Group Dynamics: Theory, Research, and Practice
2010, Vol. 14, No. 4, 332–349

1089-2699/10/$12.00 DOI: 10.1037/a0018822

Mental Model Metrics and Team Adaptability: A Multi-Facet Multi-Method Examination

Christian J. Resick
Drexel University

Toshio Murase, Wendy L. Bedwell, Elizabeth Sanz, Miliani Jiménez, and Leslie A. DeChurch
University of Central Florida

This paper empirically examines the convergent, discriminant, and predictive validity of three team mental model measurement approaches. Specifically, this study measures the similarity (MM-similarity) and quality (MM-quality) facets of team strategy-focused mental models using structural networks, priority rankings, and importance ratings. The convergent and divergent relationships among the three mental model metrics are then examined via a multi-facet multi-method matrix. Finally, the relative utility of each metric for understanding the relationships between team mental models, team adaptability, and decision effectiveness are compared. The study was conducted in a laboratory setting, modeling 56 four-person decision-making teams. Results indicate little convergent and extensive discriminant validity across the three mental model metrics. In addition, only mental models measured using the structural networks metric were found to have predictive validity in relation to team adaptation and performance. The quality and similarity of team structural networks were found to have interactive effects in relation to adaptation such that mental model quality was most strongly related to adaptation for teams with low mental model similarity and unrelated to adaptation for teams with high similarity. In turn, adaptation was critical for team decision effectiveness.

Keywords: teams, cognition, team mental models, team adaptability

Knowledge work is increasingly becoming team-based as the scope and complexity of problems require experts with distinct specialties to bridge their knowledge and experience to generate quality solutions. Two specific types of teams that frequently engage in knowledge work within organizations are management teams and project teams. One primary activity of these teams is the formation and enactment of strategies to accomplish collective goals, ranging from how a specific project should be accomplished to how an organization should be run. Furthermore, knowledge teams often need to make decisions and solve problems in dynamic environments (Marks, Zaccaro, & Mathieu, 2000; Thomas-Hunt & Phillips, 2003), requiring teams to adjust their strategies in response to changing environmental demands to perform successfully (Burns & Stalker, 1961). Team cognitive architecture has been consistently found to be a key determinant of team performance (DeChurch & Mesmer-Magnus, 2010), and is particularly important for enabling adaptive team performance (e.g., Burke, Stagl, Salas, Pierce, & Kendall, 2006).

Although abundant theoretical propositions and empirical studies underscore the importance of team mental models for team performance, progress has been limited by the lack of a generally accepted team mental model

Christian J. Resick, Management Department, LeBow College of Business, Drexel University; Toshio Murase, Wendy L. Bedwell, Elizabeth Sanz, Miliani Jiménez, and Leslie A. DeChurch, Department of Psychology, University of Central Florida.

We thank Shawn Burke, Jay Goodwin, and Joan Rentsch for helpful comments on earlier versions of this manuscript. This work was supported in part by funding from the Army Research Institute for the Behavioral and Social Sciences (Contract No. W91WAW-08-C-0028). The views expressed in this work are those of the authors and do not necessarily reflect official views of the U.S. Army.

Correspondence concerning this article should be addressed to Christian J. Resick, Drexel University, Management Department, LeBow College of Business, Philadelphia, PA 19104. E-mail: cresick@drexel.edu

metric (Mohammed, Ferzandi, & Hamilton, 2010; Mohammed, Klimoski, & Rentsch, 2000). Published studies have employed a variety of different methods for capturing mental models, making it difficult to cumulate findings in this area. In response, researchers have called for direct empirical comparisons of mental model measurement techniques to determine whether they measure the same underlying aspects of cognition (e.g., Smith-Jentsch, 2009). The current study addresses two critical questions, the answers to which are important to advancing knowledge on the role of mental models in team performance. First, to what extent do different metrics for capturing and representing team mental models evidence convergent or discriminant construct validity? Drawing upon work by Campbell and Fiske (1959), we create a multi-facet multi-method matrix to compare convergent and discriminant relationships among three common metrics. Second, what is the relative predictive utility of team mental metrics? Using each metric separately, we investigate the relationships between team mental models, team adaptability and, ultimately, team effectiveness.

Team Mental Models

Mental models are cognitive representations that individuals form regarding how the systems they interact with operate (e.g., Rouse & Morris, 1986). Applied to the team level, team mental models reflect a shared understanding among team members of particular aspects of their work environment, most commonly focusing on tasks or interactions among teammates (Cannon-Bowers & Salas, 2001; Cannon-Bowers, Salas, & Converse, 1993; Mohammed et al., 2000). Two important facets of team mental models are *similarity* between team members' mental models (i.e., shared mental models or MM-similarity) and *accuracy* (i.e., quality or MM-quality) of team mental models. MM-similarity reflects the extent to which team members hold similar cognitive representations of their performance context and goals. MM-quality reflects the extent to which team members' mental models adequately represent their performance context and goals (Edwards, Day, Arthur, & Bell, 2006; Smith-Jentsch, Campbell, Milanovich, & Reynolds, 2001). Further, researchers have argued the importance of examining both the main and interactive effects of MM-similarity and MM-quality (e.g., Mathieu, Heffner, Goodwin, Cannon-Bowers, & Salas, 2005).

Knowledge teams are charged with making decisions and developing strategies that drive team performance. As such, mental representations of key strategic decision points, relationships among these key decisions, and, finally, implications of these decision points are likely the most important mental model content for knowledge teams. The similarity and quality of member's knowledge sets ought to provide a platform for effective decision-making and strategy execution (Kellermanns, Walter, Lechner, & Floyd, 2005; Knight et al., 1999).

Mental Model Metrics

Although numerous team mental model measurement methods exist, a generally accepted approach has yet to emerge (Langan-Fox, Code, & Langfield-Smith, 2000; Mohammed et al., 2010). Mohammed and colleagues (2000) highlight four underlying features of mental model measurement techniques: content, elicitation of content, mental model structure, and representation of emergence among members. Content describes the focus of the mental model (e.g., tasks, strategies, team interactions). Elicitation involves measuring (or eliciting) the understanding of the content from team members. Mental model structure describes the modeling of the cognitive organization of the content. Finally, representation of emergence refers to the approach used to represent the team-level mental model.

In the current study, we measure strategy mental models using three common metrics (i.e., structural networks, priority rankings, and importance ratings), compare the convergence and divergence among the three metrics, and examine the utility of each metric for understanding team success. In each approach, we keep the mental model content consistent, representing the key decisions that need to be considered to achieve the collective goal. We now describe these approaches in terms of elicitation, structure, and representation of emergence.

Structural networks. We define the structural networks metric as the network of relationships among key decisions associated with achieving the collective goal. Structural networks can be elicited using any number of techniques that involve comparing decisions and actions to one another. In the current study, we elicit structural networks using pairwise ratings. Team members rate the extent to which each key decision is related to each of the remaining key decisions regarding accomplishment of the team's goal. The structure of the mental model is analyzed using Pathfinder, a network scaling algorithm that calculates a network of relationships among concepts known as a Pathfinder Network or PFnet. Emergence is represented using the Pathfinder *C* metric (i.e., metric of closeness) that calculates the degree of similarity between two PFnets. This approach has been widely used in team mental model studies (e.g., Edwards et al., 2006; Marks, Sabella, Burke, & Zaccaro, 2002; Stout, Cannon-Bowers, Salas, & Milanovich, 1999).

Priority rankings. We define the priority rankings metric as the relative importance rankings among key decisions associated with achieving a collective goal. In the current study we elicit priority rankings by asking team members to rank key decisions from highest to lowest priority for accomplishing the team's goal. The structure of priority rankings is the rank order of key decisions. Emergence is represented by the rank order correlations among team members' responses. Several prior studies have used ranking metrics of team mental models (e.g., Reger, 1990; Shobe, Fiore, & Carr, 2004; Smith-Jackson & Wogalter, 2007).

Importance ratings. We define the importance ratings metric as the overall importance of key decisions associated with achieving a collective goal. In the current study, we elicit importance ratings by asking team members to rate the importance of each key decision independent of other key decisions in achieving the team's collective goal. In contrast to network structures and priority rankings, importance ratings do not capture or reflect the organization or arrangement of information. The representation of emergence is the correlation among the team members' ratings. Ratings have been used to examine team mental models in several prior team mental model studies (e.g., Smith-Jentsch, Cannon-Bowers, Tannenbaum, & Salas, 2008; Webber, Chen, Payne, Marsh, & Zaccaro, 2000).

While each of these approaches has been used in prior team mental model research, little is known about the comparability of these three metrics. Therefore, in the current study we seek to examine the following question.

Research question 1. To what extent do different strategy-focused mental model metrics evidence convergent or divergent construct validity?

Team Adaptation

Although team mental models have been shown to be important performance drivers across studies examining varying types of teams operating in a variety of conditions (DeChurch & Mesmer-Magnus, 2010), theory posits that mental models are especially critical in dynamic contexts by enabling teams to be adaptive. Team adaptation is defined as "a change in team performance, in response to a salient cue or cue stream, that leads to a functional outcome for the entire team" (Burke et al., 2006, p. 1190). This adaptive theme was clear in the earliest formulation of team mental models by Cannon-Bowers and Salas (1990), in subsequent empirical investigations by Marks, Zaccaro, and Mathieu (2000), and elaborated on further in very recent theoretical work on team adaptation (Burke et al., 2006) and on implicit coordination (Rico, Sanchez-Manzanarez, Gil, & Gibson, 2008). The core logic is that when team members share a common understanding of the task, they are better able to anticipate one another's needs, provide useful information, and coordinate their inputs.

Given this core tenet of team mental model research—that mental models are essential enablers of adaptive team performance—it is particularly important to examine predictive validity in the context of knowledge teams whose task requires adaptation. Knowledge teams face unforeseen environmental events ranging from work process changes (e.g., downsizing requiring work effort consolidation) to disruptive events (e.g., order confusion resulting in the delayed delivery of key operating components) to more large-scale unforeseen crises (e.g., a hurricane causing major damage to a division operating within a specific region). When teams face disruptive events, factors critical to team success change (Thomas-Hunt & Phillips, 2003) and established modes of operation may no longer facilitate team success (e.g., LePine, 2003). Teams must adapt to environmental demands by developing new strategies and adjusting operating processes (Marks et al., 2000). Therefore, we seek to examine the second research question.

Research question 2. To what extent do different strategy-focused mental model metrics evidence similar predictive validity for understanding team adaptation and success?

Method

Participants included 224 students who formed 56 four-person teams. Participants were drawn from the undergraduate psychology research pool at a large southeastern university and received credit for their participation. The majority of participants were female (62.1%) and the average participant age was 20 years ($SD = 2.8$). Participants identified with a diverse range of ethnic backgrounds including Hispanic (68.8%), Black/African American (11.2%), Caucasian (10.3%), Asian (3.6%), Middle Eastern (0.9%), Pacific Islander (0.4%), and Other/Not Specified (4.8%).

Task

Teams performed a decision-making simulation based on the pc-game SimCity4 Deluxe Edition (EA Games, 2004). The game was displayed via an on-screen map with interfaces that displayed information, such as the population, funding to various city department, tax rates, and resident opinion polls. Participants used a keyboard and mouse to make changes in the city; types of changes that could be implemented included zoning or rezoning areas of the city into commercial, residential, and industrial zones; reallocating the funding to various departments; setting the tax rate; and constructing buildings, utility plants, bridges and other structures. The game is programmed by the developer so that all changes ultimately affect the desirability of the city. This task was selected as it requires a variety of information to be gathered and processed effectively for the city to thrive.

We developed a cross-functional team decision-making simulation task based on this game by first conducting a task analysis of the game. The task analysis was informed by (a) reviewing the SimCity guide book (Kramer, 2003), (b) playing the game, and (c) interviewing individuals with game experience. Similar tasks and activities were clustered together resulting in the following distinct roles: Financial Officer (maintaining the city budget), City Planner (zoning land and managing transportation systems), Public Works Officer (managing public utilities and public safety), and Social Welfare Officer (managing educational and public health services). The resulting team was structured so that each member had a unique area of responsibility and was provided with specialized knowledge about that area. Further, this combination of roles was needed to effectively manage the city, ensuring that expertise was distributed and that members needed to depend upon one another to accomplish objectives. Teams were placed as the city council of the simulated city of Pantherville and were responsible for making and implementing decisions regarding the management of all aspects of city life, including urban design, funding public agencies, taxation, and so forth. Teams were self-managed and there was no specified hierarchy among roles. Teams were charged with the goal of making Pantherville a desirable place to live and work, and thus to increase city population.

Procedure

Research sessions lasted approximately four hours. Upon arrival, participants completed demographic measures and were then informed they were recently appointed to the city council of Pantherville for a simulated 3-year term in office. Participants then completed two computer-based training (CBT) modules. The first was identical across the four roles, providing an introduction to SimCity (general features of the game, major decisions to be made, and the location of information about the city's status). Participants then completed a role-specific CBT module that outlined the responsibilities of their respective role, and provided instructions on (a) how to use specific functions of the game, (b) how to retrieve and monitor information, and (c) the social and economic impact of various strategic decisions on their specific areas of responsibility. Participants were provided a handout to use during the remainder of the session that contained role-specific information covered during training. To assess participants' learning of the training content, immediately after the participant completed the CBT, an experimenter asked the participant to demonstrate a series of tasks that were covered in training. If a participant was unable to complete a given task, the experimenter demonstrated how to complete the task and then asked the participant to demonstrate it again. All participants were able to correctly demonstrate all tasks by the second trial. The purpose of this procedure was to ensure that all participants had acquired a basic level of knowledge necessary for the simulation. Training lasted approximately one hour.

Following training, participants were given a five-minute break and then reconvened in a conference room where they were seated at a round table with the simulation presented on a 32-inch monitor. A 56 in. × 36 in. color map of Pantherville posted on the wall identified important buildings, such as power plants, hospitals, schools, police stations, and fire stations. An experimenter informed teams that their goal was to make the city as desirable a place to live and work as possible. The experimenter explained that the game is programmed such that decisions that increased city desirability resulted in a corresponding increase in city population. Likewise, governing decisions that decreased city desirability resulted in a decrease in city population. Teams were informed that they could monitor their progress toward goals by examining the city's population.

Teams were then provided an initial period of 15 minutes to examine Pantherville (e.g., determine the funding levels, population, tax rate, overall layout of the city), identify any problems or areas requiring changes, and make decisions. During this time, the simulation was paused; however, participants still had the ability to maneuver through the city without making any changes. Upon completion of the 15-min planning and discussion period, teams were allowed to implement their plans and make changes to the city for eight minutes. The simulation was then started and allowed to progress for six simulated months. Teams then completed two more planning and decision-making cycles in which (a) the simulation was paused, (b) teams were given three minutes to review the city and make plans, (c) teams were given five minutes to implement plans, and (d) the simulation was started and progressed for six simulated months.

At the conclusion of the third decision-making cycle, the experimenter introduced the disruptive event by automatically switching the monitor to a video of a simulated newscast. An actor playing the role of a news anchorwoman provided an alert that an earthquake had struck the city. The anchorwoman summarized the damage and displayed footage of an earthquake. The video lasted approximately two minutes. During this time, an experimenter loaded a second version of Pantherville from a remote location. The experimenter prompted the team to think about their plans for restoring the city and accomplishing their overall goal. This approach created a baseline common to all teams from which to examine the relationship between mental models and team performance. In real world teams, the capacity to adapt is clearly impacted by prior success making adaptation easier for high performing teams and more difficult for their lower performing counterparts. However, since the aim of the study was to examine the relationships between cognition and adaptive performance, it was necessary to hold the task conditions constant across teams in the sample. This phase required teams to determine the extent of damage to city infrastructure and prioritize post-disaster actions. Teams needed to adapt strategies by (a) recognizing damage to critical areas, (b) reprioritizing issues, and (c) identifying and implementing new strategies.

The teams then took over the post-disaster version of Pantherville. Each team was provided seven minutes to review the city and plan changes and three minutes to implement those changes. The simulation was then progressed for another six simulated months. The game was paused, and each team was given three minutes to plan changes, followed by two minutes to implement those changes before the city was once again allowed to progress six simulated months and then paused. Next, teams were asked to complete three mental model measures. Mental models were measured at this point in the study to ensure that enough time elapsed for similarity and quality to emerge based on repeated discussion and planning sessions and members' common set of observations about the implications of their decisions. At this time, teams were not aware of their final performance level. When all members completed the measures, teams were given another three minutes to plan changes and two minutes to implement those plans. The simulation was then started again, and allowed to progress for a final six simulated months. In total, each team completed three pre-disaster and three post-disaster decision-making cycles.

Measures

Control variable—game experience. Participants were asked two questions regarding game experience: "How frequently do you play SimCity?" and "How frequently do you play PC-video games?" Responses to both questions were measured on a scale with the following options *1 = only once or twice in the last 5 years, 2 = a few times per year, 3 = a few times per month, 4 = a few times per week,* and *5 = daily*. We suggest that these two variables are complementary and that as a set provide a useful operationalization of the game experience construct in the current study. We average together each participant's responses to the two items to represent individual game experience. Then, we calculated the team mean-level of game experience and used this variable in subsequent analyses.

Control variable—pre-disaster performance. Team success during pre-disaster decision-making intervals may affect the team's ability or willingness to adapt its strategic decision-making, as well as the similarity and quality of team mental models. As such, we controlled for pre-disaster performance as indicated by the city population after the first three decision-making cycles (i.e., 18 simulated months). The average pre-disaster population score was 51,172 (SD = 8,002).

Mental models. The three mental model elicitation measures were distributed to participants after the fifth decision-making cycle. Mental model content was identical across all three measures and included 10 key decisions (refer to Appendix A). We identified these decisions by conducting a task analysis of the decision options associated with improving city desirability.

Structural networks. To elicit mental models representing structural networks, we presented participants with a matrix comparing each of the 10 key strategic decisions to one another. Participants read each pair of strategic decisions and rated relationships on a seven point scale ranging from *1 = totally unrelated* to *7 = strongly related.* In all, each team member made a total of 45 ratings, which represent their structured knowledge regarding how governing decisions are related to one another in achieving the team's goal. Mental model structure was represented using PFnets (Schvaneveldt, 1990). The Pathfinder algorithm represents the geometric distance between each pair of concepts. Concepts rated as being more highly related by participants are separated by fewer links and are closer in proximity, while concepts rated as being unrelated to one another are separated by more links and represented by a greater distance.

Team MM-similarity was calculated using Pathfinder's metric of closeness (C), which calculates the degree of similarity between two PFnets. For each team, six C scores were calculated by comparing members' PFnets to one another. Two networks without any common links would have $C = 0$ and two networks with identical network structures would have $C = 1$. The six scores were then averaged together to create the team's MM-similarity score. To calculate MM-quality, we constructed an expert mental model; three members of the research team who were highly familiar with the SimCity game and the strategies for success in the post-disaster city independently completed the mental model questionnaire. The experts then met to review their ratings, discuss differences, and incorporate changes. The experts generally agreed on most of the relationships, and where disagreements were present, they were easily resolved through discussion. The experts then constructed a matrix representing their consensus view of the strategic relationships. For each team, four C scores were then calculated by comparing each member's PFnet to the expert PFnet. The fours scores were then averaged together to create a team MM-quality score.

Priority rankings. To elicit mental models representing priority rankings, we presented participants the list of 10 key strategic decisions and asked them to rank these decisions based on their priority of importance in helping the team to achieve its goals. Team MM-similarity was calculated by computing six Spearman rank-order correlations, comparing each team member's rankings to the rankings of each of the other team members. The six correlations were then averaged to obtain a team MM-similarity score. For MM-quality, we again constructed an expert mental model based on rankings of the same three SMEs who ranked the key decisions, reviewed their rankings, and reached consensus. Four rank-order correlations were calculated by comparing each team member's rankings to the expert rankings. Team MM-quality is the average of the four rank-order correlations.

Importance ratings. To elicit mental models representing importance ratings, we presented participants with the same list of 10 important strategic decisions and asked them to indicate the importance of each decision independently of the other decisions for achieving the team's goal using a 5-point rating scale ranging from *1 = not at all important* to *5 = very important.* Team MM-similarity was calculated by correlating each team members' ratings with each of the other team members. The correlations were then averaged together to obtain the team MM-similarity score. For MM-quality, we constructed a final expert mental model based on the ratings of the key decisions by the SMEs who independently rated the key decisions, reviewed their ratings, and reached consensus. Four correlations were calculated by comparing each team member's ratings to the SME rankings. Team MM-quality is the average of the four correlations.

Team adaptation. The pre-disaster and post-disaster stages were designed to require the team to make and implement different strategic decisions to achieve the goal of making Pantherville a desirable place to live (and thus increasing population). In the pre-disaster stage, teams needed to focus on growth by reducing taxes and funding public services. However, in the post-disaster stage, teams needed to quickly recognize key areas of city infrastructure that were damaged and restore operations *before* focusing on growth to avoid a mass exodus from the city. We conducted a detailed analysis of the impact of strategic decisions on the post-disaster city and determined that three decisions needed to be implemented to restore city operation or the city experienced steep population declines. It is important to note that these three strategies were not critical to success during the pre-disaster decision-making cycles so their utilization during the post-disaster decision-making cycles represented strategy adaptation. The three decisions included (a) building new power plants, (b) clearing damaged land, and (c) building new hospitals. During each session, two trained observers used a stopwatch and recorded, in 15-s increments, how long it took the team to implement each of the three critical decisions. These three scores were then averaged to represent the amount of time elapsed before the team addressed the three critical needs. The mean adaptation score across teams was 25.67 seconds (SD = 5.02). It is important to note that adaptation was operationalized as the amount of elapsed time before the team began to address key city needs, lower adaptation scores represent better adaptation.

Team decision effectiveness. Team decision effectiveness was operationalized as Pantherville population at the conclusion of the third post-disaster decision-making cycle. The average post-disaster population in the current sample was 24,883 (SD = 14,333).

Analytical Approach

To address Research Question 1, we calculated zero-order correlations to examine evidence of convergent and discriminant validity among the similarity and quality facets of strategy-focused mental models measured using the three metrics. To address Research Question 2, we used hierarchical multiple regression. First, the MM-similarity and MM-accuracy variables were centered, and then a MM-similarity $\times$ MM-accuracy interaction term was calculated for each of the three metrics. We entered control variables in Step 1, followed by the centered main effect variables in Step 2 and the centered interaction term in Step 3.

Results

Convergent & Discriminant Validity

Means, standard deviations and zero-order correlations among variables are presented in Table 1. Next, consistent with Campbell and Fiske's (1959) multi-trait, multi-method approach, we created a multi-facet, multi-method matrix (summarized in Table 2) to compare correlations among the two mental model facets (i.e., similarity and quality) and three measurement methods (i.e., structural networks, priority rankings, and importance ratings). For convergent validity, it is necessary to look at correlations among the same trait (or facet in this case),

Table 1
Mean Scores, Standard Deviations, and Zero-Order Correlations Among Variables

Variable	1	2	3	4	5	6	7	8	9	10
Game experience	—									
Pre-disaster decision effectiveness	.25†	—								
MM-similarity-priority rankings	−.19	−.18	—							
MM-quality-priority rankings	−.01	−.23	.63**	—						
MM-similarity-importance ratings	−.06	−.07	.02	−.07	—					
MM-quality-importance ratings	.04	.04	.20	.37*	.62**	—				
MM-similarity-structural networks	.03	.04	.42**	.08	.23†	.23†	—			
MM-quality-structural networks	.01	−.14	.08	−.05	.16	.09	.37**	—		
Team adaptation	−.42**	−.23†	−.06	−.06	.01	−.14	−.25†	−.31*	—	
Decision effectiveness	.39**	.10	.05	.01	.02	.18	.13	.33*	−.69**	—
Mean	1.46	51171.80	.42	.43	.49	.54	.27	.28	25.67	24883.04
SD	.46	8002.22	.27	.22	.20	.17	.08	.05	5.02	14332.88

Note. MM-similarity = mental model similarity; MM-quality = mental model quality.
† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$. Sample size ranges from 40 to 56.

as measured using different methods. Regarding MM-similarity, some evidence of convergence exists between structural networks and priority rankings (r = .42) and weaker evidence of convergence between structural networks and importance ratings (r = .23) but not between priority rankings and importance ratings (r = .02). Regarding MM-quality, the opposite pattern appeared; results indicate some degree of convergence between priority rankings and importance ratings (r = .37) but not between structural networks and priority rankings (r = −.05) or between structural networks and importance ratings (r = .09). Across both similarity and quality, the magnitude of the correlations was generally low to moderate in nature, suggesting limited evidence of convergent validity among the three metrics.

To determine discriminant validity, Campbell and Fiske (1959) suggest comparing (a) the hetero-trait, mono-method correlations, that is the relationships among different traits (or facets in this case) with the same measurement technique (i.e., comparing similarity and quality facets measured using structural networks), and (b) the hetero-trait, hetero-method correlations, the relationships among different traits (or facets) measured with different methods (i.e., similarity of structural networks with the quality of priority rankings). Consistent with prior research (e.g., Mathieu et al., 2005), we expect to find significant positive correlations of a moderate magnitude between the similarity and quality facets measured using the same metric. However, results indicate significant positive correlations of a moderate magnitude between

Table 2
Multi-Facet, Multi-Method Matrix of Shared Mental Model Metrics

Metric	Facet	1	2	3	4	5	6
Structural networks	1. Similarity						
	2. Quality	.37**					
Priority rankings	3. Similarity	.42**	.08				
	4. Quality	.08	−.05	.63**			
Importance ratings	5. Similarity	.23†	.16	.02	−.07		
	6. Quality	.23†	.09	.20	.37*	.62**	

† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$.

the similarity and quality facets only within the structural networks metrics (r = .37, $p \leq .01$). The magnitude of correlations within the priority rankings (r = .63, $p \leq .01$), and importance ratings (r = .62, $p \leq .01$) metrics were of a much stronger magnitude, indicating little discriminant validity between similarity and quality facets measured using the ranking and rating metrics.

Finally, further evidence of discriminant validity is provided if the correlations between different facets measured with different metrics are low and not statistically significant. Results provide evidence of discriminant validity as the hetero-facet hetero-method correlations were generally weak and nonsignificant. The one exception is the correlation between MM-similarity measured using structural networks and MM-quality measured using importance ratings (r = .23, $p \leq .10$).

In response to our first research question on the convergent and discriminant validity of mental model metrics, results indicate that in general, the different strategy-focused mental model metrics are not commensurate. While some evidence of convergent validity was found (e.g., MM-similarity measured using structural networks and priority rankings, r = .42), a general pattern of weak and nonsignificant correlations across measurement approaches was also found suggesting the three metrics may be measuring different underlying constructs.

Predictive Validity

Next we examined the predictive validity of the mental model metrics. We tested predictive validity against two criteria: team adaptation (a proximal indicator) and decision effectiveness (a distal indicator). Table 3 presents the results of regressing post-disaster decision effectiveness onto pre-disaster population and game experience (control variables) in step one followed by adaptation in step two. Results indicate that the block of control variables in step one accounted for a significant amount of variance in post-disaster performance (R^2 = .15, $p \leq .01$). The addition of adaptation in step two accounted for significant incremental variance in post-disaster performance (ΔR^2 = .34, $p \leq .01$), indicating that adaptation was explaining significant variance in team decision effectiveness.

Structural networks. Table 4 presents regression results examining relationships between team mental models measured via structural networks and team adaptation and decision effectiveness. The block of control variables in step one explained a moderate and significant amount of variance in adaptation (R^2 = .19, $p \leq$ .01) with only game experience significantly related to adaptation ($\beta = -.39$, $p \leq .01$). The addition of the main effects for MM-similarity and MM-quality at step two also explained a significant amount of incremental variance in adaptation (ΔR^2 = .12, $p \leq .05$). Examination of the regression coefficients indicated that MM-quality was significant related to adaptation ($\beta = -.28$, $p \leq .05$), while MM-similarity was not ($\beta = -.13$, *ns*). Finally, the addition of the MM-similarity × MM-quality interaction at step three explained a small but significant amount of incremental variance (ΔR^2 = .06, $p \leq .05$).

Next, the MM-similarity × MM-quality interaction was graphed, following procedures presented by Cohen, Cohen, West, and Aiken (2003). The interaction is depicted graphically

Table 3
Regression of Decision Effectiveness on Team Adaptation

Variable	R^2	F	ΔR^2	β	B	SE
Step 1						
Pre-disaster performance	.15	4.62**		.00	.01	.24
Game experience				.38	11886.22**	4048.72
Step 2						
Pre-disaster performance	.49	16.81**	.34	−.08	−.15	.19
Game experience				.13	4097.62	3417.91
Adaptation				−.65	−1116.66**	188.14

Note. N = 56. Lower adaptation scores represent faster adaptation.
† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$.

Table 4
Regression of Team Adaptation and Decision Effectiveness on Structural Networks of Team Mental Models

Variable	R^2	F	ΔR^2	β	B	SE
Adaptation						
Step 1	.19	6.26**				
Pre-disaster performance				−.13	.00	.00
Game experience				−.39	−6.98**	2.30
Step 2	.31	5.75**	.12			
Pre-disaster performance				−.17	.00	.00
Game experience				−.37	−6.66**	2.17
MM-similarity-structural networks				−.13	−14.65	14.04
MM-quality-structural networks				−.28	−47.70*	21.91
Step 3	.37	5.83**	.06			
Pre-disaster performance				−.18	.00	.00
Game experience				−.35	−6.24**	2.11
MM-similarity-structural networks				−.09	−9.55	13.78
MM-quality-structural networks				−.28	−47.75*	21.19
MM-quality × MM-similarity				.25	615.43*	288.77
Decision effectiveness						
Step 1	.15	4.62**				
Pre-disaster performance				.00	.00	.24
Game experience				.38**	11886.22**	4048.72
Step 2	.26	4.36**	.11			
Pre-disaster performance				.05	.10	.227
Game experience				.37	11361.72**	3865.27
MM-similarity-structural networks				−.01	−2567.16	24993.05
MM-quality-structural networks				.34	98837.18**	39010.75
Step 3	.26	3.57**	.01			
Pre-disaster performance				.06	.11	.23
Game experience				.36	11089.57	3900.46
MM-similarity-structural networks				−.03	−5821.63	25484.44
MM-quality-structural networks				.34	98868.05*	39188.02
MM-quality × MM-similarity				−.09	−392417.42	534160.72

Note. $N = 42$. Lower adaptation scores represent faster adaptation. MM = Mental Model.
† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$.

in Figure 1. Results indicate that for teams with highly similar mental models (i.e., 1 *SD* above the mean), MM-quality had virtually no relationship with team adaptation. However, for teams with low MM-similarity (i.e., 1 *SD* below the mean), MM-quality was strongly related to strategic adaptation (i.e., increases in quality were associated with faster adaptation time).

Next, we examined the relationships with decision effectiveness. In step one of the regression analysis, the block of control variables was significantly related to decision effectiveness ($R^2 = .15$, $p \leq .01$). Examination of the regression coefficients indicated that this relationship was largely due to the significant relationship with game experience ($\beta = .38$, $p \leq .01$). The addition of the MM-similarity and MM-quality main effects in step two explained a significant amount of incremental variance in decision effectiveness ($\Delta R^2 = .11$, $p \leq .01$). Similar to adaptation, MM-similarity was not related to decision effectiveness ($\beta = -.01$, *ns*), while MM-quality was significantly related to decision effectiveness ($\beta = .34$, $p \leq .01$). Finally, the addition of the MM-similarity × MM-quality interaction at step three did not explain a significant amount of incremental variance in decision effectiveness ($\Delta R^2 = .01$, *ns*).

Priority rankings. Table 5 presents regression results examining relationships between team mental models measured via priority rankings and team adaptation and decision effectiveness. In subsequent analyses, we will not repeat the relationships between the control variables and adaptation or decision effectiveness in step one of the regression analyses. The addition of

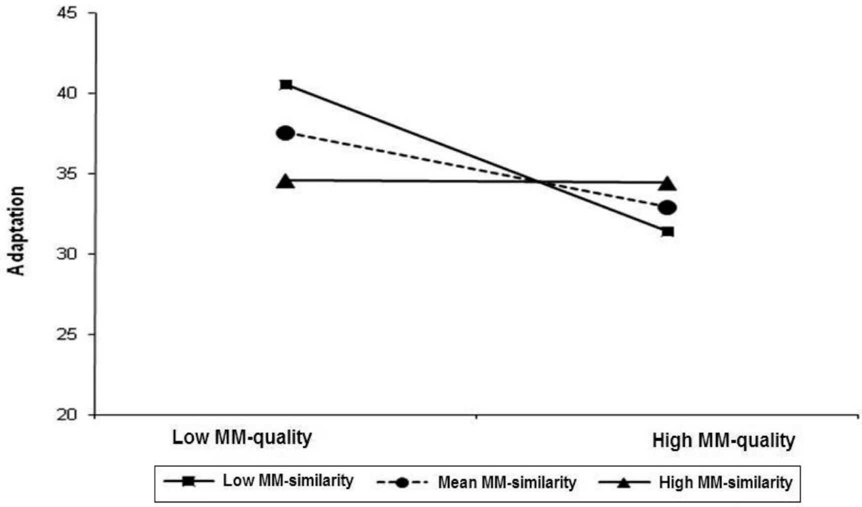

Figure 1. The interactive effects of MM-quality and MM-similarity in relation to team adaptation—structural networks metric.

the MM-similarity and MM-quality main effect variables in step two explained a small but non-significant amount of incremental variance in adaptation (R^2 = .03, *ns*). Neither MM-similarity (β = −.21, *ns*) nor MM-quality (β = .06, *ns*) was significantly related to adaptation. The addition of the MM-similarity × MM-quality interaction term in step three did not explain any additional incremental variance (R^2 = .00, *ns*).

Regarding decision effectiveness, the MM-similarity and MM-quality main effect variables entered at step two again did not explain a significant amount of incremental variance (ΔR^2 = .02, *ns*). Neither MM-similarity (β = .20, *ns*) nor MM-quality (β = −.12, *ns*) was significantly related to decision effectiveness. Finally, the addition of the MM-similarity × MM-quality interaction at step three did not explain a significant amount of additional incremental variance in decision effectiveness.

Importance ratings. Table 6 presents regression results examining relationships between team mental models measured via importance ratings and team adaptation and decision effectiveness. In step two of the regression analyses, the addition of the MM-similarity and MM-quality main effect variables explained a small but nonsignificant amount of incremental variance in adaptation (ΔR^2 = .02, *ns*). Neither MM-similarity (β = .08, *ns*) nor MM-quality (β = −.17, *ns*) was significantly related to adaptation. Similarly, the addition of the MM-similarity × MM-quality interaction at step 3 also explained a small but nonsignificant amount of incremental variance (ΔR^2 = .02, *ns*).

Regarding decision effectiveness, the addition of the MM-quality and MM-similarity main effect variables at step two did not explain a significant amount of variance (ΔR^2 = .03, *ns*). Neither MM-similarity (β = −.10, *ns*) nor MM-quality (β = .23, *ns*) was significantly related to decision effectiveness. Finally, the addition of the MM-similarity × MM-quality interaction at step three did not explain a significant amount of incremental variance in decision effectiveness.

In response to our second question regarding the predictive validity of mental model metrics, our results suggest strategic mental models measured using the structural networks approach were the only useful metric for understanding the relationships between strategy-focused team mental models and adaptive team decision effectiveness. The remaining metrics were unrelated to either adaptation or decision effectiveness.

Discussion

Cannon-Bowers and Salas's (1990) initially formulated the shared mental model construct

Table 5
Regression Analyses of Team Adaptation and Decision Effectiveness on Priority Rankings of Team Mental Model

Variable	R^2	F	ΔR^2	β	B	SE
Adaptation						
Step 1	.29	8.05**				
Pre-disaster performance				−.00	.00	.00
Game experience				−.54	−8.04**	2.10
Step 2	.32	4.39**	.03			
Pre-disaster performance				−.02	−.00	.00
Game experience				−.57	−8.54**	2.18
MM-similarity-priority rankings				−.21	−5.69	4.95
MM-quality-priority rankings				.06	1.82	5.88
Step 3	.32	3.45**	.00			
Pre-disaster performance				−.02	−.00	.00
Game experience				−.58	−8.62**	2.22
MM-similarity-priority rankings				−.18	−4.90	5.56
MM-quality-priority rankings				.05	1.62	5.98
MM-quality × MM-similarity				.05	6.48	19.80
Decision effectiveness						
Step 1	.15*	3.45*				
Pre-disaster performance				−.05	−.07	.23
Game experience				.40	10982.71*	4253.85
Step 2	.17	1.94	.02			
Pre-disaster performance				−.05	−.07	.24
Game experience				.44	12004.68**	4436.51
MM-similarity-priority rankings				.20	10146.64	10089.31
MM-quality-priority rankings				−.12	−7368.97	11971.44
Step 3	.17	1.51	.00			
Pre-disaster performance				−.05	−.07	.25
Game experience				.43	11969.65**	4520.02
MM-similarity-priority rankings				.21	10527.99	11350.69
MM-quality-priority rankings				−.12	−7466.15	12200.23
MM-quality × MM-similarity				.01	3129.65	40400.09

Note. $N = 42$. Lower adaptation scores represent faster adaptation. MM = Mental Model.
† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$.

to explain the seamless coordination they observed in expert military teams. They posited that such synchronization was the result of team members' common understanding of key elements of their performance environment. They went on to argue that this cognitive similarity enabled these teams to anticipate one another's needs, to interpret incoming information in a compatible manner, and to make decisions jointly. This conceptualization of the importance of team cognition sparked tremendous interest in both the applied and scientific communities who work with teams; several empirical insights have been gained into both the factors that facilitate the emergence of team mental models and the role of team mental models in team success (e.g., Edwards et al., 2006; Marks et al., 2002; Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000; Resick, Dickson, Mitchelson, Allison, & Clark, in press); numerous theoretical reviews have examined the state of this science (Klimoski & Mohammed, 1994; Mohammed, et al., 2010); and now two recent meta-analyses have empirically cumulated this research (DeChurch & Mesmer-Magnus, 2010; DeChurch & Mesmer-Magnus, in press). Despite these advances, the Achilles heel of the shared mental model construct is measurement (Mohammed et al., 2000; Rentsch, Small, & Hanges, 2008). Multiple approaches ranging from network-based structural techniques to perceptual attitude scales have been used, with little understanding of the comparability of these approaches for measuring the underlying construct. The current study moves the team mental model construct a step further

Table 6
Regression Analyses of Team Adaptation and Decision Effectiveness on Importance Ratings of Team Mental Models

Variables	R^2	F	ΔR^2	β	B	SE
Adaptation						
Step 1	.19	6.06**				
Pre-disaster performance				−.14	.00	.00
Game experience				−.38	−6.92**	2.35
Step 2	.21	3.25*	.02			
Pre-disaster performance				−.12	.00	.00
Game experience				−.38	−6.78**	2.37
MM-similarity-importance ratings				.08	3.31	6.89
MM-quality-importance ratings				−.17	−8.50	8.34
Step 3	.23	2.90*	.02			
Pre-disaster performance				−.13	.00	.00
Game experience				−.36	−6.43**	2.38
MM-similarity-importance ratings				.23	9.65	8.67
MM-quality-importance ratings				−.18	−9.17	8.33
MM-quality × MM-similarity				.21	41.99	35.13
Decision effectiveness						
Step 1	.15	4.56*				
Pre-disaster performance				−.01	−.01	.24
Game experience				.39	12054.84**	4111.29
Step 2	.19	2.77*	.03			
Pre-disaster performance				−.02	−.04	.24
Game experience				.38	11750.50**	4124.72
MM-similarity-importance ratings				−.10	−6935.36	11967.96
MM-quality-importance ratings				.23	19732.58	14494.02
Step 3	.19	2.25†	.01			
Pre-disaster performance				−.02	−.04	.24
Game experience				.37	11474.97**	4185.97
MM-similarity-importance ratings				−.17	−11994.446	15231.30
MM-quality-importance ratings				.23	20263.52	14632.03
MM-quality × SMM-similarity				−.10	−33547.59	61734.05

Note. $N = 54$. Lower adaptation scores represent faster adaptation. MM = Mental Model.
† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$.

by examining the convergent, discriminant, and relative predictive validity of three common elicitation metrics of the same content.

Convergent and Discriminant Validity

One important contribution of the current study concerns the convergent validity of mental model metrics. In general, the mono-facet, hetero-method correlations were small and nonsignificant suggesting little convergence among the three metrics. These findings provide evidence that different metrics measure different underlying constructs and raise concerns about the construct validity of the various metrics. Therefore, teams researchers need to carefully consider the approach used to measure team mental models in future studies to ensure that metrics have adequate construct validity.

Regarding discriminant validity, results indicate some evidence of discriminant validity between the similarity and quality mental model facets measured using the structural networks metric. However, strong correlations between the facets were found for priority rankings and importance ratings metrics (r = .63 and .62, respectively). These stronger correlations suggest that the priority rankings and importance ratings metrics do not adequately discriminate between the similarity and quality facets and, thus, do not provide adequate discriminant validity. In contrast, the evidence indicates that the structural networks metric does provide teams researchers with a viable metric for distinguishing between the similarly and quality facets.

Finally, the hetero-facet hetero-method correlations were generally low and not statistically significant, providing evidence of discriminant validity of the three metrics across mental model facets. However, one exception was found as the structural networks metric of similarity was moderately correlated with the importance ratings metrics of quality. This finding suggests some degree of overlap among these two metrics or facets.

Predictive Validity

Turning now to the predictive validity of the three mental model metrics, results suggest that measuring mental models using a structural networks approach provides a useful metric for understanding the relationships between team strategy-focused cognition, adaptation, and performance. Unexpectedly, neither the priority rankings nor the importance rating metrics were related to team adaptation or decision effectiveness. All three metrics capture team member cognition, though there are marked differences in the degree to which perceptions are then represented according to the underlying structure or organization of knowledge, versus at the other extreme, modeling solely the perceptions devoid of any structure (Rentsch et al., 2008). Network indices maximally capture the arrangement of knowledge; rankings capture some structure but less so than network indices; and ratings reflect perceptions without modeling the structure of those perceptions. Thus, one substantive explanation for this observed difference in predictive validity is that representing knowledge arrangement is a critical aspect of team mental models. This conclusion seems most plausible in teams resembling those in the current study, this is, teams with distributed expertise performing a knowledge-based task that requires strategy adaptation. The extent to which this conclusion holds in other team types remains unclear.

Results indicate that the quality of team member structural networks had a significant main effect relationship with both adaptation and decision effectiveness while the similarity of structural networks did not. That is, as team members formed higher quality structural understandings of the relationships among key decision alternatives and their implications for achieving the teams' goals, teams were able to adapt their strategies more efficiently and make decisions more quickly. At the same time, an interactive effective of similarity and quality in relation to adaptation was also found. Consistent with prior research (e.g., Marks et al., 2000), we examine the relationship between structural network quality and adaptation at high (i.e., 1 *SD* above the mean), average (mean), and low (i.e., 1 *SD* below the mean) levels of similarity. Findings indicate that structural network quality was most strongly related to adaptation when team members had less similar structural networks and virtually unrelated to adaptation for teams with highly similar structural networks. When teams had less similar structural networks, having a high quality understanding of strategic alternatives appeared to enable teams to more quickly identify and respond to critical needs. Alternatively, the similarity among members' structural networks appeared to enable teams with lower quality cognitive networks to adapt efficiently.

Given the nature of knowledge work, these results suggest that when there are multiple ways of attacking a problem, teams need to have high quality structural networks or highly similar structural networks, but not necessarily both. In addition, the importance of various team members' roles may change in response to the nature of a disruptive event. Perhaps individuals who hold critical information become more critical to team adaptability, and the quality of that person's (or persons') strategy mental model becomes a key driver of team success. Using averaging methods to operationalize team mental models places equal weight on all members' mental models (Smith-Jentsch, 2009). Future research could measure member status or ability and weight individual mental models by these factors to provide a perspective on team mental models that perhaps captures the degree of influence among individual members.

Taken together, the convergent, discriminant, and predictive validity evidence suggests that team mental models are best represented using structural network metrics. This finding is inline with DeChurch and Mesmer-Magnus' (in press) meta-analytical finding that structured approaches to assessing team cognition were more strongly related to objective indicators of team performance than perceptual measures, of which the priority ranking and importance ratings approaches could be considered.

The other elicitation techniques demonstrated little utility in understanding adaptation and team performance in the current study. Given the generally low levels of convergent validity among the three metrics, perhaps priority rankings and importance ratings are measuring different emergent cognitive states. Other explanations for the lack of predictive validity of priority rankings and importance ratings are certainly plausible. First, the relatively brief duration of time the teams worked together may not have been enough time for these metrics to affect adaptation or decision effectiveness. Second, perhaps features of the decision task rendered them less relevant than they would be on, for example, a multi-issue negotiation task where members have to tradeoff on particular issues. In the current task, there was a high level of goal congruence across team members, whereas on a multi-party negotiation task, the goal conflict within the team may render the similarity of priority rankings much more pivotal to decision effectiveness than structural networks.[1] As such, before ruling out the use of ranking and rating metrics, future research should examine these relationships across different types of teams, and over a longer duration of time.

Limitations

As with all studies, there are several limitations to the current study. A number of these limitations stem from the use of a laboratory setting. The simulation lacked some of the mundane realism that project and management teams experience working in business or government settings where their decisions have a substantial social or economic impact. In particular, while the disaster was designed to capture many of the psychological features that a team would face when responding to an actual disaster, participants did not face the consequences, trauma, or psychological stressors of an actual disaster setting. Further, team members interacted with one another over a relatively short period of time. This may have limited the complexity of the mental models that teams formed and ultimately impacted the magnitude of relationships among variables.

However, we also suggest that the SimCity simulation captured the psychological realism of situations that many knowledge-based teams experience. Moreover, the laboratory context enabled us to control extraneous effects and to obtain a clearer picture as to the relationships between the constructs. As such, we suggest that the current simulation provided a practical method for examining mental model metrics. Future research should examine these relationships in various types of knowledge-based teams within organizational settings and over a longer period of time.

Another limitation of the current study was the use of a single expert consensus model of mental model quality. Although this is the typical method for representing mental model accuracy, it precludes the possibility of multiple accurate models, which is necessary to make an ideal comparison of the distinct effects of accuracy versus similarity. The observed correlation between mental model accuracy and similarity for the network measure was .37. Thus, clearly there was some correspondence such that as models become more similar they also become more accurate and vice versa. Future research is needed to explore methods of operationalizing quality or accuracy in a manner that enables multiple accurate models so that independent effects of quality (or accuracy) versus similarity can be examined.

Another limitation is the length of each session, which lasted approximately four hours, requiring sustained task attention. This may have led to decrements in the amount of effortful attention members were able to provide (Grier et al., 2003). Similarly, it is possible that simultaneously completing the three mental model measures could have been taxing for participants; particularly the pairwise comparisons used for the structural networks metric. As a result, fatigue effects may have affected the quality of participant responses impacting the nature and magnitude of relationships.

Finally, as a result of missing data, there were unequal sample sizes across the elicitation techniques, ranging from 42 (priority rankings) to 56 (structural networks) teams. The additional power associated with the structural networks metric may have contributed to the

[1] We thank the Editor for suggesting this possibility.

ability to detect statistical significance in the predictive validity analyses. However, results of a power analysis indicated that we had more than adequate statistical power in the reduced sample of 42 teams. Future research should replicate this study using more consistent sample sizes across metrics.

Implications

The selection of mental model metrics is not a trivial matter for teams' researchers. Past reviews have emphasized that researchers should utilize a metric that most validly captures the underlying mental model construct. A clear implication of these findings, which is consistent with the meta-analytic findings of DeChurch and Mesmer-Magnus (in press), is that cognition is most predictive when it is measured in such a way as to represent the underlying structure of cognition. As an illustration of this point, consider Kilduff, Angelmar, and Mehra's (2000) study of strategic consensus and top management team decision-making in which a relationship was not found between strategic consensus and team performance; however, strategic consensus was assessed using team members' perceptions of consensus, a technique devoid of structural representation. Perhaps the conclusion regarding top management team strategic consensus and performance may have changed had a network-based metric been employed. We use this merely to illustrate the point that conclusions about the nomological net of team mental models are largely influenced by the degree to which measures of cognition represent structure. We see this as a valuable realization, which hopefully will prompt greater consideration of network representations in the next generation of collective cognition metrics.

References

Burke, C. S., Stagl, K. C., Salas, E., Pierce, L., & Kendall, D. L. (2006). Understanding team adaptation: A conceptual analysis and model. *Journal of Applied Psychology, 91,* 1189–1207.

Burns, T., & Stalker, G. M. (1961). *The management of innovation.* London, England: Tavistock Publications.

Campbell, D. T., & Fiske, D., W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cannon-Bowers, J. A., & Salas, E. (1990). *Cognitive psychology and team training: Shared mental models in complex systems.* Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Miami, FL.

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

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

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

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of team effectiveness: A meta-analysis. *Journal of Applied Psychology, 95,* 32–53.

DeChurch, L. A., & Mesmer-Magnus, J. R. (in press). Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research, and Practice.*

EA Games. (2004). *SimCity 4 Deluxe Edition.* Redwood City, CA: Electronic Arts.

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

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., & Parasuraman, R. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Human Factors, 45,* 349–360.

Kellermanns, F. W., Walter, J., Lechner, C., & Floyd, S. W. (2005). The lack of consensus about strategic consensus: Advancing theory and research. *Journal of Management, 31,* 719–737.

Kilduff, M., Angelmar, R., & Mehra, A. (2000). Top management team diversity and firm performance: Examining the role of cognitions. *Organization Science, 11,* 21–34.

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

Knight, D., Pearce, C. L., Smith, K. G., Olian, J. D., Sims, H. P., Smith, K. A., & Flood, P. (1999). Top management team diversity, group process, and strategic consensus. *Strategic Management Journal, 20,* 445–465.

Kramer, G. (2003). *Prima's official strategy guide: SimCity4 Deluxe Edition.* Roseville, CA: Random House.

Langan-Fox, J., Code, S., & Langfield-Smith, K. (2000). Team mental models: Techniques, methods, and analytical approaches. *Human Factors, 42,* 242–271.

LePine, J. A. (2003). Team adaptation and postchange performance: Effects of team composition in terms of members' cognitive ability and personality. *Journal of Applied Psychology, 88,* 27–39.

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

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

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

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

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

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

Reger, R. K. (1990). Managerial thought, structures and competitive positioning. In A. Huff (Ed.), *Mapping strategic thought* (pp. 71–88). Chichester, England: Wiley.

Rentsch, J. R., Small, E. E., & Hanges, P. J. (2008). Cognition in organizations and teams: What is the meaning of cognitive similarity? In D. B. Smith (Ed.), *The people make the place: Dynamic linkages between people and organizations* (pp. 127–155). New York, NY: Erlbaum.

Resick, C. J., Dickson, M. W., Mitchelson, J. K., Allison, L., & Clark, M. A. (in press). Team composition, cognition, and effectiveness: Examining mental model similarity and accuracy. *Group Dynamics: Theory, Research, and Practice.*

Rico, R., Sanchez-Manzanares, M., Gil, F., & Gibson, C. (2008). Team implicit coordination process: A team knowledge-based approach. *Academy of Management Review, 33,* 163–184.

Rouse, R., & Morris, N. M. (1986). On looking into the blackbox: Prospects and limits in the search for mental models. *Psychological Bulletin, 100,* 349–363.

Schvaneveldt, R. W. (1990). *Pathfinder associative networks: Studies in knowledge organization.* Westport, CT: Ablex Publishing.

Shobe, K., Fiore, S., & Carr, W. (2004). *Development of shared mental models for submarine officers* (NSMRL Tech. Rep. No. No. 1235). Groton, CT. Naval Submarine Medical Research Laboratory.

Smith-Jackson, T. L., & Wogalter, M. S. (2007). Application of a mental models approach to MSDS design. *Theoretical Issues in Ergonomics Science, 8,* 303–319.

Smith-Jentsch, K. A. (2009). Measuring team cognition: The devil is in the details. In E. Salas, J. Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (SIOP Frontiers Series; pp. 491–508). Mahwah, NJ: Erlbaum.

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

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

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.

Thomas-Hunt, M. C., & Phillips, K. W. (2003). Managing teams in the dynamic organization: The effects of revolving membership and changing task demands on expertise and status in groups. In R. S. Peterson, & E. A. Mannix (Eds.), *Leading and managing people in the dynamic organization* (pp. 115–133). Mahwah, NJ: Erlbaum Publishers.

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

Appendix A

Ten Key Decisions Used in the Mental Models Measures

1. Cut Spending
2. Increase Transportation Funding
3. Beautify the City
4. Build Power/Water Plants
5. Increase Public Safety Funding
6. Zone/Rezone Areas
7. Decrease Taxes
8. Reduce Air/Water Pollution
9. Build/Rebuild Hospitals/Clinics
10. Revitalizing Existing Buildings/Neighborhoods

Received August 4, 2009
Revision received December 3, 2009
Accepted December 15, 2009 ■