Group Dynamics: Theory, Research, and Practice
2010, Vol. 14, No. 1, 1–14

1089-2699/10/$12.00 DOI: 10.1037/a0017455

Measuring Shared Team Mental Models: A Meta-Analysis

Leslie A. DeChurch
University of Central Florida

Jessica R. Mesmer-Magnus
University of North Carolina Wilmington

Although shared team mental models are believed to be important to team functioning, substantial interstudy differences in the manner in which mental models are operationalized has impeded progress in this area. We use meta-analysis to cumulate 23 independent studies that have empirically examined shared mental models (SMMs) in relation to team process and performance and test three aspects of measurement as potential moderators: elicitation method, structure representation, and representation of emergence. Results indicate the way in which SMMs are measured and represented at the team level of analysis reveal meaningful distinctions in observed relationships. Specifically, shared mental model operationalization impacts the observed relationship between SMMs and team process; importantly, only methods that model the structure or organization of knowledge are predictive of process. Conversely, while the magnitude of the relationship differed across measurement method, SMMs were positively related to team performance regardless of the manner of operationalization. In summary, knowledge structure is predictive of team process, and both knowledge content and structure are predictive of team performance.

Keywords: team, group, mental model, cognition, meta-analysis

Teams are increasingly being utilized as the basic unit of work accomplishment, and both academics and practitioners have become particularly interested in the cognitive architecture that enables effective coordination and collaboration in work teams (Cannon-Bowers & Salas, 1990, 2001). At one end of the spectrum, action-oriented teams such as those used in medical and military settings perform highly interdependent time-sensitive tasks in dynamic environments (Burke, Salas, Wilson-Donnelly, & Priest, 2005). These teams require a pattern of cognitive similarity that enables them to anticipate one another's needs and actions and to synchronize their work in a way that is synergistic toward meeting the team's ultimate goals (Cannon-Bowers, Salas, & Converse, 1993; Cooke, Gorman, Duran, & Taylor, 2007; Marks, Zaccaro, & Mathieu, 2000). At the other end of the spectrum are more decision-oriented knowledge-based teams such as those engaged in software design and management consulting projects. These teams also rely on a pattern of cognitive similarity to effectively retrieve and share information (Ensley & Pearce, 2001; Faraj & Sproull, 2000). Across the diverse range of tasks that teams perform, research has consistently identified cognitive similarity as a key foundational element for success (DeChurch & Mesmer-Magnus, in press; Kozlowski & Ilgen, 2007; J. E. Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000). The thorn in the side of this research stream has been measuring and representing cognition at the team level (Smith-Jentsch, 2009). The current study cumulates past research on shared mental models (SMMs) and team outcomes to address the question: How should shared mental models be measured?

One type of cognitive architecture in teams that has received substantial research attention is the shared team mental model. Shared team mental models have their roots in the mental model construct from cognitive psychology. At the individual level, mental models reflect organized knowledge that enables humans to understand the basic

Leslie A. DeChurch, Department of Psychology, University of Central Florida; Jessica R. Mesmer-Magnus, Department of Management, University of North Carolina Wilmington.

We thank Eduardo Salas, Christian Resick, Ronald Piccolo, and Toshio Murase for their helpful suggestions on earlier versions of this work.

Correspondence concerning this article should be addressed to Leslie A. DeChurch, Department of Psychology, University of Central Florida, PO Box 161390, Orlando, FL 32816-1390. E-mail: ldechurc@mail.ucf.edu

functioning of systems—their purpose and form—and to form predictions and expectations about future states (Rouse & Morris, 1986). Cannon-Bowers and Salas (1990) were the first to extend the relevance of the mental model construct to teams. They invoked the concept of a team mental model to explain the performance of expert teams who demonstrated the ability to seamlessly coordinate their actions without the need for overt communication (Cannon-Bowers & Salas, 2001; Cannon-Bowers et al., 1993). *Shared team mental models*, then, are defined as "knowledge structures held by members of a team that enable them to form accurate explanations and expectations for the task, and in turn, to coordinate their actions and adapt their behavior to demands of the task and other team members" (Cannon-Bowers & Salas, 2001, p. 228).

Since their introduction to team effectiveness research in the early 1990s, the concept of shared team mental models has featured prominently in input–process–outcome models of team performance. Marks, Mathieu, and Zaccaro (2001) posited shared team mental models as an emergent state that both shapes and is shaped by the behavioral interaction processes playing out within teams. In this way, shared team mental models are mechanisms, reciprocally related to processes, that transmit or explain the effects of various input variables (e.g., composition, leadership, and training) on valued team outcomes (e.g., performance and viability).

Operationalizing Shared Team Mental Models

For more than a decade, researchers have voiced concerns about issues of measurement in shared mental model research (Cannon-Bowers & Salas, 2001; Klimoski & Mohammed, 1994; Mohammed, Klimoski, & Rentsch, 2000). Several excellent works have addressed the measurement issue conceptually and delineate important considerations and various methods for operationalizing team mental models (Cooke et al., 2007; Mohammed et al., 2000). Importantly, different operationalizations of the shared mental model construct vary in the extent to which they capture meaningful aspects of the nature and arrangement of cognition. These underlying differences are believed to impact the validity of cognition in predicting team process and performance (Klimoski & Mohammed, 1994; Mohammed et al., 2000; Smith-Jentsch, 2009). We test this idea, examining whether or not observed relationships between SMM and team process and performance differ on the basis of how cognition is operationalized. A cursory scan of the extant team mental model research shows substantial interstudy variation in the way in which team mental models are measured. Fortunately, enough empirical research on team mental models has now accumulated to permit posing the "how best to measure team mental models" question as an empirical one.

Theoretical work on the measurement of mental models generally outlines three important characteristics that permit measurement of the degree of convergence or similarity among team members' organized knowledge representations (Mohammed et al., 2000): (a) elicitation method, (b) structure representation, and (c) representation of emergence. The *elicitation method* refers to the technique used to determine the content or components of the model. Commonly used elicitation techniques are similarity ratings, concept maps, rating scales, and card sorting tasks (Mohammed et al., 2000). Presumably, all of these approaches to elicitation would be based on the results of a thorough task analysis that identifies the essential elements of the team's task. Different elicitation methods vary in the manner in which they present that information to participants. As an example, a mental model of the task of playing a computer flight simulation game could be elicited by using any of the above approaches. Similarity ratings could be used by presenting participants with a grid and then by requesting that they consider each pair of task nodes and report their perceptions of the relation between the two nodes. A rating scale would elicit the content of the model by asking participants to respond to questions about the task on fixed-response formats (e.g., strongly agree to strongly disagree). A concept map would elicit content by asking participants to place actions into some meaningful organization scheme.

Although elicitation methods capture the content of knowledge, they do not necessarily represent the structure of that knowledge. Because mental models are organized knowledge structures, a key aspect of their operationalization involves the degree of correspondence between how the knowledge contained in the model is represented in the mind and how the knowledge representation is modeled by the researcher. As such, *structure representation* constitutes a second crucial aspect of team mental model measurement (Klimoski & Mohammed, 1994; Mohammed et al., 2000). Similarity ratings typically prompt respondents to think in terms of the degree of association between distinct components of their team or task. In this way, they capture associative networks of knowledge. At the other extreme, rating scales capture levels of knowledge but do not model the structure or organization of that knowledge. Concept maps have been used to capture the sequencing of team actions; sequencing reflects an organization or structure of knowledge, though sequencing inherently models less structure than do network-based approaches (Marks et al., 2001).

The third distinguishing aspect of how mental models are operationalized, *representation of emergence*, concerns how individuals' mental models are collectively considered as constituents of a team mental model (Kozlowski & Klein, 2000). SMMs align with Chan's (1998) description of a dispersion construct. The multilevel nature of SMMs is similar to that of climate strength (Dickson, Resick, & Hanges, 2006), in which, as opposed to considering the content of the climate at the group level, it is the degree of sharedness in perceptions of climate that is of interest. Thus, because climate can be conceptualized as having both content (e.g., climate for safety) and strength (i.e., degree of consensus at focal level of analysis), so to can SMMs. When SMMs are considered at the team level, it is typically the degree of similarity in models (analogous to climate strength) as opposed to the particular content of the team model (analogous to climate type) that is of interest. Researchers have tended to index SMMs at the team level by using indices reflecting the level of similarity in the group. For example, in studies using pairwise comparison data, network analysis algorithms such as Pathfinder's C or UCINET's QAP correlation (Borgatti, Everett, & Freeman, 1992) are typically used to compare the overlap of team member's models. Rwg(j) has been used (James, Demaree, & Wolf, 1984) to reflect agreement within the team. Euclidean distance has been reported as a geometric representation of the separation or closeness of the model.

In summary, we propose that these three aspects of how mental models are operationalized (elicitation, structure representation, and representation of emergence) will moderate the strength of observed relations between mental models and outcomes because these strategies are likely to articulate the relationships with varying degrees of reliability and clarity. Because SMMs have been featured prominently in the input–process–outcome models of team performance, we explore the moderating potential of mental model measurement in relation to both team process and team performance.

Method

Database

Twenty-three independent studies reported in 22 journal articles, dissertations, and conference presentations (total number of groups = 1511; total *N* approximately 5668) examining SMMs were included in this meta-analysis. To ensure a comprehensive search, we located these studies by using the following strategies: (a) a computerized search of the PsycInfo, ABI Inform, and ERIC databases, using appropriate key words and phrases (e.g., *team* or *group* and *cognition*, *mental models*, *shared cognition*, *schemas*, *knowledge structure*, *cognitive structure*, *cognitive map*, *conceptual framework*, etc.); (b) a manual search for references cited in studies included in this meta-analysis; (c) soliciting relevant unpublished manuscripts from authors currently doing research in team cognition; and (d) obtaining related studies from recent conference presentations (e.g., Society for Industrial Organizational Psychology and Academy of Management). Our objective in examining studies from the recent conferences was to incorporate relevant peer-reviewed research results that had not yet been incorporated into the extant literature.

These search strategies resulted in an initial 225 empirical studies with the potential to report relationships relevant to our research questions. These manuscripts were carefully reviewed for relevance; only 23 of these studies were found to report sufficient information about relevant constructs to be retained in this meta-analysis. In order to be retained, studies must have (a) assessed SMMs in relation to team process, team performance, or both and (b) provided sufficient information to compute a correlation between these variables (means and standard deviations, effect sizes, etc.).

Only studies examining task-performing teams were included in the meta-analysis (Kozlowski & Ilgen, 2007). Although teams and groups both represent meaningful social units, teams possess the additional qualifier of interdependence. Teams are a subset of groups whose members share a common goal and must cooperate in order to attain shared goals (Guzzo & Shea, 1992; Wageman, 1995). We included both the terms *team* and *group* in our search to ensure that we identified all appropriate articles, regardless of the label applied by the author or authors.

Furthermore, because SMMs generally describe the extent to which team members' mental models are overlapping or similar (Cooke et al., 2007; Edwards, Day, Arthur, & Bell, 2006; Mohammed et al., 2000), to fit within the context of this meta-analysis, team cognition constructs examined within the primary studies must have attempted to explicate (a) the content of known elements of the team members' mental models (elicitation) and (b) the collective representation (similarity) of the team members' elemental content (representation of emergence; Klimoski & Mohammed, 1994; Kozlowski & Klein, 2000; Mohammed et al., 2000).[1] Studies were omitted if they dealt with constructs that differed from those examined in this research or did not report sufficient information to assess the mental model measurement strategy used (in terms of elicitation method, structure representation, or representation of emergence). For example, research in the transactive memory tradition assesses the distribution of team cognition but does not attempt to assess the degree of overlap among team members' mental models of task or team concepts. As such, transactive memory research was not included in this meta-analysis. Similarly, research exploring the accuracy of mental models typically examines the overlap between a team member's mental model and that of an expert's but does not necessarily examine the overlap in cognition among team members. These types of studies were also excluded. Finally, because SMMs are meaningful at the team level, we excluded studies that only reported relationships at the individual level. This practice was more common, for example, in the human factors' literature than in the Industrial/Organizational (I/O) psychology and management literatures.

When authors reported separate correlations for different samples, those correlations were examined separately. The manuscripts included in this meta-analysis are listed in the references prefixed with an asterisk.

Coding Procedure

We each undertook an independent effort to code the 23 studies that met criteria for inclusion in the analysis by using a jointly developed coding scheme. Intercoder agreement was very high (95.4% across all variables coded), likely because of the objective nature of the data coded. Table 1 reports intercoder reliabilities for each of the study's key variables. Coding disagreements were resolved through discussion. Data coded included study sample size, number of groups included, sample characteristics, method used to operationalize SMMs (in terms of the elicitation method, structure representation, and representation of emergence used), and relationships between SMMs and team process and team performance. When reported, we also coded reliability estimates of team cognition, process, and performance. A summary of study and sample characteristics (sample, team task, study design and setting, and shared mental model measurement moderators) for the primary studies is included in the Appendix.

Coding of process and performance. We coded relationships between team SMMs and team process and team performance. *Team process* describes verbal and behavioral acts that team members engage in to transform inputs to outcomes (Marks et al., 2001). Examples of team process included in these studies are coordination, backup behavior, and planning. Although several forms of team process have been examined (transition, action, and interpersonal process; Marks et al., 2001), an insufficient number of studies was available to permit an examination of the measurement moderators for each type of process. As such, we collapsed across form of process, computing composite

[1] Researchers in the shared mental model tradition also frequently attempt to explicate the relationships between team cognitive elements (structure representation). As this is not uniformly done in shared mental model research, we included articles that did not assess structure representation in our database but examined this measurement strategy moderator whenever possible.

Table 1
Summary of Coder Reliabilities for Key Study Variables

Variable	% agreement	κ	95% CI κ
Team process	100	1.0	1.0/1.0
Team performance	100	1.0	1.0/1.0
Elicitation method	100	1.0	1.0/1.0
Structure representation	92.3	.89	.75/1.0
Representation of emergence	96.2	.95	.86/1.0

Note. CI = confidence interval. The 95% CI κ refers to the 95% confidence interval around κ. Cohen's κ is a measure of interrater reliability often used in conjunction with percentage of agreement to index intercoder consistency in meta-analyses (Lipsey & Wilson, 2001). All κ coefficients were significant at $p < .001$. The κ coefficients greater than .80 are considered high or outstanding (Landis & Koch, 1977).

correlations as appropriate, to examine (a) the role of SMMs in team process overall and (b) the role of the measurement moderators in the shared mental model–team process relationship.

Team performance is an objective or subjective judgment of how well a team meets valued objectives (Salas, Rosen, Burke, & Goodwin, 2009). Team performance was typically operationalized as task performance, completion, or proficiency. We first examined the role of SMMs in team performance overall and then examined the role of the measurement moderators in the shared mental model–team performance relationship.

Coding of mental model measurement. In order to examine the potential moderating role of mental model measurement method on team process and performance relationships, we coded the way in which authors in the primary studies assessed the degree of convergence or similarity in knowledge among team members (Mohammed et al., 2000). Specifically, whenever sufficient information was provided, we coded the method of *elicitation* (technique by which content or elements of a mental model are assessed or elicited) and *structure representation* (technique by which the relationships between the mental model elements are assessed; the way in which structure of relationships between data is revealed) used in the primary studies. We also coded the method by which *emergence* was represented (i.e., the technique used to represent the elemental content at the team level, i.e., the degree of "sharedness" across members of a team; Kozlowski & Klein, 2000).

Elicitation was typically assessed with similarity ratings, concept mapping, or card sorting, questionnaires, or content analysis of cognitive maps. Importantly, researchers had two key approaches to eliciting mental models: (a) providing respondents predefined categories of cognitive content to rate, sort, or map (i.e., similarity ratings, concept mapping or card sorting, and questionnaire or rating) or (b) inducing respondents to provide their mental models without using predefined categories (i.e., cognitive map or content analysis). The key difference between the two approaches lies in whether the researchers generate the categories and components of the mental model for the respondent or whether they induce the respondent to provide their mental model content themselves. An example of the first approach might involve a researcher providing a respondent a list of task concepts and asking them to rate the similarity among concepts (e.g., J. E. Mathieu, Heffner, & Goodwin, 2005). An example of the second approach might involve a researcher content-analyzing a respondent's description of their mental model (e.g., Fleming, Wood, Bader, & Zacarro, 2003).

Structure representation was typically assessed with Pathfinder, UCINET, multidimensional scaling (MDS), concept mapping, or card sorting. The objective of these approaches is to determine the way in which categories and concepts in the team's mental model are organized. Approximately half of the studies in our database did not report any attempt at structure representation.

Representation of emergence indexes the extent to which team members' mental models converge and was typically accomplished with agreement indices (percentage overlap and within group inter-rater reliability (Rwg)), team member consistency (*r*, Intraclass Correlation Coefficient (ICC), and alpha), a concept mapping scoring system, Pathfinder (C), UCINET convergence indices (QAP), Euclidean distance (MDS), or consistency with Euclidean distance (*r*, ICC, alpha, and MDS).

Analysis

The meta-analytic methods outlined by Hunter and Schmidt (2004) were used to analyze this data. Because reliability estimates for team cognition and its relevant correlates were not reported in all studies, corrections for unreliability were accomplished by using artifact distribution meta-analysis. Artifact distribution meta-analysis recognizes that it is often the case that every study in a meta-analysis does not report all necessary information to correct for attenuation on all the artifacts that impact relationships in that study (Hunter & Schmidt, 2004). Thus, correction for these artifacts (in this case, reliability) is accomplished with a distribution of artifact values collected from studies that do provide relevant artifact information. Corrections were made for unreliability in both SMMs and correlate measures (process and performance). Given the possibility of a file-drawer effect (wherein significant findings are more likely to be published; Rosenthal, 1979), we also conducted a file-drawer analysis (Hunter & Schmidt, 2004) to estimate the number of studies reporting null findings that would be required to reduce reliability-corrected correlations to a specified lower value (we used $\rho = .05$).

Results

Tables 2 and 3 present meta-analytic results examining team cognition relations. For each analysis, we report the number of independent samples included in the meta-analysis (k), the total number of groups across these samples (N), the sample size weighted mean observed correlation (r), the sample size weighted standard deviation associated with the observed mean (SDr), the reliability corrected correlation (ρ), the standard deviation associated with the ρ, the 80% credibility interval around ρ, the 90% confidence interval around ρ, percentage of variance due to sampling error (%SEV), and the percentage of variance due to all statistical artifacts (%ARTV).[2]

Measurement Strategy as a Moderator of the SMM Similarity–Team Process Relationship

We first examined the operationalization of SMMs as a moderator of the relationship between SMMs and team process. Table 2 presents results examining elicitation method, structure representation, and representation of emergence as moderators of the SMM–team process relationship. Results indicate that SMMs are most strongly related to team process when models are elicited by using similarity ratings ($\rho = .27$, $SD\rho = 0$). The next strongest relationship results from models elicited with concept mapping or card sorting ($\rho = .17$, $SD\rho = .06$). Notably, these two confidence intervals do not overlap, suggesting that elicitation does moderate the size of the observed SMMs–team process relationship. When models were elicited with questionnaires or rating scales, the relationship between SMMs and team process was small and negative ($\rho = -.05$, $SD\rho = .37$), with a credibility interval ranging from $-.52$ to .43.

Next we examined the SMMs–team process relationship on the basis of how model structure was represented. The strongest effects were evidenced with the Pathfinder PFNets ($\rho = .31$, $SD\rho = 0$). Positive, albeit smaller, effects were seen when structure was represented with a card-sorting or concept-mapping approach ($\rho = .17$, $SD\rho = .06$). Notably, the confidence intervals do not overlap, suggesting that structure representation moderates this relationship. Lastly, when the method used to operationalize SMMs did not enable the representation of structure, there was essentially no observed effect between SMMs and team process ($\rho = .05$, $SD\rho = .28$).

Lastly, we compared the size of the SMM–team process relationship on the basis of the metric used to represent the "shared" aspect of team mental models. Notably, effect sizes were highest when using Pathfinder C ($\rho = .30$, $SD\rho = 0$) but were also high when using consistency or Euclidean distance indices ($\rho = .20$, $SD\rho = 0$). However, studies using an index of within-group agreement showed no relation between SMMs and team process ($\rho = -.06$, $SD\rho = .35$).

[2] We report both the credibility and confidence intervals, because each provides unique information about the nature of ρ (Hunter & Schmidt, 2004; Whitener, 1990). Specifically, the credibility interval (CV) provides an estimate of the variability of corrected correlations across studies. Wide credibility intervals or those that include zero suggest the presence of a moderator. An 80% CV that excludes zero indicates that more than 90% of the corrected correlations are different from zero (10% lie beyond the upper bound of the interval). The confidence interval (CI) provides an estimate of the accuracy of our estimation of ρ (Whitener, 1990); in other words, the confidence interval estimates the variability around ρ due to sampling error. A 90% CI that excludes zero indicates that if our estimation procedures were repeated many times, 95% of the estimates of ρ would be larger than zero (5% would fall beyond the upper limit of the interval).

Table 2
Mental Model Measurement as a Moderator of the Shared Mental Model–Team Process Relationship

Meta-analysis	*k*	*N*	*r*	*SDr*	ρ	*SD*ρ	80% CV	90% CI	% SEV	% ARTV	FD *k*
SMMs–team process	17	1,238	.20	.23	.22	.22	−.06/.50	−.14/.58	25.49	25.83	58
Elicitation method											
Similarity ratings	6	271	.24	.06	.27	0	.27/.27	.27/.27	100	100	26
Concept mapping/card sorting	3	208	.16	.13	.17	.06	.10/.24	.07/.27	82.10	82.49	7
Questionnaire/rating scale	5	302	−.04	.34	−.05	.37	−.52/.43	−.66/.56	14.53	14.53	—
Cognitive map/content analysis	1	20	.29	—	—	—	—	—	—	—	—
Structure representation											
Pathfinder	3	116	.28	.02	.31	0	.31/.31	.31/.31	100	100	16
UCINET	1	56	.17	—	—	—	—	—	—	—	—
Multidimensional scaling	—	—	—	—	—	—	—	—	—	—	—
Concept mapping/card sorting	3	208	.16	.13	.17	.06	.10/.24	.07/.27	82.10	82.49	7
None	9	538	.05	.28	.05	.28	−.31/.42	−.41/.51	20.96	20.98	—
Representation of emergence											
Agreement	5	320	−.05	.33	−.06	.35	−.50/.39	−.64/.52	14.15	14.16	1
Consistency	4	219	.19	.05	.20	0	.20/.20	.20/.20	100	100	12
Concept map scoring system	1	63	.23	—	—	—	—	—	—	—	—
Pathfinder	2	106	.28	.01	.30	0	.30/.30	.30/.30	100	100	10
UCINET convergence	1	56	.17	—	—	—	—	—	—	—	—
Euclidean distance	1	29	.15	—	—	—	—	—	—	—	—
Consistency with Euclidean distance	4	219	.19	.05	.20	0	.20/.20	.20/.20	100	100	12

Note. This table reports the results of an artifact distribution meta-analysis examining the measurement of shared mental models as a moderator of the relationship between team process and shared mental models. SMMs = shared mental models; *k* = number of correlations meta-analyzed; *N* = total number of groups across the correlations meta-analyzed; *r* = sample size–weighted mean observed correlation; *SDr* = sample size–weighted standard deviation of the correlations; ρ = sample size–weighted mean observed correlation corrected for unreliability in both measures; *SD*ρ = standard deviation of ρ; 80% CV = 80% credibility interval around ρ; 90% CI = 90% confidence interval around ρ; % SEV = percentage of variance due to sampling error; % ARTV = percentage of variance due to all other artifacts; FD *k* = file drawer *k*; MDS = multidimensional scaling.

Table 3
Mental Model Measurement as a Moderator of the Shared Mental Model–Performance Relationship

Meta-analysis	*k*	*N*	*r*	*SDr*	ρ	*SD*ρ	80% CV	90% CI	% SEV	% ARTV	FD *k*
SMMs–team performance	19	1,363	.29	.13	.33	.09	.22/.44	.18/.48	65.87	68.33	106
Elicitation method											
Similarity ratings	9	483	.20	.15	.23	.07	.14/.32	.11/.35	80.74	81.35	32
Concept mapping/card sorting	6	411	.28	.10	.31	0	.31/.31	.31/.31	100	100	31
Questionnaire/rating scale	6	566	.32	.16	.39	.16	.18/.59	.13/.65	32.88	33.95	41
Cognitive map/content analysis	2	47	.50	.14	.56	0	.56/.56	.56/.56	100	100	61
Structure representation											
Pathfinder	5	287	.28	.08	.31	0	.31/.31	.31/.31	100	100	26
UCINET	1	56	.14	—	—	—	—	—	—	—	—
Multidimensional scaling	1	41	.48	—	—	—	—	—	—	—	—
Concept mapping/card sorting	6	411	.28	.10	.31	0	.31/.31	.31/.31	100	100	31
None	9	756	.27	.16	.32	.15	.14/.51	.07/.57	40.19	41.52	49
Representation of emergence											
Agreement	3	467	.36	.11	.39	.09	.27/.49	.24/.54	41.04	44.05	20
Consistency	6	289	.13	.13	.16	0	.16/.16	.16/.16	100	100	13
Concept map scoring system	4	266	.32	.10	.35	0	.35/.35	.35/.35	100	100	24
Pathfinder	5	287	.28	.08	.31	0	.31/.31	.31/.31	100	100	26
UCINET convergence	1	56	.14	—	—	—	—	—	—	—	—
Euclidean distance	3	117	.27	.15	.31	.03	.26/.35	.26/.36	95.41	96.27	16
Consistency with Euclidean distance	8	377	.17	.16	.20	.07	.11/.29	.08/.32	84.08	84.50	24

Note. This table reports the results of an artifact distribution meta-analysis examining the measurement of shared mental models as a moderator of the relationship between shared mental models and team performance. SMMs = shared mental models; *k* = number of correlations meta-analyzed; *N* = total number of groups across the correlations meta-analyzed; *r* = sample size weighted mean observed correlation; *SDr* = sample size weighted standard deviation of the correlations; ρ = sample size weighted mean observed correlation corrected for unreliability in both measures; *SD*ρ = standard deviation of ρ; 80% CV = 80% credibility interval around ρ; 90% CI = 90% confidence interval around ρ; % SEV = percentage of variance due to sampling error; % ARTV = percentage of variance due to all other artifacts; FD *k* = file drawer *k*; MDS = multidimensional scaling.

Measurement Strategy as a Moderator of the SMM Similarity–Team Performance Relationship

Next, we examined operationalization as a moderator of the relationship between SMMs and team performance. Table 3 reports the results of meta-analyses examining measurement strategy as a moderator of the relationships between SMM similarity and team performance. First, examining the SMM–team performance relationship on the basis of the method used to elicit the mental model shows that all elicitation methods reveal positive relationships between SMMs and team performance. The strongest relationship was found when cognitive mapping or content analysis was used ($\rho = .56$, $SD\rho = 0$).

Next, examining the SMM–team performance relationships on the basis of the representation of structure shows highly comparable estimates using Pathfinder ($\rho = .31$, $SD\rho = 0$), concept mapping and card sorting ($\rho = .31$, $SD\rho = 0$), and even estimates in which mental model structure is not captured ($\rho = .32$, $SD\rho = .15$).

Lastly, examining the SMM–team performance relationship on the basis of representation of emergence shows that positive effects are evidenced across all indices. The strongest effects were observed when either an index of agreement ($\rho = .39$, $SD\rho = .09$) or a concept mapping scoring system was used ($\rho = .35$, $SD\rho = 0$), followed by either Pathfinder C ($\rho = .31$, $SD\rho = 0$) or Euclidean distance ($\rho = .31$, $SD\rho = .03$), and lastly by consistency with Euclidean distance ($\rho = .20$, $SD\rho = .07$).

Discussion

Does it really matter how SMMs are operationalized? The current results suggest that the answer depends on whether one wishes to predict team process or team performance. Interestingly, aspects of mental model measurement appear to be more paramount when studying relations to team process than to performance, a difference that appears to tie to the differential necessity of representing model structure. The current meta-analysis suggests that SMMs only predict team process when the measurement technique enables the structure of individuals' mental models to be revealed. In contrast, SMMs predict team performance across measurement techniques.

With team process, the strongest relationships were evident when similarity ratings were used to elicit the content of the model, the Pathfinder network analysis algorithm was used to represent the structure of the model, and Pathfinder's C was used as an index of team mental model similarity. Weaker but positive validity coefficients were also obtained when researchers elicited the mental model by using a concept map or card-sorting task, represented structure using a subject matter expert-constructed scoring system, and indexed team similarity with a consistency metric. Traditional rating-scale techniques did not show a relationship between mental model similarity and team process. This is likely due to their deficiency in representing the structure of knowledge (Mohammed et al., 2000).

The pattern differed when SMMs were used to predict team effectiveness. Across the elicitation methods, structure representation approaches, and representation of emergence used in past research, positive relations were observed between the similarity in mental models and team performance. In contrast to predicting team process, shared team mental models show a positive effect on team performance even when more traditional, nonstructural measurement techniques were used (i.e., questionnaires and rating scales). Measures of mental models contain varying parts of knowledge content and knowledge structure. This pattern of findings suggests that knowledge content is predictive of team performance but not of team process. Knowledge structure is predictive of both team process and team performance.

Given the moderate effect size between team process and team performance ($\rho = .31$; LePine, Piccolo, Jackson, Mathieu, & Saul, 2008), the differential validity found here with process and performance further suggests that measurement may represent unique aspects of cognition that are incrementally predictive of team effectiveness. Whereas all representations of cognition are comparatively predictive of team performance, structured cognition is more predictive of process. This may be indicative of the fact that there are meaningful conceptual differences in what is captured, and so an interesting avenue for future research is to explore multiple methods for operationalizing mental models within the same sample of teams. Perhaps mental models are most diagnostic when represented using a variety of operationalization tools.

Limitations

Although the current study provides practical insight into how best to measure SMMs, there are several important limitations to consider. As with any meta-analysis, this study is limited by the availability of reported effect-size estimates. Some of the relationships we examined had very little data available for cumulation, resulting in small *k*s. We chose to report as much detail as possible regarding effect sizes broken out by particular combinations of measurement features, though we recognize that this resulted in numerous instances of $k = 1$ or small k analyses. We recognize that small-*k* meta-analyses are prone to second-order sampling error (Hunter & Schmidt, 2004), and thus we refrained from specifically interpreting these analyses. Although we believe that the take-away message of these analyses as a whole—that structure is a relevant component of mental model criteria—can be seen by cumulating extant research, we urge against drawing more fine-grained conclusions about particular comparisons across methods until additional research is available.

A second limitation is that we were not able to explore differences in effect sizes on the basis of measurement characteristics for different types of mental models (e.g., accuracy). The studies contained in this meta-analysis all examined the similarity or congruence (Rentsch, Small, & Hanges, 2008) of team mental models. An open question is to what extent knowledge structure is an important aspect of additional team cognitive criteria such as accuracy and complementarity.

Future Directions

The current findings support the relevance of structure to SMM criteria. This finding has relatively straightforward implications for future laboratory investigations of mental models; care should be taken to design and implement measures of mental models that capture not only the content but also the structure of the mental model. The reality is that collecting data with pairwise comparison matrices and card-sorting techniques is time consuming, is labor intensive, takes longer to analyze (and thus do not render ready feedback for constituencies in the field), and may be viewed more skeptically in terms of content validity. An important task for future research is to refine current measures or develop new measurement approaches that enable the relevant aspects of mental model criteria (i.e., structure) to be captured while also enabling their use in a variety of field settings. This problem is further compounded by the fact that many of the disaster response, military, and medical settings in which mental models are thought to be both important and dynamic are precisely the settings in which cumbersome measurement methods such as card sorts and pairwise comparisons are the least feasible.

Conclusion

Although shared team mental models were identified as important drivers of team effectiveness over 15 years ago (Cannon-Bowers et al., 1993), the complexity involved in capturing this collective cognitive construct has prompted researchers to use a variety of different measurement approaches. This variation in methodology poses a challenge to the aggregate interpretation of findings. The current study used meta-analysis to empirically cumulate past research linking shared team mental models to team process and team performance and finds that measurement does impact the magnitude of effects observed between mental models and outcomes. Future research on shared team mental models would be well served to utilize methods than enable structure to be captured.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Borgatti, S., Everett, M. G., & Freeman, L. C. (1992). *UCINET IV network analysis software.* Columbia, SC: Analytic Technologies.

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,* 96–104.

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

Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on team 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 N. J. Castellan, Jr. (Ed.), *Individual and group decision making: Current issues* (pp. 221–246). Hillsdale, NJ: Erlbaum.

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

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.

*DeChurch, L. A. (2003). Teams leading teams: Examining the role of leadership in multi-team systems. *Dissertation Abstracts International: Section B. Sciences and Engineering, 64*(01), 450.

DeChurch, L. A., & Mesmer-Magnus, J. R. (in press). The cognitive underpinnings of team effectiveness: A meta-analysis. *Journal of Applied Psychology*.

Dickson, M. W., Resick, C. J., & Hanges, P. J. (2006). When organizational climate is unambiguous, it is also strong. *Journal of Applied Psychology, 91,* 351–364.

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

*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.

Ensley, M. D., & Pearce, C. L. (2001). Shared cognition in top management teams: Implications for new venture performance. *Journal of Organizational Behavior, 22,* 145–160.

Faraj, S., & Sproull, L. (2000). Coordinating expertise in software development teams. *Management Science, 46,* 1554–1568.

*Fleming, P. J., Wood, G. M., Bader, P. K., & Zaccaro, S. J. (2003). *The locus of SMM: Whence does the sharedness come?* Paper presented at the annual meeting of the Society for Industrial Organizational Psychology, Orlando, FL.

Guzzo, R. A., & Shea, D. (1992). Group performance and intergroup relations in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, pp. 269–313). Palo Alto, CA: Consulting Psychologists Press.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). New York, NY: Sage.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69,* 85–98.

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.

Kozlowski, S. W. J., & Ilgen, D. R. (2007). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest, 7,* 77–124.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In S. W. J. Kozlowski & K. J. Klein (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). San Francisco, CA: Jossey-Bass.

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

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.

*Levesque, L. L., Wilson, J. M., & Whorley, D. R. (2001). Cognitive divergence and shared mental models in software development project teams. *Journal of Organizational Behavior, 22,* 135–144.

*Lim, B. C., & Klein, K. (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.

*Lim, R. A. (2004). Patterns of cognition in information systems development teams. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 65*(05), 1862.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review, 26,* 356–376.

*Marks, M. A., Sabella, M. J., Burke, C. S., & Zaccaro, S. (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., Maynard, M. T., Rapp, T. L., & Mangos, P. M. (2006). *Interactive effects of team and task shared mental models as related to air traffic controller's collective efficacy and effectiveness.* Paper presented to the annual meeting of the Academy of Management, Atlanta, GA.

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

*Minionis, D. P. (1995). Enhancing team performance in adverse conditions: The role of shared team mental models and team training on an interdependent task. *Dissertation Abstracts International: Section B. Sciences and Engineering, 56*(02), 1139.

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

*Mohammed, S., & Ringseis, E. (2001). Cognitive diversity and consensus in group decision making: The role of inputs, processes, and outcomes. *Organizational Behavior and Human Decision Processes, 85,* 310–335.

*Probber, J. K. (2000). Shared mental models as an antecedent of work group effectiveness. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 60*(08), 3018.

*Rapert, M. I., Velliquette, A., & Garretson, J. A. (2002). The strategic implementation process: Evoking strategic consensus through communication. *Journal of Business Research, 55,* 301–310.

*Rentsch, J. R., & Klimoski, R. J. (2001). Why do "great minds" think alike? Antecedents of team member schema agreement. *Journal of Organizational Behavior, 22,* 107–120.

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

*Resick, C. J., Dickson, M. W., & Mitchelson, J. K. (2005). *Cognitive ability, personality, and shared mental models in teams.* Paper presented to the annual meeting of the Society for Industrial Organizational Psychology, Los Angeles, CA.

*Rittman, A. L. (2004). A comparison of team training strategies for team effectiveness. *Dissertation Abstracts International: Section B. Sciences and Engineering, 65*(05), 2679.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

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

Salas, E., Rosen, M., Burke, C. S., & Goodwin, G. F. (2009). The wisdom of collectives in organizations: An update of the teamwork competencies. In E. Salas, G. F. Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations.* New York, NY: Taylor & Francis.

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.* New York, NY: Taylor & Francis.

*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.

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

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

Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology, 75,* 315–321.

Received July 31, 2008
Revision received February 20, 2009
Accepted February 23, 2009 ■

Appendix

Meta-Analytic Database Study Characteristics

Author (year)	Sample	Team task	Design	Setting	Elicitation method	Structure representation	Representation of emergence
DeChurch (2003)	Undergraduates	*Total War* flight simulation	Experiment	Lab	Questionnaire/rating	None	Consistency
Edwards, Day, Arthur, and Bell (2006)	Undergraduates	*Space Fortress* video game	Experiment	Lab	Similarity ratings	Pathfinder	Pathfinder (C)
Ellis (2006)	Undergraduates	*Distributed Dynamic Decision-Making* simulation	Experiment	Lab	Concept mapping/ card sorting	Concept mapping/ card sorting	Concept map scoring system
Fleming, Wood, Bader, and Zaccaro (2003)	Undergraduates	*JSTARS* battle simulation	Experiment	Lab	Questionnaire/rating	Concept mapping/ card sorting	Concept map scoring system
Knight et al. (1999)	Top management teams in high tech firms	Various	Nonexperiment	Field	Questionnaire/rating	None	Agreement
Levesque, Wilson, and Whorley (2001)	Undergraduate	Develop software product	Non-experiment	Lab	Questionnaire/rating	None	Agreement
R. A. Lim (2004)	Graduate students	Design computer system architecture documents	Nonexperiment	Lab	Cognitive map/ content analysis	None	Consistency
B. C. Lim and Klein (2006)	Military combat teams from Singapore Armed Forces	Military combat tasks	Nonexperiment	Field	Similarity ratings	Pathfinder	Pathfinder (C)
Marks, Zaccaro, and Mathieu (2000)	Undergraduates	Tank wargame simulation	Experiment	Lab	Concept mapping/ card sorting	Concept mapping/ card sorting	Concept map scoring system
Marks, Sabella, Burke, and Zaccaro (2002)	Undergraduates	*Longbow2* (a) and *TWIST* (b) military flight/tank simulations	Experiment	Lab	Similarity ratings (a); concept mapping/card sorting (b)	Pathfinder (a); none (b)	Pathfinder C (a); agreement (b)
Mathieu, Heffner, Goodwin, Salas, and Cannon-Bowers (2000)	Undergraduates	Flight simulator	Nonexperiment	Lab	Similarity ratings	UCINET	Consistency

(*Appendix continues*)

Appendix (*continued*)

Author (year)	Sample	Team task	Design	Setting	Elicitation method	Structure representation	Representation of emergence
Mathieu, Heffner, and Goodwin (2005)	Undergraduates	Flight simulator	Nonexperiment	Lab	Similarity ratings	None	Consistency
Mathieu, Maynard, Rapp, and Mangos (2006)	Air traffic control teams	Coordinate takeoff/landing, ensure route timing, etc.	Nonexperiment	Field	Questionnaire/rating	None	Consistency
Minionis (1995)	Undergraduates	Tank battle simulation	Experiment	Lab	Concept mapping/ card sorting	Concept Mapping/ Card Sorting	Agreement
Mohammed and Ringseis (2001)	Undergraduates	Towers market multi-issue negotiation task	Experiment	Lab	Questionnaire/rating	None	Agreement
Probber (2000)	Undergraduates	Computer simulation of hotel management teams	Experiment	Lab	Similarity ratings	None	Euclidean distance–MDS
Rapert, Velliquette, and Garretson (2002)	CEO/marketing executive teams	Various	Nonexperiment	Field	Questionnaire/rating	None	Agreement
Rentsch and Klimoski (2001)	Intact work teams from U.S. Department of Defense organization	Various	Nonexperiment	Field	Similarity ratings	Multidimensional Scaling	Euclidean distance–MDS
Resick, Dickson, and Mitchelson (2005)	Undergraduates	Freelancer battle simulation	Nonexperiment	Lab	Similarity ratings	Pathfinder	Pathfinder (C)
Rittman (2004)	Undergraduates	*Red Alert* multiplayer videogame	Experiment	Lab	Concept mapping/ card sorting	Concept mapping/ card sorting	Concept map scoring system
Smith-Jentsch, Mathieu, and Kraiger (2005)	Air traffic control teams	Coordinate takeoff/landing, guide aircraft, ensure route timing	Nonexperiment	Field	Questionnaire/rating	None	Euclidean distance–MDS
Stout, Cannon-Bowers, and Salas (1999)	Undergraduates	*Gunship* helicopter simulation	Nonexperiment	Lab	Similarity ratings	Pathfinder	Pathfinder (C)

Note. CEO = chief executive officer; MDS = multidimensional scaling.