Enhancing Shared Mental Models: A Systematic Review and Meta-Analysis of Randomised Controlled Trials

Robin L. J. Lines[1], Benjamin L. Hoggan[2], Sasha Nahleen[2], Philip Temby[2], Monique Crane[3], and Daniel F. Gucciardi[1,4]

[1]*Curtin School of Allied Health, Curtin University*

[2]*Land Division, Defence Science and Technology Group*

[3]*School of Psychology, Macquarie University*

[4]*Curtin enAble Institute, Faculty of Health Sciences, Curtin University*

Author Notes

*Address correspondence to Daniel Gucciardi, Curtin School of Allied Health, Curtin University, GPO Box U1987, Perth, Australia, 6845. Email: daniel.f.gucciardi@gmail.com

Funding statement: The Commonwealth of Australia supported this research through the Australian Army and a Defence Science Partnerships agreement of the Defence Science and Technology Group, as part of the Human Performance Research Network. Our prospectively registered protocol, data, and analysis script are located on the project page on the Open Science framework (OSF; https://osf.io/yzdxn/).

,

Lines, R.L.J., Hoggan, B., Nahleen, S., Temby, P., Crane, M., & Gucciardi, D.F. (in press). Enhancing shared mental models: A systematic review and meta-analysis of randomised controlled trials. *Sport, Exercise, and Performance Psychology*. doi: 10.1037/spy0000288

Abstract

Having everyone on the 'same page' in teams working towards a common objective is essential to effective teamwork, yet an integrative understanding of factors that enhance these shared mental models is absent from the evidence base. We addressed this gap in the literature via a prospectively registered (https://osf.io/yzdxn/) systematic review of five databases to identify eligible studies and to statistically synthesise evidence from 36 lab or field experiments (131 effect sizes, $N_{participants}$ = 6,209, N_{teams} = 1,912) that tested the effectiveness of team development interventions for enhancing shared mental models among adults where participants were randomised to experimental groups. Via a three-level random effects meta-analysis, we found a positive and significant medium-to-large overall effect of team development interventions on shared mental models (g = .61, 95% CI = .41, .82); sensitivity and meta-bias analyses (e.g., risk of bias, GRADE assessment) generally supported the robustness of this overall effect. Moderator analyses indicated that outcome assessment method meaningfully altered the overall pooled effect, with stronger effects observed when outcomes were researcher-assessed. Nevertheless, our assessment indicated low certainty in the quality of the evidence and 'noisiness' in the overall estimate (i.e., prediction interval of -0.66 and 1.89). Overall, this study contributes new knowledge on the antecedents of shared mental models that can inform theory regarding the nomological network of this concept, as well as methodological insights that can improve the evidence-base in future work.

Keywords: cognitive map; shared cognition; team cognition; three-level meta-analysis

Enhancing Shared Mental Models: A Systematic Review and Meta-Analysis of Randomised Controlled Trials

Collective action underpins the safety, health, security, and success of societies and their citizens worldwide. As one representation of collective action, the science of teamwork offers important insights into the capacities, states, and processes that optimise the coordinated efforts of two or more individuals (e.g., team composition, cognitions, debriefing; Driskell et al., 2018). From a cognitive standpoint, there exists substantial interest across numerous occupational settings in the concept of shared mental models (SMM) as an essential component of effective teamwork. Formally defined, SMM reflect degree of overlap across team members regarding the knowledge structures that characterise members' roles and responsibilities, the tasks and procedures that need to be implemented, and how members work interactively to achieve a common objective (Cannon-Bowers et al., 1993). Essentially, when members of a team 'are on the same page' regarding key taskwork and teamwork elements they are best positioned to anticipate and react effectively to situational demands as well as the needs, duties, and actions of their team members and, in so doing, deliver high performance (Cannon-Bowers & Salas, 2001). Narrative reviews (Mohammed et al., 2010) and statistical syntheses of empirical data (DeChurch & Mesmer-Magnus, 2010a, 2010b) support these expectations regarding the salience of SMM as a determinant of a range of collective states, processes, and outcomes. Yet our knowledge of factors that can enhance SMM remains underdeveloped and therefore insufficient for theory development and practice, primarily because the evidence is fragmented across diverse scientific disciplines (e.g., organisational behaviour, education) and occupational contexts (e.g., defence, healthcare). Knowledge integration is necessary for expanding conceptual discussions beyond the dominant focus on SMM as a determinant of key team processes and outcomes to incorporate knowledge on its antecedent factors. For these reasons, we aimed to characterise existing knowledge on the antecedents of SMM by narratively and statistically synthesising work that has utilised team development interventions (TDI) as a means by which to enhance SMM.

Conceptual Foundations

Foundations of Shared Mental Models

Scholars have proposed and tested numerous team cognition concepts that reflect knowledge-building processes (e.g., process-based group learning, information sharing) or emergent mental representations (e.g., team situation awareness, transactive memory systems) of the content, structure, and interrelationships among knowledge components that underpin collective action (Mohammed et al., in press). SMM and transactive memory systems are two of the most studied team cognition concepts (DeChurch & Mesmer-Magnus, 2010a; Mesmer-Magnus et al., 2017). SMM, as the name suggests, reflect knowledge structures of the task and teamwork elements that overlap in their representation across members of the collective (Cannon-Bowers et al., 1993). In contrast, transactive memory systems reflect the division of unique knowledge among individual members and a collective awareness of how that information is distributed among the team (Ren & Argote, 2011; Wegner, 1987). In essence, the distinction is akin to 'members being on the same page' versus 'knowing who knows what'. SMM and transactive memory systems differ regarding breadth and depth of information (DeChurch & Mesmer-Magnus, 2010a). First, SMM cover a broader range of content (e.g., taskwork, teamwork, time elements) than transactive memory systems (e.g., knowledge of who knows what; Mohammed et al., 2015). Second, SMM encompass knowledge content and structure among the individual elements (DeChurch & Mesmer-Magnus, 2010a; Resick et al., 2010). Unsurprisingly, breakdowns in team coordination and therefore collective effectiveness often result from failures in members being on the same page regarding taskwork and teamwork knowledge (Bearman et al., 2010; Rafferty et al., 2010).

In addition to the content of mental models (taskwork and teamwork), the form and property of cognition are two important conceptual details for the operationalisation of SMM. The form of cognition has implications for the ways in which SMM are elicited (Rentsch et al., 2008). Perceptual cognitions represent individual members' self-reported evaluations (e.g., beliefs, attitudes) of the cognition of the team as a collective or as individual members (Rentsch et al., 2008). This form of cognition is typically assessed via individuals' ratings of the key taskwork and teamwork elements (e.g., declarative or procedural knowledge) in terms of their similarity, accuracy, or overall effectiveness (DeChurch & Mesmer-Magnus, 2010a). In contrast, structured forms of cognition reflect the organisational properties of knowledge availability among team members, and the degree of similarity or sharedness of these patterns between team members (Mesmer-Magnus et al., 2017). This form of cognition is typically assessed via pairwise comparisons and concept mapping, which are statistically interrogated and summarised via programs or frameworks such as Pathfinder, network analysis, or computational modelling. Mental models have two main properties, namely their similarity and their accuracy among members, both of which are considered compositional properties because they capture convergence in perceptions among members (DeChurch & Mesmer-Magnus, 2010a). Similarity reflects the extent to which individual members' mental models overlap or converge with other members of the team (Cannon-Bowers et al., 1993; Rentsch et al., 2008), whereas accuracy characterises the degree to which mental models adequately cover the essential elements of a performance domain (Edwards et al., 2006). Operationalised in this manner, teams may possess a high level of similarity in the mental models among individual members, yet their representation of the problem and performance space may be suboptimal. Thus, teams who possess high degrees of similarity and accuracy in their mental models are expected to perform optimally (Resick et al., 2010).

To What Extent Do TDI Enhance Shared Mental Models?

Narrative reviews (Mohammed et al., 2010) and statistical syntheses of empirical data (DeChurch & Mesmer-Magnus, 2010a, 2010b; Mesmer-Magnus et al., 2017; Niler et al., 2021) support expectations regarding the importance of SMM for a range of collective states (e.g., motivational elements such as cohesion and confidence), processes (e.g., behaviours that underpin task-goal accomplishment such as mission analysis, monitoring goal progress), and outcomes (e.g., indices of performance outcomes that reflect effectiveness and efficiency, such as number of objectives achieved or time to complete tasks). Thus, knowledge on how best to enhance this emergent concept has the potential to inform theory and practice. SMM are underpinned by knowledge of individual members' mental models that manifest at the collective level as an emergent concept when individuals' cognitive representations of the content, structure, and interrelationships of key elements of the task and environment are shared among team members (Kozlowski et al., 2006). Conceptualised as an emergent concept, therefore, SMM can be enhanced via individual-level (e.g., role clarity) or team-level inputs (e.g., norms) and team-level processes (e.g., performance monitoring) that involve some degree of interaction among members. Efforts designed to optimise such inputs and processes are typically characterised as TDI, which formally defined, represent "actions taken to alter the performance trajectories of organisational teams" in ways that foster returns to, maintenance of, enhancement of, and diversification of the healthy functioning of the unit (Shuffler et al., 2018, p. 689). Unsurprisingly, TDI have received widespread attention for the enhancement of SMM (Mohammed et al., 2010).

Evidence supports the utility of targeted (e.g., leadership) or multicomponent TDI as an effective means by which to optimise team functioning across a variety of domains such as the military, education, healthcare, and sport (e.g., Klein et al., 2006; Lacerenza et al., 2017; Lines et al., in press; McEwan et al., 2017). An important question for the science and practice of SMM development therefore is not whether TDI enhance this element of team cognition, but rather by how much they effect change or development. From a theoretical standpoint, knowledge of the magnitude of an effect via a point estimate and/or range of plausible values is essential for generating high-quality theoretical summaries of human phenomena because directional hypotheses that exclude specification of statistical benchmarks are imprecise and therefore evade falsification (Edwards & Christian, 2014). We require knowledge of the smallest effect size upon which to make judgements about the theoretical meaningfulness of empirical findings (Lakens, 2014). Statistical syntheses of existing bodies of work, particularly when they minimise publication bias via the incorporation of unpublished evidence, represent one key approach to estimating the existence and robustness of effects for theory development (for others, e.g., see Anvari & Lakens, 2021). Thus, our first contribution summarises causal evidence regarding the magnitude and direction of effects that is necessary for theory development regarding the antecedents of SMM among teams. We expected the overall pooled estimate regarding the effectiveness of TDI on SMM to show that they represent a small yet worthwhile approach to getting team members ‘on the same page’.

Which Type of TDI are Most Effective for Enhancing Shared Mental Models?

An overall estimate tells us little about the types of TDI that are most effective for enhancing SMM. The answer to this question is of theoretical and practical significance. Theoretically, knowledge of which type of TDI are most effective for enhancing SMM can clarify the necessary and sufficient conditions of antecedent factors. Practically, organisations can utilise such knowledge to invest strategically in interventions that are most likely to offer them the best outcomes for enhancing SMM. Of course, small effects can be practically meaningful when considered in conjunction with contextual factors, such as the resources required to produce those effects (Prentice & Hall, 1992). Meta-analyses are well suited to allowing comparisons between different intervention types, given the burdensome nature of experimental trials (e.g., financial resources) means it is often impractical to examine multiple types of TDI within a single primary study. Thus, we set out to explore the types of

TDI that have been used to enhance SMM, and empirically estimate the direction and magnitude of effectiveness of different types of TDIs.

Given the diversity of ways by which a team's trajectory of functioning can be altered, we leveraged Shuffler and colleagues' (2018) integrative framework of TDI to calibrate findings from a diverse body of work so that our findings contribute meaningfully to theory on SMM development. This integrative framework of TDI encompasses 10 categories of TDI according to the classic input-process-output (IPO) framework. With regard to targeting *team inputs*, one might identify key knowledge, skills, and abilities required for successful job and task performance (task analysis); compose members of a team in ways that align knowledge, skills, and abilities of individuals with the requirements for collective effectiveness (team composition); specify and structure roles, tasks, and goals of the collective to align with team and organisational objectives (team work design); and generate and document explicit guidelines, rules, and policies that govern what members do and how they do it (team charters). Observing, recording, and evaluating actions that precede or influence the attainment of collective goals and metrics that quantify goal attainment (team performance monitoring and assessment) are used to address *team processes*. Interventions in which teams are guided to exert conscious and intentional effort towards evaluating and learning from experiential activities (team debriefs) primarily target *team outcomes*. Some TDI target multiple IPO elements, including in/formal efforts to foster social relations among members and clarify roles (team building); formal learning experiences in which teams participate in structured activities guided by a curriculum and pre-set objectives targeting key knowledge, skills, or competencies for collective effectiveness (team training); coaching from people external to the team on how members coordinate their resources for collective objectives (team coaching); and efforts to enhance capabilities of leaders for defining the collective's future (e.g., objectives, vision) and organising members structurally and procedurally toward this end state (team leadership). Within each of these categories, TDI designed to enhance SMM might do so by targeting cognitive representations of the content, structure, and interrelationships of key elements of the task and environment in/directly.

What Components of TDI are Most Effective for Enhancing Shared Mental Models?

Knowledge of which types of TDI are most effective for enhancing SMM would address an important gap in our understanding of the antecedents of SMM, yet the limited resolution of such categorisations means we are restricted to a broad, overall snapshot for theory and practice. Within the 10 different categories of TDI outlined by Shuffler et al. (2018), there are a multitude of ways by which scholars and practitioners might operationalise those programs into practice. In terms of work design interventions, for example, organisations might enrich (e.g., increase autonomy or decision-making) and/or enlarge (e.g., job rotation, increase task variety) jobs with or without employee input, enhance employee's perceptions of the significance of their work, empower teams with job autonomy, or implement system-wide changes (e.g., information and communication systems) (Knight & Parker, 2021). Although TDI are operationalised in diverse ways, particularly regarding the active ingredients or components that drive change in target outcomes, this degree of specificity is often absent from statistical syntheses of the effectiveness of TDI on team processes and outcomes (e.g., Lacerenza et al., 2017; Lines et al., in press; McEwan et al., 2017). From a theoretical standpoint, knowledge of 'what' drives change is fundamentally important for clarifying the mechanisms of action or 'how' an intervention works, that is, the nature of the change or development process through which the active ingredients of an intervention affect target outcomes (Connell et al., 2019). Scholars have relied on numerous different theoretical or conceptual foundations to explain why and how TDI might influence SMM, such as social information processing theory (Salancik & Pfeffer, 1978), social impact theory (Latane, 1981), information processing theory (Schroder, Driver, & Streufert, 1967), and functional leadership theory (Kozlowski, Gully, Salas, & Cannon-Bowers, 1996) just to name a few. However, individual studies tell us little about the robustness of such theoretical expectations, or their relative effectiveness compared to alternative yet complementary explanations. For these reasons, we set out to characterise the key components of TDI employed to enhance SMM and empirically examine their meaningfulness as moderators of the overall pooled effect to shine a spotlight on the which elements offer the greatest benefit.

Overview of Contributions

Via a narrative and statistical review of the literature on the antecedents of SMM, we generate new insights that illuminate the building blocks for theory development and evidence-based practice regarding the enhancement of SMM among teams in three ways. First, we provide the first statistical summary of the *magnitude* of the effect of TDI on SMM via a meta-analytic summary of lab and field experiments across over two decades of research (1996-2020). We focused our efforts on primary studies in which teams were randomly assigned to treatment and comparator conditions because experimental designs are considered the 'gold standard' for inferences regarding causality (Podsakoff & Podsakoff, 2019). This contribution is important because there is a reliance on *directional* propositions within theory in the psychological sciences; these propositions, which neglect the expected magnitude of effects, are suboptimal because any nonzero statistical summary (e.g., d = .01) can be considered supportive of theoretical propositions and therefore evade falsification (Edwards & Christian, 2014). Second, primary experimental studies of the antecedents of SMM often focus on individual determinants in isolation (e.g., knowledge exchange); our comprehensive review of the literature permits an evaluation of multiple determinants assessed across individual studies that is necessary for articulating the conceptual underpinnings required for a holistic theory of SMM. Third, we empirically test substantive (e.g., type of TDI) and methodological (e.g., operationalisation of SMM) moderators of the effectiveness of TDI for optimising SMM, several of which have yet to be tested within individual studies and are resource intensive to test in primary research. Such tests provide essential knowledge for interpretations regarding the extent to which the effectiveness of TDI for enhancing SMM generalises across samples and contexts and, in so doing, generate new insights into boundary conditions for theory and practice. Knowledge of which types of TDI and active components of those interventions are most beneficial, for example, can shine a spotlight on theoretical explanations that might augment knowledge on the antecedents of SMM. Practically, theoretical propositions that encompasses expectations of magnitude of effects alongside knowledge of methodological and substantive moderators of effectiveness offers a degree of precision that can be used to guide decisions regarding the investment of organisational resources and efforts (e.g., prioritisation of one intervention over another).

Methods

We prospectively registered the protocol for this systematic review and meta-analysis on 8th June 2020 via the Open Science framework (OSF; https://osf.io/yzdxn/), using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses-Protocol template (PRISMA-P; Shamseer et al., 2015). We report the results of this work in accordance with the PRISMA 2020 guidelines (Page et al., 2021).

Literature Search

RL conducted a comprehensive search from inception until 9th June 2020 via the following electronic databases: Web of Science, PsycInfo, Scopus, Business Source Complete, and ProQuest Dissertations. We utilised the following combination of search terms across these databases: ("mental models" OR "situation awareness" OR "cognitive map" OR "knowledge map" OR "knowledge structure" OR "cognitive structure") AND (team OR group OR collective OR shared) OR ("team cognition" OR "team knowledge" OR "shared cognition" OR "shared knowledge" OR "collective cognition" OR "collective knowledge") AND (intervention OR trial* OR experiment* OR train* OR development OR program*). Full details of the search protocol are provided in our registered PRISMA-P document. We also conducted a manual backward and forward search in which we examined the reference lists of eligible studies and all papers that had cited the final sample of eligible studies.

Eligibility Criteria

We considered studies for inclusion when they: (i) tested the effectiveness of a TDI, training program, or experimental manipulation and incorporated the assessment of SMM within teams; (ii) randomised teams into experimental and control conditions; (iii) included an adult population (18 years of age or older); and (iv) provided sufficient information in the published paper to extract the required data for effect size calculations, or this information was available via the authors. Papers were excluded when: (i) they utilised non-experimental designs (e.g., non-random assignment); (ii) they excluded an assessment of SMM as an outcome variable; (iii) the article was written in any language other than English; (iv) the full-text was unavailable via our university library subscriptions or directly from the corresponding author; (v) the information required for analysis was unavailable in the document and via direct requests from the corresponding author; (vi) the article was a protocol, guideline, review, or a duplicate (e.g., thesis that was subsequently published); or (vii) the results were published as a conference abstract rather than a full-text (e.g., pre-print).

Population

Teams composed of adults aged 18 years and older were the focus of this systematic review and meta-analysis. For the purposes of this review, teams are defined as a collective of two or more individuals who work interdependently for a specified timeframe to achieve a common and valued outcome or objective (Sundstrom et al., 1990). We considered teams who were sampled naturalistically from occupational contexts (e.g., sport, emergency services) or brought together for the purposes of an empirical study (e.g., student teams).

Intervention

Our focus was on TDI directed towards the enhancement of SMM. For this review, we used the definition of TDI as “actions taken to alter the performance trajectories of organisational teams" in ways that foster returns to, maintenance of, enhancement of, and diversification of the healthy functioning of the unit (Shuffler et al., 2018, p. 689).

Comparators

We considered all types of comparators, including waitlist controls, no contact controls, placebo control, and active controls.

Outcomes

The primary outcome of interest was SMM. Formally defined, SMM are "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 behaviour to demands of the task and other team members" (Cannon-Bowers et al., 1993, p. 228).

Article Screening

All references identified via the electronic database search were imported into a citation management program (Endnote) and subsequently exported into Research Screener (https://researchscreener.com), a web based program which semi-automates the screening process using a machine-learning algorithm without sacrificing accuracy. Research Screener initially prioritises abstracts, using deep learning and natural language processing methods, based upon a selection of seed articles which are representative of eligible articles. We used four seed articles for the purposes of this review (Boies & Fiset, 2018; Gurtner et al., 2007; Marks et al., 2002; Toader et al., 2019). We selected these seed articles because they approximated the breadth of research in the area and a broad temporal span of research we were interested in examining in this meta-analysis. Research Screener presents the 50 most relevant article abstracts from which the researcher makes a judgment as to whether they are (ir)relevant and therefore to be flagged for retention and full text screening. Research Screener then re-ranks the remaining articles by relevance based upon the selection of flagged articles, and presents the next 50 article abstracts. Guided by simulation evidence with Research Screener (Chai et al., 2021), RL screened 50% of the total sample of abstracts.

This threshold is conservative based on simulation work; RL flagged no articles for full text review in the final 36 rounds of 50 articles (*n* = 1,800). A second reviewer [BH] used Research Screener to screen 20% of the total initial sample (*N* = 1,600); RL and BH discussed discrepancies and when a decision was unable to be made based upon the title and abstract the paper was retained for full text review. Three reviewers [RL, BH, and SN] conducted the full text review stage separately, with a separate member of the research team [DG] judging the eligibility of studies when there was disagreement. The categories for study exclusion are summarised in the PRISMA flow chart (see Figure 1).

Data Extraction

RL carried out the data extraction from the final sample of eligible papers; DG assessed a random sample of 30% of data extraction forms to check accuracy and consistency, with discrepancies discussed and revisited across the entire pool of eligible studies. In cases where the data was unavailable in the full text, we sent an e-mail to the corresponding author requesting the required data. We initially planned to send requests out on two occasions separated by two weeks; however, considering the global coronavirus (COVID-19) pandemic the decision was made to send a third request, again separated by two weeks. The complete data extraction sheet is located on the OSF project page (https://osf.io/yzdxn/).

Statistical Analyses

Coding of Studies

We utilised a coding system to record the key characteristics of the interventions, outcome variables, studies, and samples. SMM were characterised and measured in several ways across the final sample of studies. We coded the nature of interventions according to the (i) 10 team development intervention categories proposed by Shuffler et al. (2018) and (ii) primary component of the TDI (described below in ‘Moderator, Sensitivity, and Meta-Bias Analyses’). We also coded whether technology was used in the provision of the intervention (i.e., present vs absent). In terms of SMM, we coded the property (i.e., similarity or accuracy) and content (i.e., teamwork, taskwork, or combined team and taskwork) of SMM being assessed. We also coded all outcome variables in terms of the way in which they were assessed (i.e., self-reported, objective, or researcher assessed). For study characteristics, we coded type of comparator (i.e., active, no treatment, and treatment as usual), publication type (i.e., peer reviewed paper vs thesis), and risk of bias (see below for full details). Sample characteristics coded included total sample size (continuous variable), number of teams (continuous variable), team size (continuous variable; the mean value was used for the four studies that reported a range), percentage female (continuous variable), and mean age (continuous variable). Continuous variables were mean centred prior to moderation analyses.

Calculation of Effect Sizes

We utilised pooled standardised mean differences due to our expectation of heterogeneity between studies. We computed the standardised mean difference as a summary measure of effect size (ES) to quantify the effect of the intervention in comparison with comparators, thereby enabling the synthesis of an outcome variable (e.g., model similarity) even when eligible studies employed different measures. Hedges' *g* was used as the ES unit because it accounts for relative sample sizes (Lakens, 2013). We computed ES from the extracted means, standard deviations, and the number of teams in experimental and control conditions. When the study design included two or more assessment points after the delivery of the intervention, we utilised the distal (final) assessment point for the calculation of the ES. In cases where means and standard deviations were unavailable within the paper or via data requests from the authors, we used *F* statistics, *t* scores, and *p* values to calculate ES (Borenstein et al., 2009). The final dataset is available on the OSF project page (https://osf.io/yzdxn/). When required, we reversed coded effects so that a positive sign represents the scenario where the intervention improved or resulted in greater similarity or accuracy in SMM for intervention groups relative to their comparators.

Statistical Synthesis of Effect Sizes

Two-thirds of eligible studies (66.6%) included two or more ES (e.g., multiple dimensions of SMM, multiple comparator groups), which violates the assumption of independence among effects that is core to most meta-analytic techniques (Cheung, 2014; Van den Noortgate et al., 2013). Treating dependent effects as independent underestimates the standard errors of the overall effect and therefore biases any statistical conclusions from the analysis (López-López et al., 2018). Therefore, we used a three-level random effects model with restricted maximum-likelihood estimation to account for dependency among ES from within the same study (Cheung, 2019). The three-level random effects meta-analytical model allows for the decomposition of sampling variance of individual effects (level 1) as well as variance within (level 2) and between studies (level 3; Cheung, 2014). We first estimated the overall effect of TDI on all SMM outcomes within a single model. We utilised log-likelihood ratio tests to determine the magnitude and significance of heterogeneity within (level 2) and/or between (level 3) studies (Assink & Wibbelink, 2016). When the variance within or between studies is significant ($p < .05$), the distributions of the ES are considered heterogeneous and therefore worthy of further investigation via moderator analyses[1]. We conducted all statistical analyses detailed below using the *metafor* (Viechtbauer, 2010) and *metaviz* (Kossmeier et al., 2020a) packages in the R statistical platform (R Development Core Team, 2019).

Moderator, Sensitivity, and Meta-Bias Analyses

We examined several possible moderators of the effect of TDI on SMM, including the type of TDI (i.e., task analysis, composition, work design, charters, performance monitoring and assessment, debriefing, team building, team training, coaching, and leadership), TDI

[1] We also conducted meta-analytic investigations of the individual components of SMM, namely similarity of SMM, accuracy of SMM, teamwork SMM, taskwork SMM, and combined team and taskwork SMM. Interested readers can view the results and output of these analyses on the OSF project page (https://osf.io/yzdxn/).

components (i.e., alterations to human capital resources, critical thinking, inspirational motivation, interpositional knowledge or training, knowledge exchange, knowledge representation and structure, performance feedback and guidance, reward structures, role clarity, sensegiving or sensemaking, shaping knowledge, and strategy planning), property of the SMM (i.e., similarity, accuracy), type of SMM (i.e., taskwork, teamwork, and combined), and outcome measurement type (i.e., self-reported, researcher assessed, objective). We also conducted moderator analyses to assess the robustness of the findings across study design characteristics, including number of teams, team size, percentage of female participation, presence or absence of technology, type of comparator (e.g., active, no treatment), and risk of bias. Moderator analyses provide essential information for understanding boundary conditions and auxiliary hypotheses for theory development (Hagger, Gucciardi, & Chatzisarantis, 2017). For the sensitivity analyses, we examined the outlier influence in two ways, namely effects whose residuals were more than three standard deviations times the mean or with a Cook’s distance more than three times the mean (Viechtbauer & Cheung, 2010). For meta-bias, we examined the influence of sample size, publication type (i.e., peer reviewed vs thesis), and study quality (i.e., risk of bias) on the overall pooled effect. We also calculated the multilevel extension of Egger’s test as an approximation of funnel plot symmetry (Fernández-Castilla et al., 2021) and examined ‘sunset’ funnel plots that incorporate the statistical power of each individual study (Kossmeier et al., 2020b).

Statistical Heterogeneity

We estimated statistical heterogeneity in three ways. First, I^2 provides the proportion of total variance of the effect due to statistical heterogeneity as opposed to sampling error (0%-40% = might not be important; 30%-60% = may represent moderate heterogeneity; 50%-90% = may represent substantial heterogeneity; and 75%-100% = considerable heterogeneity; Higgins et al., 2003). Second, we decomposed variance according to within study (τ^2_{within}) and between study ($\tau^2_{between}$) levels, whereby zero reflects no heterogeneity.

Third, we calculated a complementary assessment of between-study heterogeneity using 95% prediction intervals to compute the range in which the effect of estimates of future studies will lie (IntHout et al., 2016).

Confidence in Cumulative Evidence

The quality of evidence and strength of recommendations were assessed using the GRADE approach (Guyatt et al., 2008; see https://gradepro.org/) across four domains: risk of bias, inconsistency, indirectness, and imprecision. The revised Risk of Bias tool (RoB2; Sterne et al., 2019) was used to assess the risk of bias of each outcome across five domains: randomisation process, deviations from intended intervention, missing outcome data, measurement of the outcome, and selection of the reported result.

Narrative Analysis of Intervention Content

We also narratively synthesised the characteristics of effective interventions, that is, interventions where the ES was larger than g = .40 and the lower bound confidence interval was above zero because it reflects a medium effect within the context of social psychological research (Lovakov & Agadullina, in press). Of particular interest was the nature of those interventions (e.g., TDI category and components) identified as effective in enhancing SMM in teams. In cases where multiple effects were present for a single study, we classified interventions as 'effective' when 100% of effects met our criterion, and 'promising' when at least 50% of effects satisfied the criterion.

Deviations from Pre-Registered Protocol

We deviated from our pre-registered protocol in three ways. First, we originally planned to send two reminder emails, each 14 days apart when the information required to compute ES was unavailable in the full text. Due to the COVID-19 pandemic, many people were working from home with restricted access to offices and increased workload translating face-to-face pedagogy to online formats, so we decided to send a third request two weeks later. Second, we explored the effect of residual outliers and Cook's distance outliers on the overall pooled effect, as additional pieces of information regarding potential meta-bias. Third, regarding the narrative synthesis of intervention content, we (i) changed our criterion for assessing interventions as effective to incorporate the confidence interval alongside the point estimate to take into consideration the precision of the effect, and (ii) added a category of 'promising' to account for scenarios where authors utilised multiple assessments of SMM.

Results

Literature Search Overview

An overview of the search and selection process is depicted in Figure 1. In total, 39 studies were identified from the electronic database search as fulfilling the eligibility criteria, with an additional seven eligible studies identified via backward and forward scanning. Of the 46 studies identified, the data needed to calculate ES were unavailable in 10 studies resulting in a final sample of 36 studies included in the meta-analysis. The 36 eligible studies were published between 1996 and 2019, and yielded a total of 160 ES of which 131 were considered relevant for inclusion. The final sample included 6,209 participants who were members of 1,912 teams, with a mean team size of 3.4 members. Participants were, on average, 23.1 years of age and females accounted for 53.1% of participants. An overview of these studies is provided in Supplementary Table 1 (http://bit.ly/smm-supptable1).

Overall Effect of TDI on Shared Mental Models

The overall effect (131 ES, $k = 36$) was moderate in magnitude ($g = .61$, $SE = .10$, 95% CI [.41, .82]; see Table 1 and Figure 2). The 95% prediction intervals suggest that for a new study there is a 95% chance that the effect will be between -0.66 and 1.89 (Hedges' *g*). The log-likelihood ratio tests (LRT) demonstrated significant variance in ES within (level 2; LRT = 20.81, $p < .001$) and between studies (level 3; LRT = 38.11, $p < .001$), explaining 33.18% and 43.1% of the variance, respectively (see Table 1). Due to substantial heterogeneity among the effect sizes ($I^2 = 76.19\%$; Higgins et al., 2003), we carried out moderator analyses to examine factors that may explain the variance between studies.

Sensitivity Tests

Two effects within two studies produced residuals exceeding three standard deviations (Andres, 2013; Burke, 2000); their removal reduced the overall effect of interventions on SMM by 0.04 (g = .57, SE = .09, 95% CI [.38, .75]). Eight effects within five studies had a Cook's distance over three times the mean (Andres, 2012; Andres, 2013; Burke, 2000; Crespin, 1996; Dalenberg et al., 2009; Ouverson, 2019); exclusion of these effects resulted in a reduction in the overall pooled effect by 0.12 (g = .49, SE = .08, 95% CI [.32, .65]). Collectively, these exploratory analyses indicated that the influence of outliers or influential studies was small-to-moderate in nature, though the overall conclusion regarding the effectiveness of interventions remained the same.

Moderator Effects

Results of the moderator effects are provided in Table 2. Only outcome assessment method moderated the overall effect, $F(2, 128) = 6.46$, $p = .002$, whereby the effectiveness of TDI was strongest for researcher assessed outcomes (2 effects; g = 3.38, 95% CI [1.81, 4.96]) followed by self-reported outcomes (126 effects; g = 0.59, 95% CI [0.37, 0.79]), with objective outcomes being statistically inconsequential (3 effects; g = 0.30, 95% CI [-0.46, 1.06]). Inclusion of the outcome assessment method with the overall model produced a significant reduction in heterogeneity, Cochran's $Q(130) = 443.76$, $p < .001$, although residual heterogeneity remained statistically significant, $Q_E(128) = 424.88$, $p <.001$.

Meta-Bias Assessment

The multilevel extension of Egger's test, F (1,129) = 5.39, p = .02, suggested asymmetry in the funnel plot. Visual examination of the funnel plot showed that the effects were unevenly distributed around the mean effect, with a slight propensity for smaller studies to produce stronger effects on SMM development (see Figure 3). Of the 131 effects, 30 (22.9%) were located outside of the 95% confidence interval. The sunset enhanced funnel plot (see Figure 3) showed the median power of primary studies, assuming an effect of g =

.50 (p = .05), is around 27.5%; the true underlying effect size for realising median power levels of 33% (g = 0.56) and 66% (g = 0.87); and a low probability of replication (R-index = 20%). No publication biases were observed for publication status (p = .843), sample size (p = .915), or study quality (p = .515).

Risk of Bias

An overall summary can be seen in Figure 4, with a breakdown of each outcome provided in Supplementary Table 2 (http://bit.ly/smm-supptable2). In total, 52 outcomes received a rating of low risk of bias; the remaining 79 were rated as having some concerns. Of the five risk of bias categories, the only source of bias within outcomes was related to the selection of the reported results. The main reason for studies receiving a rating of some concerns in this category was a lack of a data analysis section within the paper; the inclusion of a pre-specified analysis plan would have resulted in all outcomes examined in this meta-analysis receiving a low risk of bias rating.

GRADE Assessment

A summary of all assessments is presented in Supplementary Table 3 (http://bit.ly/smm-supptable3). For all analyses the level of evidence was downgraded for the category risk of bias, due to most outcomes receiving a rating of some risk of bias. Similarly, all analyses were downgraded for inconsistency as there were high levels of heterogeneity in effect sizes across each of the meta-analyses (I^2 > 68.9%). For the category of imprecision, the level of evidence for SMM accuracy was downgraded due to wide 95% confidence intervals around the main effect. Overall, the level of evidence was graded as very low quality for one outcome (SMM accuracy) and low quality for five outcomes (overall SMM, SMM similarity, teamwork SMM, taskwork SMM, and combined SMM).

Narrative Synthesis of Effective TDI Interventions

All but three experiments tested the effectiveness of TDI on SMM with university students. We classified 11 TDI as effective, 10 as promising, and 15 as non-effective.

Effective interventions covered six of the ten TDI categories, namely team leadership (*n* = 3), team training (*n* = 2), work design (*n* = 2), team building (*n* = 2), team debriefing (*n* = 1), and performance monitoring and assessment (*n* = 1). In terms of TDI components, effective interventions targeted sensemaking or sensegiving (*n* = 3), knowledge representation and structure (*n* = 2), shaping knowledge (*n* = 1), strategy planning (*n* = 1), critical thinking (*n* = 1), interpositional knowledge or training (*n* = 1), knowledge exchange (*n* = 1), and performance feedback and guidance (*n* = 1). Of the 11 effective studies, we assessed six as low risk of bias and five with some concerns. Studies involving effective interventions employed a no treatment comparison group only (*n* = 11), primarily did not use technology to deliver the intervention (*n* = 7), published the results in a peer-reviewed journal (*n* = 9), relied on self-reported assessments (*n* = 9), and assessed varying aspects of the content (taskwork = 5, teamwork = 4, and combined = 2), property (similarity = 15 effects, accuracy = 5 effects), and form (perceptual = 7 effects, structured = 13 effects) of cognition. Finally, effective frameworks were informed theoretically by functional leadership theory (*n* = 2), cognitive flexibility theory (*n* = 1), mathematical graph theory (*n* = 1), media synchronicity theory (*n* = 1), process-oriented theory (*n* = 1), social impact theory (*n* = 1), or a combination of theories including information sharing, expertise, computer mediation, and knowledge objects (*n* = 1); in three cases, there was no explicit mention of guiding theory.

Discussion

Shared mental models are an essential component of effective teamwork (Boies & Fiset, 2018; DeChurch & Mesmer-Magnus, 2010a; Mohammed et al., 2010). Narrative reviews (e.g., Mohammed et al., 2010) and meta-analytical evidence (e.g., DeChurch & Mesmer-Magnus, 2010a; Mesmer-Magnus et al., 2017; Niler et al., 2021; Turner et al., 2014) support the benefits of SMM as a key determinant of team processes (e.g., goal setting), emergent states (e.g., collective efficacy), and outcomes (e.g., performance). However, this integrative knowledge is currently limited to reporting on the outcomes of SMM or team cognition broadly. We addressed this gap in the literature by narratively and statistically synthesising experimental work in which scholars have examined the effectiveness of TDI on the enhancement of SMM. Methodologically, we focused on randomised controlled trials or experiments because they provide the strongest evidence for inferences regarding causality. In so doing, we contribute new knowledge on the antecedents of SMM that can inform theory regarding the nomological network of this concept.

To What Extent Do TDI Enhance Shared Mental Models?

We expected TDI overall would represent a worthwhile approach to enhancing the likelihood of team members being 'on the same page', yet the magnitude of effectiveness would likely be small because of the multiplicity of potential influences on SMM as an emergent concept (e.g., team-level inputs or processes) and diversity of ways by which a team's trajectory of functioning can be altered (Shuffler et al., 2018). Our *directional* expectation regarding the usefulness of TDI was supported; the *magnitude* (g = .61) observed can be considered moderate-to-large relative to distributions observed in social psychology (Lovakov & Agadullina, in press). This magnitude is comparable to meta-analytic values reported for the specific TDI of teamwork interventions (McEwan et al., 2017), leadership training (Lacerenza et al., 2017), team training (Hughes et al., 2016), and team reflections (Lines et al., in press). Sensitivity and meta-bias analyses generally supported the robustness of the overall effect of TDI on SMM. Collectively, therefore, our findings support the idea that SMM are malleable and TDI are an effective means by which to enhance them. This finding is important because SMM are an essential prerequisite or contributor to high-quality team processes and outcomes (e.g., Mesmer-Magnus et al., 2017). Generally, therefore, our findings suggest that moderately sized effects might be considered the minimum value upon which to make judgements about the theoretical meaningfulness of empirical findings regarding the determinants of SMM. From a practical standpoint, however, it is important to consider effect size magnitude together with contextual factors, such as the resources required to produce those effects (Prentice & Hall, 1992).

Despite our encouraging findings regarding an overall positive effect that is moderate-to-large in magnitude, prediction intervals (IntHout et al., 2016) indicated a 95% chance that the effect of a future test of the effectiveness of TDI on SMM will lie between -0.66 and 1.89. In other words, if TDI are applied in a new study or population similar to those included in this meta-analysis then it is plausible that the effect could differ considerably from the overall pooled effect observed here; the plausible range of effects incorporates scenarios where effectiveness is null or even detrimental to the enhancement of SMM. This finding is theoretically important because it illuminates the first piece of empirical evidence regarding the extent to which the effectiveness of TDI for enhancing SMM may generalise across contexts. Returning to the primary question, therefore, our meta-analytic data indicated that TDI can be effective for enhancing SMM, yet their overall effectiveness (or magnitude of effect) likely depends on methodological, substantive, or contextual factors.

From a scientific standpoint, the overall strength of the body of evidence synthesised in a narrative or statistical synthesis is an important consideration for inferences regarding the robustness of theoretical propositions (Guyatt et al., 2008). Our assessment indicated low certainty in the quality of the evidence that contributed to the meta-analytic estimates reported here. Key considerations in this regard included imprecision in the estimates (e.g., wide confidence intervals of estimates across studies, non-overlap of confidence intervals within and between studies) and risk of bias (e.g., limited or absence of detail on the data analysis protocol). Insufficient detail on key elements of the scientific process in published papers is a well-known issue in the organisational behaviour and psychology literatures (e.g., Aguinis et al., 2018). Researchers are advised to consult general guidelines (e.g., Aguinis et al., 2019; Aguinis et al., 2020) and design-specific checklists (e.g., CONSORT guidelines for clustered randomised trials, Campbell et al., 2012; CONSORT-SPI for social and psychological interventions, Montgomery et al., 2018) when planning a study and reporting the results; the Equator Network is an excellent resource for checklists (https://www.equator-network.org).

Perhaps most salient for the overall assessment of the strength of evidence regarding the effectiveness of TDI for optimising SMM was the low statistical power of primary studies, even to detect a moderate-to-large effect (d = .50; see Figure 3). The median statistical power of studies incorporated in our statistical synthesis is roughly consistent with estimates obtained from 200 meta-analyses on psychological science (Stanley et al., 2018). The sensitivity of a design and test combination to detect an effect of some magnitude (e.g., smallest effect size of interest or practical value, Lakens & Evers, 2014; minimally clinical important difference, Jäschke, Singer, & Guyatt, 1996) is essential to making reliable inferences regarding the strength of evidence of individual studies. Within the context of meta-analyses, low statistical power of primary studies combined with high heterogeneity is a primary cause of failed replications (Stanley et al., 2018). Statistical power for research on teams is complicated because one needs to consider two levels corresponding to individual members (level 1) who comprise a single team (level 2) as well as fixed and random effects. Simulations indicate that there is much more to be gained by increasing sample size for teams (level 2) rather than individuals (level 1) within the context of multilevel modelling (e.g., Arend & Schäfer, 2019). Thus, scholars interested in testing TDI on SMM in future research are advised to invest effort and resources towards maximising the sample of teams included in their work.

Which Type of TDI are Most Effective for Enhancing Shared Mental Models?

Between-study heterogeneity is an important consideration for the replicability of scientific knowledge (Stanley et al., 2018). Some degree of heterogeneity between studies is to be expected because of variations in interventions, measures, designs, methods, and characteristics of the population from which researchers sample their participants. Even in scenarios where researchers have controlled for such factors where possible, there can remain notable heterogeneity between studies (e.g., Hagger et al., 2016). We considered the differential effectiveness of different categories of interventions (Shuffler et al., 2018) on the enhancement of SMM as one potential explanation of between-study heterogeneity. Moderator analyses indicated that the type of TDI collectively was inconsequential as a predictor of the overall pooled effect, meaning that differences among the eight categories were statistically indistinguishable. Thus, our expectation of differences between categories of TDI, specifically regarding direct versus indirect approaches to targeting cognitive representations of the content, structure, and interrelationships of key elements of the task and environment and their communication between members, was unsupported. From a statistical standpoint, we may have been inadequately powered to detect a meaningful effect, particularly as three TDI subgroups were characterised by one study only. Thus, caution is urged when interpreting the results of these statistical findings.

Representing the second key contribution, our narrative synthesis and inspection of the subgroup effect sizes offered partial support for our expectation regarding the differential effectiveness of TDI categories on SMM. Effective interventions covered six of the ten TDI categories, namely team leadership (3 effective of 5), team training (2 effective of 6), team building (2 effective of 5), work design (2 effective of 16), team debriefing (1 effective or 1), and performance monitoring and assessment (1 effective of 3). Proportionally, although work design interventions were among the most prevalent of effective TDI, the ‘success rate’ from the total pool of studies was low (12.5%). Inspection of the subgroup effect sizes also indicated that the multicomponent interventions (7 effective of 15), except for team coaching, evidenced the strongest and most robust effects; point estimates were the largest ($g \geq .745$) and the confidence intervals had the highest lower bounds ($g \geq .265$). Within the context of the integrative framework on TDI (Shuffler et al., 2018), the preferences among SMM researchers reflects a focus on team inputs (work design) and multicomponent interventions (team training, team leadership, and team building) that target team inputs, processes, and/or outcomes.

Guided by socio-cognitive theories of learning (e.g., Akgün et al., 2003; Langfield-Smith, 1992), multicomponent interventions likely to enhance SMM via active strategies that directly elicit or shape the content and structure of knowledge held by individual members *and* how that knowledge is communicated between members, relative to interventions that address only one of these elements in isolation (e.g., how knowledge of taskwork and teamwork is communicated rather than the content and structure). Team training, for example, often encompasses activities that prompt interactions among members in which they engage actively with knowledge of key elements of the task and environment during simulated or real-world activities (Bisbey et al., 2019). Within the context of SMM, leaders play an active role in the enhancement efforts via promoting a shared understanding (similarity) within the team of what needs to be done (accuracy) to accomplish a task. As a key architect of the collective environment, leaders drive the formation of team norms, definition of collective objectives, and what needs to be done and the ways by which individual members are organised and integrated to achieve shared goals (Klein et al., 2006; Taggar & Ellis, 2007). In contrast, work design interventions, which specify and structure roles, tasks, and goals of the collective to align with team and organisational objectives (Knight & Parker, 2021), typically would be expected to target SMM indirectly unless they incorporate features that shape how knowledge is communicated between members. The conceptual and practical implications are that 'having' (i.e., knowledge of content and structure of mental models) and 'doing' (i.e., interactions between members) components are both important features of antecedent factors of SMM.

What Components of TDI are Most Effective for Enhancing Shared Mental Models?

The conceptual and practical resolution offered via knowledge of the TDI categories is limited to a broad, overarching view of the antecedents of SMM. Consistent with recent calls to examine the effects of TDI on cognitive, affective, and motivational outcomes (Rapp et al., 2021), our third contribution focused on unpacking the components utilised in these TDI with the view to clarify key drivers of change as a foundation upon which to appreciate how these interventions work. Moderator analyses indicated that the TDI components collectively were inconsequential as a predictor of the overall pooled effect, meaning that differences among the different elements were statistically indistinguishable. As with TDI type, it is likely that we were inadequately powered to detect a meaningful effect with 12 individual components, particularly as three subgroups were characterised by one study only and two subgroups by two studies. Nevertheless, our narrative synthesis and inspection of the subgroup effect sizes (see Table 2) offered some insight into potential nuances regarding the active ingredients of TDI for enhancing SMM. Inspection of the subgroup effect sizes indicated TDI components which actively engaged members with cognitive representations and/or their communication between members (strategy planning, sensegiving or sensemaking, interpositional knowledge or training) evidenced the strongest and most robust effects; point estimates were the largest ($g \geq .839$) and the confidence intervals had the highest lower bounds ($g \geq .226$). Knowledge representation and structure as well as shaping knowledge also offered robust moderate-to-large effects (point estimate $\geq .574$, lower bound CI $\geq .104$), despite being addressed directly or indirectly, depending on the category of TDI. As an example of indirect enactment, work design interventions augmented the ways by which knowledge is represented and structured via information and communication systems that optimised the environment in which collaborative discussions occur and the ways by which they unfold (e.g., Andres, 2012, 2013). In contrast, active approaches typically occurred within the context of team training interventions focused on optimising interactions among members (Marks et al., 2000) or team building interventions where members learned about each other’s roles and how best to coordinate their actions (Rittman, 2004). Our narrative interrogation of effective interventions reinforced the salience of these TDI components as drivers of enhanced SMM. Overall, therefore, our findings reinforce the importance of actively engaging members with essential knowledge and opportunities to communicate with each other for the enhancement of SMM, or augmenting the work environment to optimise collaborative interactions among members.

Conceptual Implications

There exists substantially more research on the consequences of team cognition and specific concepts like SMM rather than the antecedents of those collective cognitions (Mohammed et al., in press; Niler et al., 2021). Knowledge of the antecedents of SMM is fragmented across diverse scientific disciplines making it insufficient for informed theory development and practice. Via a synthesis of this dispersed literature, our meta-analytic and narrative findings offer two key contributions to theory on SMM. First, we confirmed the expectation that TDI can change or enhance SMM and clarified this position for future work by revealing moderately sized effects as a potential minimum value upon which to make judgements about the theoretical meaningfulness of empirical findings regarding the determinants of SMM. In so doing, our meta-analytic findings provide a benchmark upon which to evaluate and potentially falsify directional hypotheses in future work or serve as Bayesian priors for integration with new data to update knowledge immediately with each primary study. Second, our findings potentially shine a spotlight on TDI that directly rather than indirectly target cognitive representations of the content, structure, and interrelationships of key elements of the task and environment and their communication between members as the most likely candidates for enhancing SMM. This contribution is important because existing work on the antecedents of SMM is theoretically informed by a diverse range of theories. No single theoretical framework stood out among the effective interventions within this systematic review. Thus, regardless of the theoretical orientation adopted in future work, it is essential that researchers articulate exactly how guiding directly informs the ‘having’ and ‘doing’ elements of SMM of their intervention content. Relatedly, 14 of the 36 eligible studies made no specific mention of guiding theory; three of these studies were categorised as effective, five assessed as promising; and six as non-effective. Descriptively, our findings suggest that theory-informed research is more likely to lead to the development of effective or promising interventions than when theoretical underpinnings are absent.

Strengths, Limitations, and Future Research

Key strengths of this study include the utilisation of randomised trials or experiments as the evidence source, integration of a diverse range of TDI, preregistered protocol and transparency regarding methodological deviations, statistical analysis that accounted for dependency among effect sizes within primary studies, and open access dataset (for a review of the anatomy of high-quality meta-analyses, see Steel et al., in press). Nevertheless, it is important to acknowledge the limitations of this meta-analysis when interpreting the results and evaluating their contributions to theory and practice. First, we aimed to strike a balance between breadth and depth of information when coding and categorizing the types of interventions and their components employed to enhance SMM to maximise statistical power, yet acknowledge there are potentially meaningful differences between individual approaches within an intervention category that might have been diluted by these categorisations. Second, despite an extensive search for published and unpublished literature, we identified a relatively small number of primary studies (N = 36) to test our expectations regarding the effectiveness of TDI on the enhancement of SMM. Our focus on randomised controlled trials or experiments maximised the quality of evidence for making inferences regarding cause and effect, yet limited the quantity of data available for moderator tests that provide important information regarding boundary conditions (Gonzalez-Mulé & Aguinis, 2018). Relatedly, the types of moderator tests incorporated in this study – as with any meta-analysis – relied on the information available in primary studies, which in some cases was limited. For this reason, there remained substantial unexplained variance in the overall pooled effect, thereby suggesting other factors unaccounted for in this meta-analysis play a role in understanding the differential effectiveness of TDI on SMM. There exists a need for guidelines regarding the types of information that are of substantive (e.g., developmental cycle, active ingredients of interventions) and methodological (e.g., response rates) interest for meta-analysts interested in team dynamics so that primary researchers can report this information to facilitate future syntheses that incorporate their work. Third, consistent with observations made a decade ago (DeChurch & Mesmer-Magnus, 2010b), we observed substantial heterogeneity regarding the operationalisation of SMM in scientific research. This diversity is both a strength and weakness of the knowledge base; on the one hand, it has enabled researchers to design and test interesting questions regarding the antecedents of 'SMM', yet on the other hand it has limits regarding the specificity of findings related to SMM and the extent to which this cumulative evidence contributes to a coherent story regarding the most effective means by which to enhance SMM. Finally, as all but three experiments tested the effectiveness of TDI on SMM with university students, there is a need for researchers to examine the robustness of these findings in settings where SMM are often discussed as fundamental to high performing teams (e.g., military, emergency services, healthcare). Given the challenges of studying teams (Kolbe & Boos, 2019), scholars might find inspiration in contemporary guidelines (e.g., Klonek et al., 2019; Maynard et al., 2021) and new technologies (e.g., Kazi et al., 2021; Klonek et al., 2020) to facilitate such endeavours.

Conclusions

Several decades of research supports the scientific and organisational relevance of SMM for effective teamwork (Boies & Fiset, 2018; Cannon-Bowers & Salas, 2001; DeChurch & Mesmer-Magnus, 2010a; Mesmer-Magnus et al., 2017; Mohammed et al., 2010; Turner et al., 2014). Yet absent from the scientific literature is an integrative understanding of the ways by which SMM have been and can be enhanced. Our meta-analysis therefore fills important empirical and conceptual voids in the literature and, in so doing, provides an optimistic view of the potential of TDI as a means by which to enhance SMM among teams.

Ultimately, our findings provide the necessary building blocks from which to develop and refine a unifying theoretical framework of the nomological network of SMM, which is currently absent from the scientific literature. Nevertheless, there remains several challenges (e.g., diversity in the operationalisation of SMM) and opportunities (e.g., enhanced transparency in reporting) that require attention in future scholarly work if the concept of SMM is to fulfil its potential in science and practice.

References

Aguinis, H., Hill, N. S., & Bailey, J. R. (2019). Best practices in data collection and preparation: Recommendations for reviewers, editors, and authors. *Organizational Research Methods*, 1-16. doi: 10.1177/1094428119836485

Aguinis, H., Ramani, R. S., & Alabduljader, N. (2018). What you see is what you get? Enhancing methodological transparency in management research. *Academy of Management Annals, 12*, 83-110. doi: 10.5465/annals.2016.0011

Aguinis, H., Ramani, R. S., & Alabduljader, N. (2020). Best-practice recommendations for producers, evaluators, and users of methodological literature reviews. *Organizational Research Methods*, 1-31. doi: 10.1177/1094428120943281

Akgün, A. E., Lynn, G. S., & Byrne, J. (2003). Organizational learning: A socio-cognitive framework. *Human Relations, 56,* 839 – 868.

*Andres, H. P. (2012). Technology-mediated collaboration, shared mental model and task performance. *Journal of Organizational and End User Computing, 24,* 64-81. doi: 10.4018/joeuc.2012010104

*Andres, H. P. (2013). Team cognition using collaborative technology: A behavioral analysis. *Journal of Managerial Psychology, 28,* 38-54. doi: 10.1108/02683941311298850

Anvari, F., & Lakens, D. (2021). Using anchor-based methods to determine the smallest effect size of interest. *Journal of Experimental Social Psychology, 96,* 104159. https://doi.org/10.1016/j.jesp.2021.104159

Arend, M. G., & Schäfer, T. (2019). Statistical power in two-level models: A tutorial based on Monte Carlo simulation. *Psychological Methods, 24,* 1–19. doi: 10.1037/met0000195

Assink, M., & Wibbelink, C. J. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*, 154-174. doi: 10.20982/tqmp.12.3.p154

Bearman, C., Paletz, S. B., Orasanu, J., & Thomas, M. J. (2010). The breakdown of coordinated decision making in distributed systems. *Human Factors*, *52,* 173- 188. doi: 10.1177/0018720810372104

*Bedwell, W. L. (2019). Adaptive team performance: The influence of membership fluidity on shared team cognition. *Frontiers in Psychology*, *10*, 2266. doi: 10.3389/fpsyg.2019.02266

Bisbey, T. M., Reyes, D. L., Traylor, A. M., & Salas, E. (2019). Teams of psychologists helping teams: The evolution of the science of team training. *American Psychologist, 74,* 278–289. doi: 10.1037/amp0000419

*Boies, K., & Fiset, J. (2018). Leadership and communication as antecedents of shared mental models emergence. *Performance Improvement Quarterly*, *31*, 293-316. doi: 10.1002/piq.21267

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Hoboken, NJ: John Wiley.

*Burke, C. S. (2000). *Examination of the cognitive mechanisms through which team leaders promote effective team process and adaptive team performance* (Unpublished doctoral dissertation). George Mason University, Virginia.

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ*, *345*, e5661. doi: 10.1136/bmj.e5661

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. (1993). Shared mental models in expert team decision making. In J. N. Castellan Jr. (Ed.), *Individual and group decision making* (pp. 221-246). Hillsdale, NJ: Lawrence Erlbaum.

*Cassidy, S. A., & Stanley, D. J. (2019). Getting from ‘me’ to ‘we’: role clarity, team process, and the transition from individual knowledge to shared mental models in employee dyads. *Canadian Journal of Administrative Sciences, 36,* 208-220. doi: 10.1002/CJAS.1493

Chai, K.E.K., Lines, R.L.J., Gucciardi, D.F., & Ng, L. (2021). Research Screener: A machine learning tool to semi-automate abstract screening for systematic reviews. *Systematic Reviews, 10,* 93. doi: 10.1186/s13643-021-01635-3

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19,* 211–229. doi: 10.1037/a0032968

Cheung, M. W. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review, 29,* 387-396. doi: 10.1007/s11065-019-09415-6

Connell, L. E., Carey, R. N., De Bruin, M., Rothman, A. J., Johnston, M., Kelly, M. P., & Michie, S. (2019). Links between behaviour change techniques and mechanisms of action: An expert consensus study. *Annals of Behavioural Medicine, 53,* 708-720.

*Connors, E. S. (2006). *Intelligent group interfaces: Envisioned designs for exploring team cognition in emergency crisis management* (Unpublished doctoral dissertation). Pennsylvania State University, Pennsylvania.

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

*Crespin, T. R. (1996). *Cognitive convergence in developing groups: The role of sociocognitive elaboration* (Unpublished doctoral dissertation). The Ohio State University, Ohio.

*Curtis, A. M. (2009). *From monologue toward dialogue: Using performative objects to promote collective mindfulness in computer-mediated group discussions* (Unpublished doctoral dissertation). Indiana University, Indiana.

*Dalenberg, S., Vogelaar, A. L., & Beersma, B. (2009). The effect of a team strategy discussion on military team performance. *Military Psychology*, *21*, S31-S46. doi: 10.1080/08995600903249107

*DeChurch, L. A. (2002). *Teams leading teams: Examining the role of leadership in multi-team systems* (Unpublished doctoral dissertation). Florida International University, Florida.

DeChurch, L. A., & Marks, M. A. (2006). Leadership in multiteam systems. *Journal of Applied Psychology*, *91*, 311-329. doi: 10.1037/0021-9010.91.2.311

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

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010b). Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research, and Practice*, *14*, 1-14. doi: 10.1037/a0017455

*de Vreede, T. (2012). *The effect of regulatory focus, information sharing, and shared mental models on consensus* (Unpublished master’s thesis). University of Nebraska, Omaha.

Driskell, J. E., Salas, E., & Driskell, T. (2018). Foundations of teamwork and collaboration. *American Psychologist*, *73*(4), 334-348. doi: 10.1037/amp0000241

*Dubé, G., Kramer, C., Vachon, F., & Tremblay, S. (2011). Measuring the impact of a collaborative planning support system on crisis management. In *ISCRAM*.

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. doi: 10.1037/0021-9010.91.3.727

Edwards, J.R., & Christian, M.S. (2014). Using accumulated knowledge to calibrate theoretical propositions. *Organizational Psychology Review, 4*, 279-291. doi: 10.1177/2041386614535131

*Ellwart, T., Happ, C., Gurtner, A., & Rack, O. (2015). Managing information overload in virtual teams: Effects of a structured online team adaptation on cognition and performance. *European Journal of Work and Organizational Psychology*, *24*, 812-826. doi: 10.1080/1359432X.2014.1000873

*Engelmann, T., Kozlov, M. D., Kolodziej, R., & Clariana, R. B. (2014). Fostering group norm development and orientation while creating awareness contents for improving net-based collaborative problem solving. *Computers in Human Behavior*, *37,* 298-306. doi: 10.1016/j.chb.2014.04.052

Fernández-Castilla, B., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P., & Noortgate, W. V. den. (2021). Detecting Selection Bias in Meta-Analyses with Multiple Outcomes: A Simulation Study. *The Journal of Experimental Education*, *89*(1), 125–144. https://doi.org/10.1080/00220973.2019.1582470

Gonzalez-Mulé, E., & Aguinis, H. (2018). Advancing theory by assessing boundary conditions with meta-regression: A critical review and best-practice recommendations. *Journal of Management*, 44, 2246–2273. doi: 10.1177/0149206317710723

*Grand, J. A., Braun, M. T., Kuljanin, G., Kozlowski, S. W., & Chao, G. T. (2016). The dynamics of team cognition: A process-oriented theory of knowledge emergence in teams. *Journal of Applied Psychology*, *101*, 1353-1385. doi: 10.1037/apl0000136

Gucciardi, D., Lines, R., & Crane, M. (2021). *Optimising shared mental models: A systematic review and meta-analysis of randomised controlled trials*. Retrieved from https://osf.io/yzdxn/.

*Gurtner, A., Tschan, F., Semmer, N. K., & Nägele, C. (2007). Getting groups to develop good strategies: Effects of reflexivity interventions on team process, team performance, and shared mental models. *Organizational Behavior and Human Decision Processes*, *102*, 127-142. doi: 10.1016/j.obhdp.2006.05.002

Guyatt, G.H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., et al. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, *336*: 924. doi: 10.1136/bmj.39489.470347.AD

*Hagemann, V., Kluge, A., & Badura, B. (2012). The impact of a large-screen projection of the technical process on shared mental models and team performance in a furnace control room. In D. de Waard, N. Merat, A.H. Jamson, Y. Barnard, and O.M.J. Carsten (Eds.) (2012). *Human Factors of Systems and Technology* (pp. 1 - 15). Maastricht: Shaker Publishing.

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R.,... Zwienenberg, M. (2016). A multi-lab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science, 11,* 546 –573. doi: 10.1177/174 5691616652873

Hagger, M.S., Gucciardi, D.F., & Chatzisarantis, N.L.D. (2017). On nomological validity and auxiliary assumptions: The importance of simultaneously testing effects in social cognitive theories applied to health behaviour and some guidelines. *Frontiers in Psychology, 8*: 1933: doi: 10.3389/fpsyg.2017.01933/full

Higgins, J.P., Thompson, S.G., Deeks, J.J., & Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327,* 557–560. doi: 10.1136/bmj.327.7414.557

Hughes, A. M., Gregory, M. E., Joseph, D. L., Sonesh, S. C., Marlow, S. L., Lacerenza, C. N., . . . Salas, E. (2016). Saving lives: A meta-analysis of team training in healthcare. *Journal of Applied Psychology, 101,* 1266–1304. doi: 10.1037/apl0000120

IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, *6*(7), e010247. doi: 10.1136/bmjopen-2015-010247

Kazi, S., Khaleghzadegan, S., Dinh, J. V., Shelhamer, M. J., Sapirstein, A., Goeddel, L. A., Chime, N. O., Salas, E., & Rosen, M. A. (2021). Team physiological dynamics: A critical review. *Human Factors*, *63*(1), 32–65. https://doi.org/10.1177/0018720819874160

*Kim, H., & Kim, D. (2008). The effects of the coordination support on shared mental models and coordinated action. *British Journal of Educational Technology*, *39*, 522-537. doi: 10.1111/j.1467-8535.2007.00751.x

Knight, C., & Parker, S.K. (2021). How work redesign interventions affect performance: An evidence-based model from a systematic review. *Human Relations, 74,* 69-104. doi: 10.1177/0018726719865604

Klonek, F. E., Gerpott, F. H., Lehmann-Willenbrock, N., & Parker, S. K. (2019). Time to go wild: How to conceptualize and measure process dynamics in real teams with high-resolution. *Organizational Psychology Review*, *9*(4), 245–275. https://doi.org/10.1177/2041386619886674

Klonek, F. E., Meinecke, A. L., Hay, G., & Parker, S. K. (2020). Capturing Team Dynamics in the Wild: The Communication Analysis Tool. *Small Group Research*, *51*(3), 303–341. https://doi.org/10.1177/1046496420904126

Kolbe, M., & Boos, M. (2019). Laborious but elaborate: The benefits of really studying team dynamics. *Frontiers in Psychology*, *10*, 1478. https://doi.org/10.3389/fpsyg.2019.01478

Kossmeier, M., Tran, U. S., Voracek, M. (2020a). Conducting visual inference with funnel plots using *R* package metaviz. Retrieved from: https://cran.r-project.org/web/packages/metaviz/vignettes/funnelinf.html

Kossmeier, M., Tran, U. S., & Voracek, M. (2020b). Power-enhanced funnel plots for meta-analysis: The sunset funnel plot. *Zeitschrift für Psychologie*, *228*, 43–49. doi: 10.1027/2151-2604/a000392

Kozlowski, S. W. J., Gully, S. M., Salas, E., & Cannon-Bowers, J. A. (1996). Team leadership and development: Theory, principles, and guidelines for training leaders and teams. In M. M. Beyerlein, D. A. Johnson, & S. T. Beyerlein (Eds.), *Advances in interdisciplinary studies of work teams* (Vol. 3, pp. 253–291). Greenwich, CT: JAI Press.

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

Klein, K. J., Ziegert, J. C., Knight, A. P., & Xiao, Y. (2006). Dynamic delegation: Shared, hierarchical, and deindividualised leadership in extreme action teams. *Administrative Science Quarterly, 51,* 590–621. doi: 10.2189/asqu.51.4.590

Lacerenza, C. N., Marlow, S. L., Tannenbaum, S. I., & Salas, E. (2018). Team development interventions: Evidence-based approaches for improving teamwork. *American Psychologist, 73,* 517-531. doi: 10.1037/amp0000295

Lacerenza, C. N., Reyes, D. L., Marlow, S. L., Joseph, D. L., & Salas, E. (2017). Leadership training design, delivery, and implementation: A meta-analysis. *Journal of Applied Psychology, 102,* 1686–1718. doi: 10.1037/apl0000241

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*: 418 doi: 10.3389/fpsyg.2013.00863

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology, 44*, 701–710. https://doi.org/10.1002/ejsp.2023

Lakens, D., & Evers, E. R. K. (2014). Sailing From the Seas of Chaos Into the Corridor of Stability: Practical Recommendations to Increase the Informational Value of Studies. *Perspectives on Psychological Science*, *9*, 278–292. doi: 10.1177/1745691614528520

Langfield-Smith, K. (1992). Exploring the need for a shared cognitive map. *Journal of Management Studies, 29*, 349-368. doi:10.1111/j.1467-6486.1992.tb00669.x

Latane, B. (1981). The psychology of social impact. *The American Psychologist*, *36,* 343–356. doi: 10.1037/0003-066X.36.4.343

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

Lines, R. L. J., Pietsch, S., Crane, M., Ntoumanis, N., Temby, P., Graham, S., & Gucciardi, D.F. (in press). The effectiveness of team reflexivity interventions: A systematic review and meta-analysis of randomised controlled trials. *Sport, Exercise, and Performance Psychology*. doi: 10.1037/spy0000251

*Liu, J., Qin, Y., Yang, Q., Yu, C., & Shi, Y. (2015). A Tabletop-Centric Smart Space for Emergency Response. *IEEE Pervasive Computing*, *14*, 32-40. doi: 10.1109/MPRV.2015.24

López-López, J. A., Page, M. J., Lipsey, M. W., & Higgins, J. P. T. (2018). Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods*, *9*(3), 336–351. doi: 10.1002/jrsm.1310

Lovakov, A., & Agadullina, E.R. (in press). Empirically derived guidelines for effect size interpretation in social psychology. *European Journal of Social Psychology*. doi: 10.1002/ejsp.2752

*Lowry, P. B., Roberts, T. L., & Romano Jr, N. C. (2013). What signal is your inspection team sending to each other? Using a shared collaborative interface to improve shared cognition and implicit coordination in error-detection teams. *International Journal of Human-Computer Studies*, *71*, 455-474. doi: 10.1016/j.ijhcs.2012.11.004

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

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

Maynard, M. T., Conroy, S., Lacerenza, C. N., & Barnes, L. Y. (2021). Teams in the wild are not extinct, but challenging to research: A guide for conducting impactful team field research with 10 recommendations and 10 best practices. *Organizational Psychology Review*, 2041386620986597. https://doi.org/10.1177/2041386620986597

McEwan, D., Ruissen, G. R., Eys, M. A., Zumbo, B. D., & Beauchamp, M. R. (2017). The effectiveness of teamwork training on teamwork behaviors and team performance: A systematic review and meta-analysis of controlled interventions. *PLoS ONE,* 12: e0169604. doi: 10.1371/journal. pone.0169604

Mesmer-Magnus, J., Niler, A. A., Plummer, G., Larson, L. E., & DeChurch, L. A. (2017). The cognitive underpinnings of effective teamwork: A continuation. *Career Development International, 22*, 507-519. doi: 10.1108/CDI-08-2017-0140

*Millward, S. M. (2006). Do you know your STUFF? Training collaborative modelers. *Team Performance Management, 12,* 225-236. doi: 10.1108/13527590610711787

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

Mohammed, S., Hamilton, K., Tesler, R., Mancuso, V., & McNeese, M. (2015). Time for temporal team mental models: Expanding beyond "what" and "how" to incorporate "when." *European Journal of Work and Organizational Psychology, 24,* 693-709. doi: 10.1080/1359432x.2015.1024664

Mohammed, S., Rico, R., & Alipour, K. (in press). Team cognition at a crossroad: Toward conceptual integration and network configurations. *Academy of Management Annals*. doi: 10.5465/annals.2018.0159

Montgomery, P., Grant, S., Mayo-Wilson, E., Macdonald, G., Michie, S., Hopewell, S., & Moher, D. (2018). Reporting randomised trials of social and psychological interventions: the CONSORT-SPI 2018 Extension. *Trials*, *19*, 407. doi: 10.1186/s13063-018-2733-1

*Newlon, C. M. (2016). *The effect of shared dynamic understanding on willingness to contribute information: design and analysis of a mega-collaborative interface* (Unpublished doctoral dissertation). Indiana University, Indiana.

Niler, A. A., Mesmer-Magnus, J. R., Larson, L. E., Plummer, G., DeChurch, L. A., & Contractor, N. S. (2021). Conditioning team cognition: A meta-analysis. *Organizational Psychology Review, 11*, 144-174. doi: 10.1177/2041386620972112.

*Ouverson, K. M. (2019). *Intelligent team tutoring: An analysis of communication, cognition, cooperation, and coordination* (Unpublished master's thesis). Iowa State University, Iowa.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372;* n71. doi: 10.1136/bmj.n71

* Putri, D. A. P. (2012). *The effect of communication strategy and planning intervention on the processes and performance of course material development teams* (Unpublished doctoral dissertation). Florida State University, Florida.

Podsakoff, P. M., & Podsakoff, N. P. (2019). Experimental designs in management and leadership research: Strengths, limitations, and recommendations for improving publishability. *The Leadership Quarterly*, *30*, 11–33. doi: 10.1016/j.leaqua.2018.11.002

Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin*, *112*, 160–164. https://doi.org/10.1037/0033-2909.112.1.160.

Rafferty, L. A., Stanton, N. A., & Walker, G. H. (2010). The famous five factors in teamwork: A case study of fratricide. *Ergonomics, 53*, 1187-1204. doi: 10.1080/ 00140139.2010.513450

*Randall, K. R. (2008). *Adaptation in knowledge-based teams: An examination of team composition, leader sensegiving, and cognitive, behavioral, and motivational mechanisms* (Unpublished doctoral dissertation). Florida International University, Florida.

Rapp, T., Maynard, T., Domingo, M., & Klock, E. (2021). Team emergent states: What has emerged in the literature over 20 tears. *Small Group Research*, *52,* 68–102. doi: 10.1177/1046496420956715

*Rentsch, J. R., Delise, L. A., Mello, A. L., & Staniewicz, M. J. (2014). The integrative team knowledge building training strategy in distributed problem-solving teams. *Small Group Research*, *45*, 568-591. doi: 10.1177/1046496414537690

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* (pp. 129-157). Mahwah, NJ: Lawrence Erlbaum Associates.

Resick, C. J., Dickson, M. W., Mitchelson, J. K., Allison, L. K., & Clark, M. A. (2010). Team composition, cognition, and effectiveness: Examining mental model similarity and accuracy. *Group Dynamics: Theory, Research, and Practice*, *14*, 174-191. doi: 10.1037/a0018444

*Rittman, A. L. (2004). *A comparison of team training strategies for team effectiveness* (Unpublished doctoral dissertation). George Mason University, Virginia.

Salancik, G.R., & Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. *Administrative Science Quarterly, 23,* 224–253.

Salas, E., DiazGranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008). Does team training improve team performance? A meta-analysis. *Human Factors*, *50*, 903-933. doi: 10.1518/001872008X375009.

Salas, E., Nichols, D., & Driskell, J. E. (2007). Testing three team training strategies in intact teams: A meta-analysis. *Small Group Research, 38*, 471–488. doi: 10.1177/1046496407304332

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3–29). Norwood, NJ: Ablex Publishing Corporation.

Schroder, H., Driver, M., & Streufert, S. (1967). *Human information processing*. New York, NY: Holt, Rinehart & Winston.

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ, 349,* g7647. doi: 10.1136/bmj.g7647

Shuffler, M. L., Diazgranados, D., Maynard, M. T., & Salas, E. (2018). Developing, sustaining, and maximizing team effectiveness: An integrative, dynamic perspective of team development interventions. *Academy of Management Annals*, *12*, 688-724. doi: 10.5465/annals.2016.0045

*Sikorski, E. G. (2009). *Team knowledge sharing intervention effects on team shared mental models and team performance in an undergraduate meteorology course* (Unpublished doctoral dissertation). Florida State University, Florida.

Stanley, T. D., Carter, E. C., & Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin, 144*, 1325–1346. doi: 10.1037/bul0000169

Steel, P., Beugelsdijk S., & Aguinis, H. (in press). The anatomy of an award-winning meta-analysis: Recommendations for authors, reviewers, and users of meta-analytic reviews. *Journal of International Business Studies.*

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Emberson, J. R. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ, 336*, l4898. doi: 10.1136/bmj.l4898

Sundstrom, R., de Meuse, K. P., & Futrell, D. (1990). Work teams: Applications and effectiveness. *American Psychologist, 45*, 120–133. doi: 10.1037/0003-066X.45.2.120

Taggar, S., & Ellis, R. (2007). The role of leaders in shaping formal team norms. *The Leadership Quarterly, 18*, 105–120. doi: 10.1016/j.leaqua.2007.01.002

Toader, A. F., Cantner, U., & Kessler, T. (2019). The Effect of Team Mental Models Divergence on Creative Performance During Situational Changes. *Creativity Research Journal*, *31*, 40-51. doi: 10.1080/10400419.2019.1577206

*Toader, A. F., & Kessler, T. (2018). Task Variation and Mental Models Divergence Influencing the Transfer of Team Learning. *Small Group Research, 49*, 545-575. doi: 10.1177/1046496418786429

Turner, J. R., Chen, Q., & Danks, S. (2014). Team shared cognitive constructs: A meta-analysis exploring the effects of shared cognitive constructs on team performance. *Performance Improvement Quarterly*, *27*, 83-117. doi: 10.1002/piq.21163

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45*, 576-594. doi: 10.3758/s13428-012-0261-6

*van Ginkel, W. P., & van Knippenberg, D. (2008). Group information elaboration and group decision making: The role of shared task representations. *Organizational Behavior and Human Decision Processes*, *105*, 82-97. doi: 10.1016/j.obhdp.2007.08.005

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1-48. doi: 10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*, 112-125. doi: 10.1002/jrsm.11

*Yue, C., Fong, P. S., & Li, T. (2019). Meeting the challenge of workplace change: Team cooperation outperforms team competition. *Social Behavior and Personality, 47*, e7997. doi: 10.2224/sbp.7997

Table 1

Results for Overall Mean Effect Sizes.

Outcome	**#Studies**	**#ES**	**Mean *g* (SE)**	**95% CI**	**t-statistic**	***p***	**Level 1 variance (%)**	**Level 2 variance (%)**	**Level 3 variance (%)**
Overall SMM	36	131	.615 (.104)	0.409, 0.821	5.904	< .0001	23.809	43.012***	33.178***
SMM Similarity	36	99	.558 (.096)	0.367, 0.749	5.795	< .0001	28.376	46.433***	25.191***
SMM Accuracy	9	32	.754 (.343)	0.055, 1.452	2.199	.035	13.346	49.472***	37.182*
Teamwork SMM	15	59	.705 (.157)	0.391, 1.019	4.496	< .0001	31.209	53.111	15.679***
Taskwork SMM	20	53	.284 (.103)	0.077, 0.492	2.746	.008	21.482	1.238***	78.518
Combined SMM	9	19	.708 (.218)	0.249, 1.116	3.244	.005	23.496	63.148	13.355

Note: #Studies = number of studies; #ES = number of effect sizes; Mean *g* = mean effect size; SE = standard error; CI = confidence interval; *p* = significance of mean effect size; Level 2 variance = percentage variance between effect sizes from the same study; Level 3 Variance = percentage variance in effect sizes between studies; *** = p < .001, ** = p < .01, * = p < .05.

Table 2

Moderator and Sensitivity Analyses of the Overall Effect of Shared Mental Model Interventions.

Moderator	#Studies	#ES	*g* (95%CI)	Overall[a]	*p*[b]	Level 2 variance	Level 3 variance
Study Characteristics							
Sample Size	36	131	.614 (0.405, 0.823)***	F(1, 129) = 0.048	.827	32.887***	43.732***
Number of Teams	36	131	.612 (0.403, 0.821)***	F(1, 129) = 0.128	.721	33.181***	43.297***
Team Size	36	131	.629 (0.417, 0.841)***	F(1, 129) = 0.481	.489	32.857***	43.761***
Age	26	85	.663 (0.380, 0.945)***	F(1, 83) = 0.027	.871	45.016***	38.036**
Percentage Female	30	120	.551 (0.339, 0.763)***	F(1, 118) = 0.481	.829	32.860***	41.383***
Risk of Bias				F(1, 129) = 0.419	.518	32.628***	44.076***
Low (RC)	14	52	.706 (0.361, 1.051)***				
Some Concerns	22	79	.564 (0.301, 0.827)***				
Publication Type				F(1, 129) = 0.040	.843	31.952***	44.995***
Peer-Reviewed (RC)	24	98	.633 (0.372, 0.893)***				
Thesis	12	33	.588 (0.227, 0.949)**				
Outcome Characteristics							
Outcome Measurement				F(2, 128) = 6.464	.002	34.222***	40.366***
Objective (RC)	3	3	.297 (-0.462, 1.057)				
Research Assessed	2	2	3.384 (1.813, 4.955)***				
Self-reported	31	126	.585 (0.379, 0.791)***				
Form of Cognition				F(1, 129) = 2.427	.122	31.363***	45.373***
Perceptual (RC)	22	64	.714 (0.470, 0.958)***				
Structured	19	67	.516 (0.270, 0.763)***				
Task or Teamwork Focus				F(2, 128) = 1.897	.154	35.137***	39.844***
Combined (RC)	9	19	.650 (0.265, 1.036)**				
Taskwork	20	53	.482 (0.230, 0.734)***				

Moderator	#Studies	#ES	*g* (95%CI)	Overall[a]	*p*[b]	Level 2 variance	Level 3 variance
Teamwork	15	59	.740 (0.481, 0.998)***				
Similarity or Accuracy				F(1, 129) = 2.899	.091	29.824***	46.938***
Accuracy (RC)	9	32	.849 (0.507, 1.191)***				
Similarity	36	99	.576 (0.359, 0.794)***				
Intervention Characteristics							
TDI Category				F(7, 123) = 0.789	.597	31.804**	45.063***
Team building (RC)	5	14	.975 (.417, 1.533)***				
Team coaching	1	1	.135 (-1.790, 2.060)				
Team composition	1	24	.213 (-.791, 1.218)				
Team debriefing	1	2	1.227 (-.029, 2.482)				
Team leadership	5	18	.762 (.269, 1.256)**				
Team performance monitoring and assessment	3	6	.379 (-.344, 1.102)				
Team training	6	35	.745 (.265, 1.225)**				
Work design	16	31	.425 (.076, .773)*				
TDI Component				F(11,119) = 1.173	0.313	32.586***	43.549***
Alterations to human capital resources (RC)	1	24	.213 (-.761, 1.187)				
Critical thinking	2	3	.775 (-.200, 1.751)				
Inspirational motivation	1	2	.815 (-.330, 1.961)				
Interpositional knowledge or training	3	28	.839 (.226, 1.451)**				
Knowledge exchange	5	10	.255 (-.269, .778)				
Knowledge representation and structure	10	20	.574 (.111, 1.038)*				
Performance feedback and guidance	3	3	.165 (-.686, 1.016)				
Reward structures	1	2	-.400 (-1.568, .768)				
Role clarity	2	6	.923 (.191, 1.654)*				
Sensegiving or sensemaking	4	12	1.095 (.527, 1.663)***				

Shaping knowledge	5	18	.596 (.104, 1.088)*				
Strategy planning	2	3	1.313 (.390, 2.236)**				
Use of Technology				F(1, 129) = 0.020	.888	32.585***	44.088***
Moderator	#Studies	#ES	*g* (95%CI)	Overall[a]	*p*[b]	Level 2 variance	Level 3 variance
Present (RC)	13	25	.592 (0.199, 0.986)**				
Absent	23	106	.626 (0.379, 0.872)***				
Comparator				F(2, 128) = 2.780	.066	32.579***	42.859***
Active (RC)	7	18	.265 (-0.180, 0.710)				
No Treatment	24	105	.775 (0.531, 1.019)***				
Treatment as Usual	5	8	.227 (-0.408, 0.861)				

Note: #Studies = number of studies, #ES = number of effect sizes, Mean *g* = mean effect size, CI = confidence interval, Level 2 variance = variance in effect sizes from the same study, Level 3 variance = variance in effect sizes between studies, RC = reference category, *** = p < .001, ** = p < .01, * = p < .05, [a] Omnibus test of all regression coefficients in the model, [b] p-value of omnibus test.

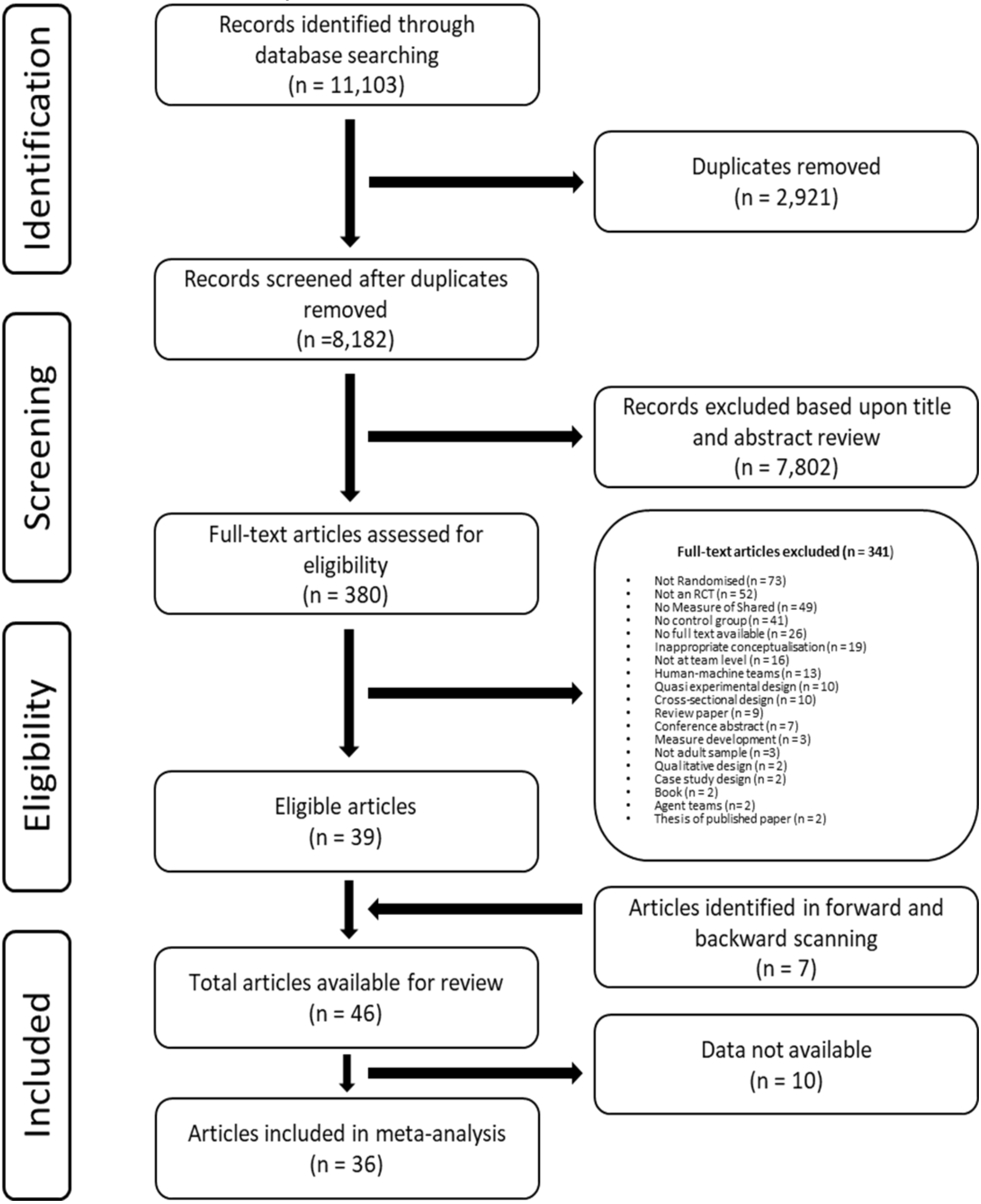


Figure 1. PRISMA flow diagram.

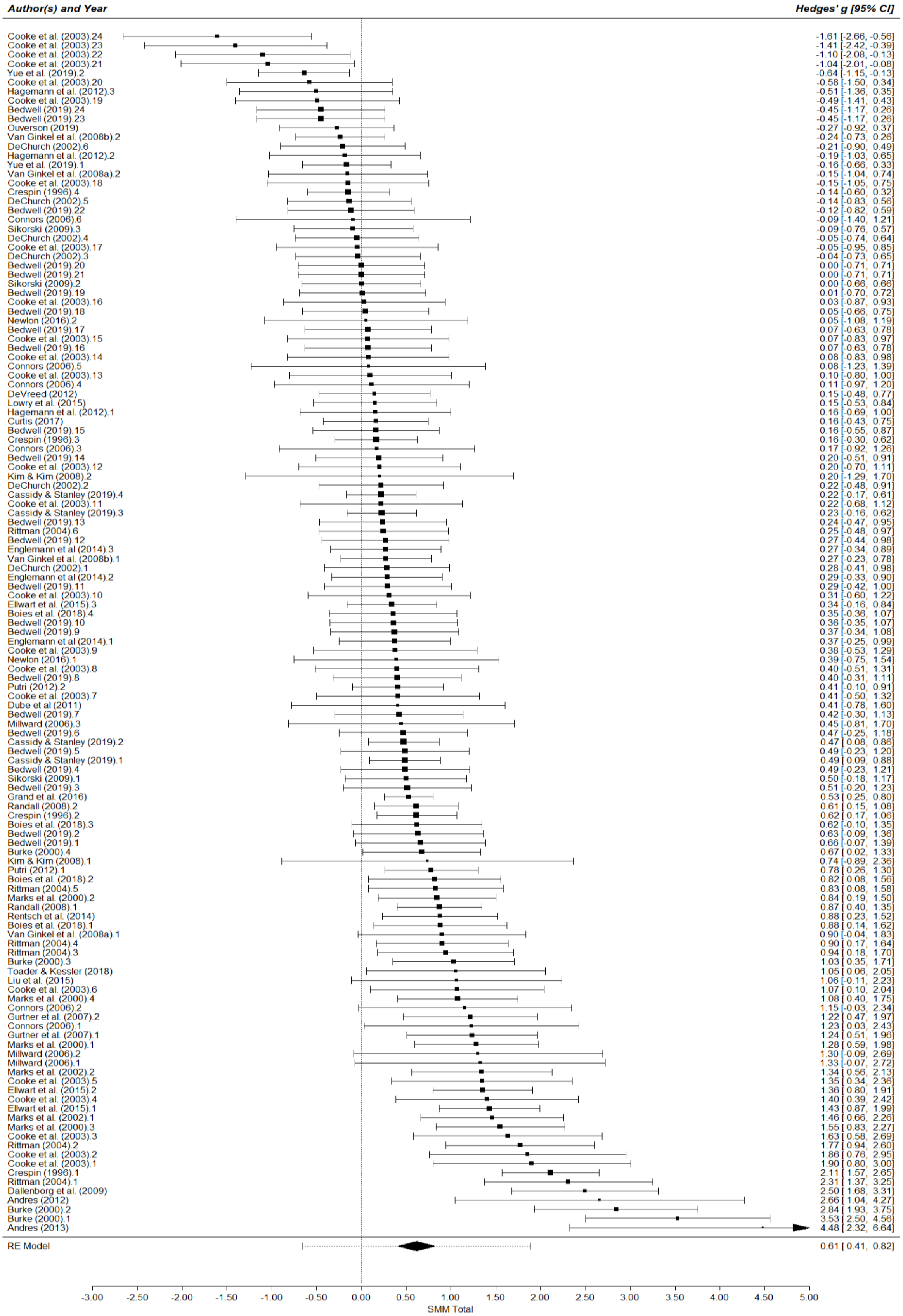
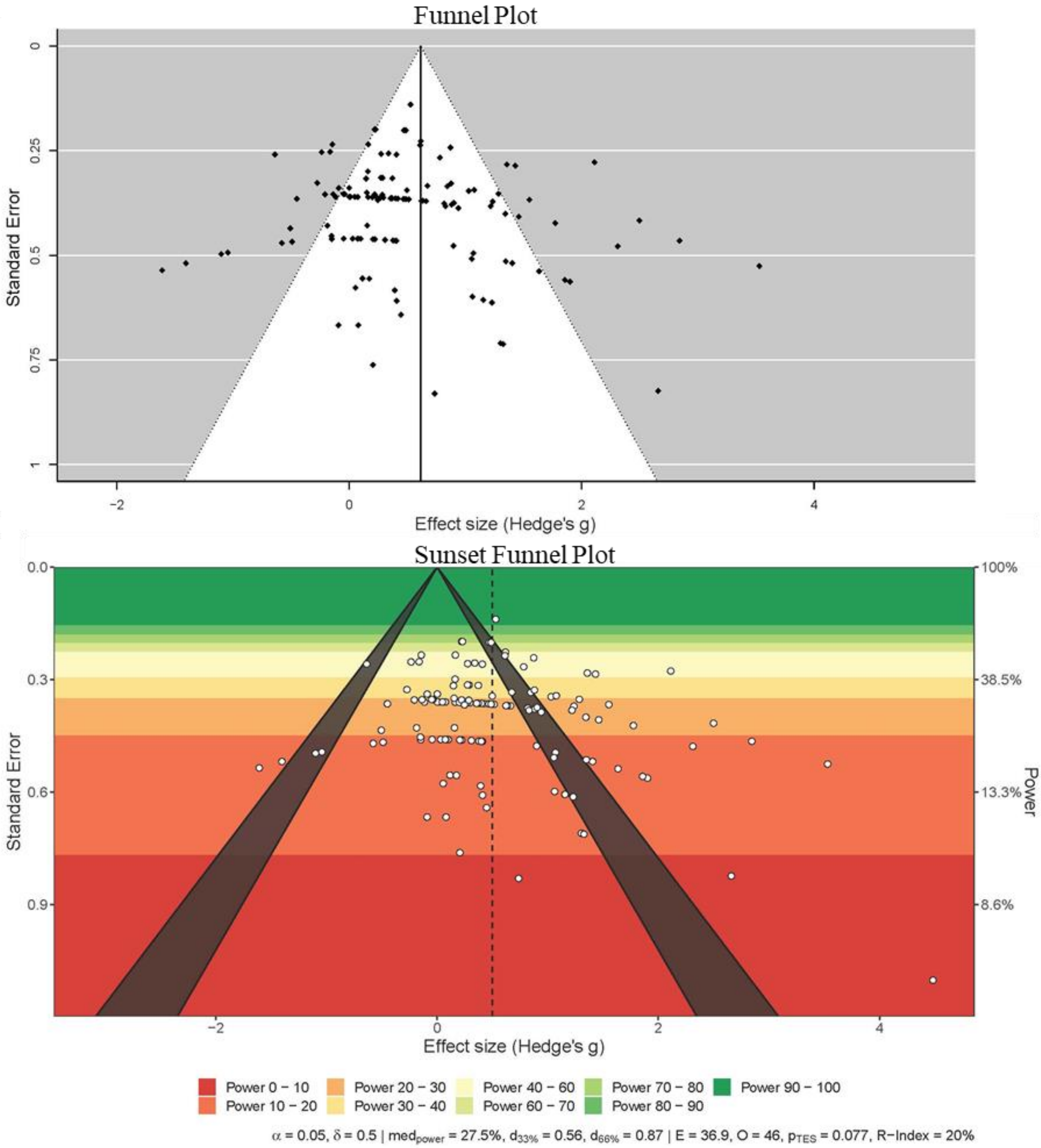


Figure 2. Forest plot of the overall effect of interventions on shared mental models (for a high resolution version, see here: http://bit.ly/smm-figure2).



Note: We assumed a true effect of *g* = .50 for the sunset funnel plot.

Figure 3. Funnel and sunset funnel plots for the overall effect of interventions on shared mental models.

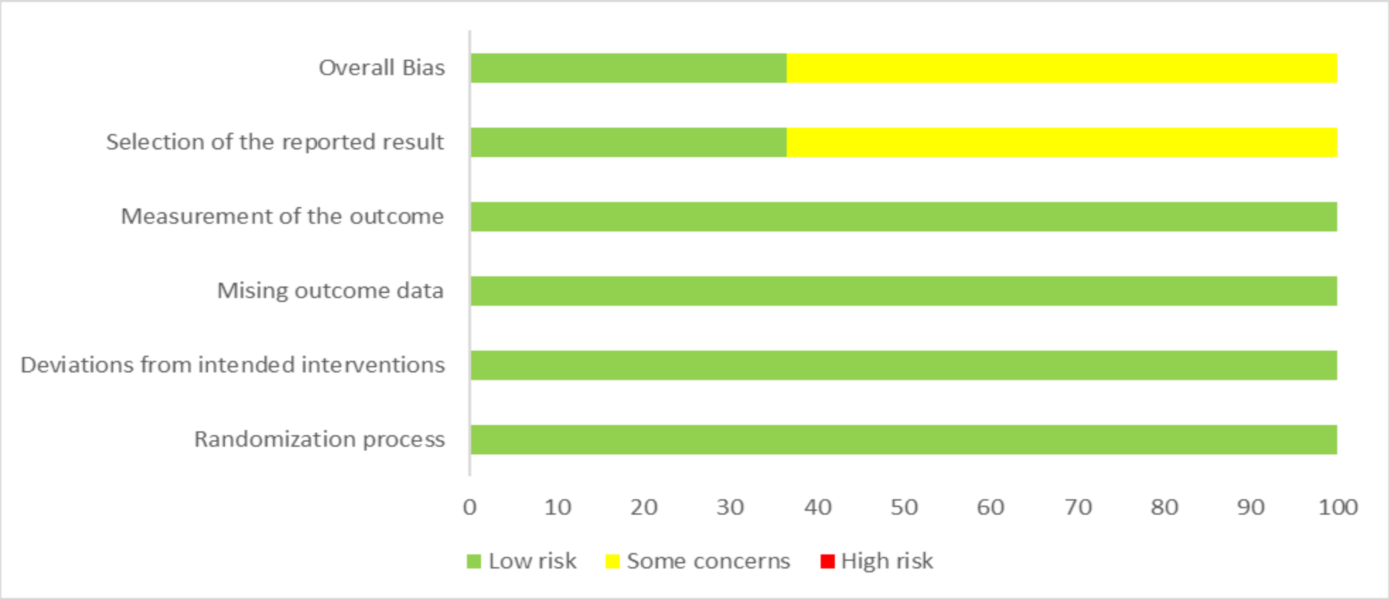


Figure 4. Risk of bias summary.