panzer, F. J. (2004). The influence of team monitoring on team processes and performance. *Human Performance, 17,* 25–41. doi:10.1207/S15327043HUP1701_2

Marks, M. A., Sabella, M. J., Burke, C. S., & Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *Journal of Applied Psychology, 87,* 3–13. doi:10.1037//0021-9010.87.1.3

Marks, M. A., Zaccaro, S. J., & Mathieu, J. E. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology, 85,* 971–986. doi:10.1037/0021-9010.85.6.971

Marsch, S. C., Muller, C., Marquardt, K., Conrad, G., Tschan, F., & Hunziker, P. R. (2004). Human factors affect the quality of cardiopulmonary resuscitation in simulated cardiac arrests. *Resuscitation, 60,* 51–56. doi:10.1016/j.resuscitation.2003.08.004

Mathieu, J. E., Heffner, T-S., Goodwin, G-F., Cannon-Bowers, J-A., & Salas, E. (2005). Scaling the quality of teammates' mental models: Equifinality and normative comparisons. *Journal of Organizational Behavior, 26,* 37–56. doi:10.1002/job.296

Mathieu, J. E., Maynard, M. T., Rapp, T., & Gilson, 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

Mazzocco, K., Petitti, D. B., Fong, K. T., Bonacum, D., Brookey, J., Graham, S., & Thomas, E. J. (2009). Surgical team behaviors and patient outcomes. *American Journal of Surgery, 197,* 678–685. doi:10.1016/j.amjsurg.2008.03.002

McGrath, J. E. (1984). *Groups: Interaction and performance.* Englewood Cliffs, NJ: Prentice-Hall.

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

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management, 36,* 876–910. doi:10.1177/0149206309356804

Mohammed, S., Klimoski, R., & Rentsch, J. R. (2000). The measurement of team mental models: We have no shared schema [Review]. *Organizational Research Methods, 3,* 123–165. doi:10.1177/109442810032001

Oken, A., Rasmussen, M. D., Slagle, J. M., Jain, S., Kuykendall, T., Ordonez, N., & Weinger, M. B. (2007). A facilitated survey instrument captures significantly more anesthesia events than does traditional voluntary event reporting. *Anesthesiology, 107,* 909–922. doi:10.1097/01.anes.0000291440.08068.21

Phipps, D., Meakin, G. H., Beatty, P. C., Nsoedo, C., & Parker, D. (2008). Human factors in anaesthetic practice: Insights from a task analysis. *British Journal of Anaesthesia, 100,* 333–343. doi:10.1093/bja/aem392

Porter, C. O. L. H., Gogus, C. I., & Yu, R. C.-F. (2010). When does teamwork translate into improved team performance? A resource allocation perspective. *Small Group Research, 41,* 221–248. doi:10.1177/1046496409356319

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31,* 437–448. doi:10.3102/10769986031004437

Rall, M., & Gaba, D. M. (2005). Human performance and patient safety. In R. D. Miller (Ed.), *Miller's anesthesia* (6 ed., pp. 3021–3072). Philadelphia, PA: Churchill Livingstone.

Reader, T. W., Flin, R., & Cuthbertson, B. H. (2007). Communication skills and error in the intensive care unit. *Current Opinion and Critical Care, 13,* 732–736. doi:10.1097/MCC.0b013e3282f1bb0e

Rico, R., Sánchez-Manzanares, M., Gil, F., & Gibson, C. (2008). Team implicit coordination processes: A team knowledge-based approach. *Academy of Management Review, 33,* 163–184. doi:10.5465/AMR.2008.27751276

Risser, D. T., Rice, M. M., Salisbury, M. L., Simon, R., Jay, G. D., & Berns, S. D. (1999). The potential for improved teamwork to reduce medical errors in the emergency department. The MedTeams Research Consortium. *Annals of Emergency Medicine, 34,* 373–383. doi:10.1016/S0196-0644(99)70134-4

Rousseau, V., Aube, C., & Savoie, A. (2006). Teamwork behaviors: A review and an integration of frameworks. *Small Group Research, 37,* 540–570. doi:10.1177/1046496406293125

Salas, E., Rosen, M. A., & King, H. (2007). Managing teams managing crises: Principles of teamwork to improve patient safety in the Emergency Room and beyond. *Theoretical Issues in Ergonomics Science, 8,* 381–394. doi:10.1080/14639220701317764

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "Big Five" in teamwork? *Small Group Research, 36,* 555–599. doi:10.1177/1046496405277134

Shapiro, M. J., Morey, J. C., Small, S. D., Langford, V., Kaylor, C. J., Jagminas, L., & Jay, G. D. (2004). Simulation based teamwork training for emergency department staff: Does it improve clinical team performance when added to an existing didactic teamwork curriculum? *Quality and Safety in Health Care, 13,* 417–421. doi:10.1136/qshc.2003.005447

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428. doi:10.1037/0033-2909.86.2.420

Smith-Jentsch, K. A. (2009). Measuring team-related cognition: The devil is in the details. In E. Salas, G.-F. Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 491–508). New York, NY: Psychology Press.

Smith-Jentsch, K. A., Cannon-Bowers, J. A., Tannenbaum, S. I., & Salas, E. (2008). Guided team self-correction: Impacts on team mental models, processes, and effectiveness. *Small Group Research, 39,* 303–327. doi:10.1177/1046496408317794

Smith-Jentsch, K. A., Mathieu, J. E., & Kraiger, K. (2005). Investigating linear and interactive effects of shared mental models on safety and efficiency in a field setting. *Journal of Applied Psychology, 90,* 523–535. doi:10.1037/0021-9010.90.3.523

Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors, 41,* 61–71. doi:10.1518/001872099779577273

Sundstrom, E., McIntyre, M., Halfhill, T., & Richards, H. (2000). Work groups: From the Hawthorne Studies to the work teams of the 1990s. *Group Dynamics: Theory, Research, and Practice, 4,* 44–67. doi:10.1037/1089-2699.4.1.44

Thomas, E. J., Sexton, J. B., & Helmreich, R. L. (2004). Translating teamwork behaviours from aviation to heathcare: Developement of behavioural markers for neonatal resuciation. *Quality and Safety in Health Care, 13,* 57–64. doi:10.1136/qshc.2004.009811

Tschan, F., Semmer, N. K., Gautschi, D., Hunziker, P., Spychiger, M., & Marsch, S. U. (2006). Leading to recovery: Group activities in medical emergency driven groups. *Human Performance, 19,* 277–304. doi:10.1207/s15327043hup1903_5

Tschan, F., Semmer, N. K., Gurtner, A., Bizzari, L., Spychiger, M., Breuer, M., & Marsch, S. U. (2009). Explicit reasoning, confirmation bias, and illusory transactive memory: A simulation study of group medical decision making. *Small Group Research, 40,* 271–300. doi:10.1177/1046496409332928

Undre, S., Healey, A. N., Darzi, A., & Vincent, C. A. (2006). Observational assessment of surgical teamwork: A feasibility study. *World Journal of Surgery, 30,* 1774–1783. doi:10.1007/s00268-005-0488-9

Waller, M., & Uitdewilligen, S. (2008). Talking to the room. Collective sense making during crisis situations. In R. Roe & M. Waller (Eds.), *Time in organizational research* (pp. 186–203). Oxford: Routledge.

Weinger, M. B., & Slagle, J. (2002). Human factors research in anesthesia patient safety: Techniques to elucidate factors affecting clinical task performance and decision making. *Journal of the American Medical Informatics Association, 9*(Suppl), 58–63. doi:10.1197/jamia.M1229

Weinger, M. B., Slagle, J., Jain, S., & Ordonez, N. (2003). Retrospective data collection and analytical techniques for patient safety studies. *Journal of Biomedical Informatics, 36,* 106–119. doi:10.1016/j.jbi.2003.08.002

Williams, A. L., Lasky, R. E., Dannemiller, J. L., Andrei, A. M., & Thomas, E. J. (2010). Teamwork behaviours and errors during neonatal resuscitation. *Quality and Safety in Health Care, 19,* 60–64. doi:10.1136/qshc.2007.025320

Wilson, K. A., Burke, C. S., Priest, H. A., & Salas, E. (2005). Promoting health care safety through training high reliability teams. *Quality and Safety in Health Care, 14,* 303–309. doi:10.1136/qshc.2004.010090

Received September 1, 2010
Revision received May 23, 2011
Accepted June 1, 2011 ■