Does Team Training Improve Team Performance? A Meta-Analysis

Eduardo Salas, Deborah DiazGranados, Cameron Klein, C. Shawn Burke, Kevin C. Stagl, Gerald F. Goodwin and Stanley M. Halpin

Human Factors: The Journal of the Human Factors and Ergonomics Society 2008 50: 903

DOI: 10.1518/001872008X375009

Does Team Training Improve Team Performance? A Meta-Analysis

Eduardo Salas, Deborah DiazGranados, Cameron Klein, C. Shawn Burke, and **Kevin C. Stagl,** University of Central Florida, Orlando, Florida, and **Gerald F. Goodwin** and **Stanley M. Halpin,** Army Research Institute, Arlington, Virginia

Objective: This research effort leveraged the science of training to guide a taxonomic integration and a series of meta-analyses to gauge the effectiveness and boundary conditions of team training interventions for enhancing team outcomes. **Background:** Disparate effect sizes across primary studies have made it difficult to determine the true strength of the relationships between team training techniques and team outcomes. **Method:** Several meta-analytic integrations were conducted to examine the relationships between team training interventions and team functioning. Specifically, we assessed the relative effectiveness of these interventions on team cognitive, affective, process, and performance outcomes. Training content, team membership stability, and team size were investigated as potential moderators of the relationship between team training and outcomes. In total, the database consisted of 93 effect sizes representing 2,650 teams. **Results:** The results suggested that moderate, positive relationships exist between team training interventions and each of the outcome types. The findings of moderator analyses indicated that training content, team membership stability, and team size moderate the effectiveness of these interventions. **Conclusion:** Our findings suggest that team training interventions are a viable approach organizations can take in order to enhance team outcomes. They are useful for improving cognitive outcomes, affective outcomes, teamwork processes, and performance outcomes. Moreover, results suggest that training content, team membership stability, and team size moderate the effectiveness of team training interventions. **Application:** Applications of the results from this research are numerous. Those who design and administer training can benefit from these findings in order to improve the effectiveness of their team training interventions.

Attention to team training has grown exponentially during the past 10–20 years. It is driven by at least the perception that the "high performance team" is a critical element in the design of organizations that will be effective in the global economy.

– Campbell & Kuncel (2001, p. 299)

INTRODUCTION

The nature of work has changed. Organizations now face increased competition and collaboration within and across organizational, geographic, and temporal boundaries; a need to engage a demographically heterogeneous workforce; a need to deal with advancements in information technology; a need to promote safety; and a need to foster enduring customer relations (Noe, 2002; Pfeffer & Sutton, 2000; Tannenbaum, 2002). One response to these changes has been the use of work teams as a preferred performance management technique. "Teams are ubiquitous. Whether we are talking about software development, Olympic hockey, disease outbreak response, or urban warfare, teams represent the critical unit that 'gets things done' in today's world" (Marks, 2006, p. i).

Both governmental agencies and private industry are increasingly relying upon work teams as a preferred performance arrangement to fulfill their visions, execute their complex missions, and accomplish their goals (Salas, Stagl, & Burke, 2004). In fact, a recent random sample of U.S. organizations indicated that nearly half (48%) used some

Address correspondence to Eduardo Salas, Institute for Simulation & Training, University of Central Florida, 3100 Technology Pkwy., Orlando, FL 32826; esalas@ist.ucf.edu. ***HUMAN FACTORS,*** Vol. 50, No. 6, December 2008, pp. 903–933. DOI 10.1518/001872008X375009.

type of team-based structure, with ongoing project teams being the most frequently reported (Devine, Clayton, Philips, Dunford, & Melner, 1999).

The increasing frequency of team-based forms of organizing presents unique opportunities and creates challenges. Opportunities via the use of teams abound. For example, collectives can leverage shared mental models, compensatory processes, and affective states such as cohesion to deal more effectively with the complex, stressful, and sometimes chaotic contexts that typify modern operations (Orasanu & Salas, 1993). The process gains yielded by teams can ultimately culminate in enhanced efficiencies, quality and safety improvements, creativity, and/or adaptation (Banker, Field, Schroeder, & Sinha, 1996; Burke, Stagl, Salas, Pierce, & Kendall, 2006; Cohen & Ledford, 1994; Foushee, 1984).

However, challenges inherent to the use of teams exist. Most notably, how should organizations compose, manage, and develop team members? More specifically, how does one design and deliver team training? What kind of team training should be implemented? And finally, does it work? This research was motivated by the need to answer these questions.

Team training targets latent teamwork knowledge, skills, and/or attitudinal competencies (KSAs) as well as manifest team processes and performance for improvement (Salas & Cannon-Bowers, 1997, 2000). Teams can be trained to make better decisions (Orasanu & Fischer, 1997), to perform better under stress (Driskell, Salas, & Johnston, 1995), and to make fewer errors (Wiener, Kanki, & Helmreich, 1993).

Enhancing teamwork through team training is a chief concern in many industries, including commercial and military aviation (Boehm-Davis, Holt, & Seamster, 2001; Helmreich & Foushee, 1993; Oser, Salas, Merket, & Bowers, 2001), health care (Baker, Salas, Barach, Battles, & King, 2007; Davies, 2001; Salas, DiazGranados, Weaver, & King, 2008), and energy (Flin, 1995; Flin & O'Connor, 2001), to name just a few. In these industries (and others), a great deal of research has been generated on team training. Unfortunately, no comprehensive source exists in the team training literature that demonstrates whether or not team training even matters and, if it does, by how much. Further, no source outlines the conditions under which team training works. This research initiative seeks to fill these voids.

The following sections of this paper are organized as follows. We first describe the rationale for the current meta-analytic initiative. Next, we address the conceptual foundation on which this research is based. Included here is a short discussion centered upon the nature of teams and team performance, as well as a longer discourse on the current state of team training efforts. Eight specific study hypotheses are then put forward, followed by a detailed accounting of the study methods. An ensuing description of meta-analytic results is subsequently presented, followed by implications for applied work with teams – a discussion that fully incorporates study limitations and directions for future research on team training.

WHY A TEAM TRAINING INTEGRATION?

This research effort includes a series of meta-analyses, each of which was conducted to gauge the effectiveness of team training for enhancing team outcomes. Specifically, we sought to (a) clarify and contribute to the science of training by providing a clarification of the nature of team training via a taxonomic integration of team training interventions; (b) establish the boundary conditions of team training for enhancing team outcomes via moderator analyses with additional variables of interest; and (c) impart valuable information for organizational stakeholders charged with designing, delivering, and evaluating team training interventions.

First, several limitations of the existing base of research on team training and development have served to impede knowledge accumulation and integration (Campbell & Kuncel, 2001). As Judge, Cable, Colbert, and Rynes (2007) recently noted, "Regardless of the quality of an idea, the ability to draw inferences about a phenomenon is constrained by the quality of the methods used to gather data about it" (p. 493). In the team training area, practical methodological limitations witnessed in primary studies that rely on relatively small samples of teams generally result in an unacceptably large sampling error – error that can adversely influence the consistency and quality of independent research studies' conclusions. At the same time, disparate effect sizes across primary studies have made it difficult to determine the true strength of the relationships between team training techniques and team outcomes.

The use of meta-analysis can help overcome these issues. Specifically, meta-analytically integrating the results of independent studies allows one to amalgamate, with increased precision and certainty, the significance, strength, and predictable variation of the effects of team training on team functioning. Although recent attempts have been made to empirically (and meta-analytically) review the influence of team development interventions on team functioning (e.g., Salas, Nichols, & Driskell, 2007; Salas, Rozell, Mullen, & Driskell, 1999), this research advances (and expands upon) prior efforts by including more primary studies, employing enhanced methodology, and examining key variables that will help clarify the existing knowledge base on team training.

We should note that most team training evaluations have been conducted in the military and in aviation. These sectors are the primary investors of funding and effort to understand team performance and how to develop and deliver team training. The current research includes team training evaluations from several sectors (i.e., military, aviation, industry, lab/classroom, and medical settings). As will be discussed at a later point, the primary studies included in this meta-analysis are heavily represented by military research.

This meta-analysis also provides clarity and understanding concerning the amorphous empirical literature on team training. In practice, team training appears in many shapes, sizes, and forms and is called by many different names and labels (e.g., assertiveness training, crew resource management, cross-training, group process training, problem-solving training, task-focused simulation training). What has long been needed in this domain is a taxonomic integration and empirical investigation of the utility of these interventions based upon the content or focus of the interventions, rather than upon the myriad of names given to these strategies by independent researchers.

Fortunately, an inclusive taxonomy has already emerged. Upon summarizing work conducted by Salas and colleagues (e.g., Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995; McIntyre & Salas, 1995; Salas, Dickinson, Converse, & Tannenbaum, 1992), Goldstein and Ford (2002) argued that [team] training methods are designed to enhance taskwork, teamwork, and process improvement skills. Meta-analytic reviews have been conducted of team interventions that focused on process improvement skills (i.e., team building; Klein et al., in press; Salas et al., 1999), so there is now a need to empirically analyze the team training literature along the teamwork-taskwork continuum. The current research investigates the unique and relative value of these two distinct focal points of team training.

A second impetus for this research is to investigate moderators of team training effectiveness. These investigations will help provide answers to a number of important questions, including the following:

- Does the effectiveness of team training vary based upon the content of the training?
- Does team training work better with intact versus ad hoc teams?
- Do larger teams benefit more from team training than smaller teams?

Answers to these questions will go a long way toward advancing the literature and providing deeper insights into the effectiveness of team training.

Finally, although there is evidence to suggest that training (and team training) has a significant, positive impact on key measures of business performance (e.g., George, Hannibal, & Hirsch, 2004; Mathieu & Leonard, 1987) and safety (Salas, Burke, Bowers, & Wilson, 2001; Salas, Wilson, Burke, & Wightman, 2006), there is a need to form a more accurate understanding of the impact of team training, specifically. Accurate estimates of the efficacy of team training are important for both researchers and practitioners because tens of millions of dollars are spent annually on team training in military, aviation, health care, and other industry settings.

In short, the present research provides a substantial contribution to the current understanding of the efficacy of team training. In comparison with other quantitative reviews (e.g., Salas, Nichols, et al., 2007) and qualitative reviews (e.g., Salas et al., 2001, 2008), the current effort uses more rigorous methods to assess a greater number of published and unpublished primary studies; distills knowledge from a wider range of interventions, outcomes, and moderator variables; and offers actionable guidance to organizational decision makers. Before describing in detail the research questions and hypotheses posed in the present initiative, a brief discussion of teams, team performance, and team training is provided.

TEAMS, TEAM PERFORMANCE, AND TEAM TRAINING

"Teams embedded in organizations exist to perform tasks" (Ilgen, 1999, p. 131). These tasks are often complex, difficult, and dynamic. They require expertise, experience, and the perspectives of multiple individuals synchronizing their work to perform collectively in the pursuit of common goals (Mohammed & Ringseis, 2001).

For the purposes of this meta-analysis, *teams* are defined as a distinguishable set of two or more individuals who interact dynamically, adaptively, and interdependently; who share common goals or purposes; and who have specific roles or functions to perform (Salas et al., 1992). Teams can be distinguished from work groups by their task, feedback, and goal interdependencies (e.g., Saavedra, Earley, & Van Dyne, 1993) as well as by their formal structure and coordination demands (Tannenbaum, Beard, & Salas, 1992). That is, groups are usually more loosely constituted. They are collectives that may be perceived as social entities, and they may even have common goals, but they also possess task connections that are less well defined (e.g., juries, committees, councils).

For the current research integration, *team performance* is defined as an emergent phenomenon resulting from the goal-directed process whereby members draw from their individual and shared resources to display taskwork processes, teamwork processes, and integrated team-level processes to generate products and provide services (Kozlowski & Klein, 2000; Salas, Stagl, Burke, & Goodwin, 2007).

Frameworks of team performance and effectiveness generally follow an input, throughput, and output format (e.g., Hackman, 1983, 1987; McGrath, 1984; Nieva, Fleishman, & Rieck, 1978; Salas, Stagl, et al., 2007; Steiner, 1972; Tannenbaum et al., 1992) in which inputs might include individual and team characteristics, capabilities, and states; throughputs include team communication, coordination, collaboration, and decision-making processes; and outputs consist of the goods or services produced by a team. Stakeholders make judgments about the effectiveness of the quality, quantity, and timeliness of the products or services produced, as well as about the changes in the team and its members that result as performance unfolds (Hackman, 2002).

Training is a systematic, planned intervention designed to facilitate the acquisition of job-related KSAs (Goldstein & Ford, 2002). Concerning *team training,* one widely cited definition positions it as "a set of tools and methods that, in combination with required [team-based] competencies and training objectives, form an instructional strategy" (Salas & Cannon-Bowers, 1997, p. 254). Task-focused team training enables team members to become aware of, learn about, and practice requisite team competencies (i.e., KSAs) and performance processes while receiving feedback on their performance. Moreover, similar to that of individual training, the science of team training involves identifying the optimal combination of tools (e.g., team task analysis), delivery methods (e.g., practice based, information based, demonstration based), and content (e.g., knowledge, skills, attitudes; Salas & Cannon-Bowers, 1997).

Many narrative and empirical evaluations of team training have been conducted (e.g., Denson, 1981; Dyer, 1984; Marks, Zaccaro, & Mathieu, 2000; Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000; Salas et al., 2001, 2006). Although the majority of these assessments have yielded ample support for team training, differences have been observed in the efficacy with which assorted team training strategies improve team outcomes. In an effort to summarize the relative contributions of various team training strategies, Salas, Nichols, et al. (2007) recently conducted a meta-analysis on the subject. Specifically, they investigated the relative efficacy of three team training strategies: cross-training, coordination and adaptation training, and team self-correction training.

Their database consisted of 28 effect sizes (from seven studies) assessing the efficacy of team training for team performance. For the overall test, there was a low, significant tendency for team training to improve team performance ($r = .29$). However, when they examined the relative efficacy of different team training strategies, cross-training did not appear to be effective ($r = -.09$), whereas the efficacy of self-correction training ($r = .45$) and team coordination and adaptation training ($r = .61$) was confirmed.

Although the findings from the Salas, Nichols, et al. (2007) team training meta-analysis suggest that coordination and adaptation training is effective for producing team performance improvements, some caution should be exercised in interpreting these results and directly comparing them with those of other meta-analyses. The meta-analytic technique that Salas, Nichols, et al. (2007) utilized was based on procedures outlined by Mullen (1989), Mullen and Rosenthal (1985), and Rosenthal (1991). That approach to meta-analytic integration is considerably different (e.g., use of transformations, not calculating confidence intervals) from the procedure we utilized in the present study, which followed the procedure provided in Hunter and Schmidt (2004).

In sum, the findings from qualitative and limited quantitative research on team training interventions have largely supported the effectiveness of these techniques. However, none is as comprehensive, robust, or systematic as this one. This is (to our knowledge) the first thorough research endeavor undertaken to gauge the impact of team training on team outcomes. Given the widespread use of team training in many important settings, it is also a timely and critical integration.

HYPOTHESES: THEIR RATIONALE AND CONCEPTUAL FOUNDATION

As noted, several research questions concerning the importance of team training for enhancing targeted team outcomes guided this effort. In an effort to address these questions, we present testable hypotheses and their theoretical rationale in the following sections. A model that graphically illustrates these research questions is offered in Figure 1.

Does Team Training Work?

The first question posed in this meta-analysis is whether team training works. This general question treats all forms of team training as equivalent and, likewise, treats all targeted outcomes of team training interventions as interchangeable. Although simplistic from a scientific standpoint, this overarching issue is the single most important question posed in the current study. It asks, in a general way, whether team training is useful as a lever to produce change. If the answer to this question is no, then an enormous amount of scientific effort and limited organizational resources have surely been wasted. Moreover, such null results would more broadly call into question the widely accepted premise that planned interventions are useful in some measurable way.

Fortunately, research on instructional design theory and, more recently cognitive, metacognitive, and macrocognitive theory, collectively suggests that team training is indeed useful for prompting meaningful change and in helping teams achieve performance levels deemed effective by organizational stakeholders. Supporting this growing base of theory are the mounting findings from primary

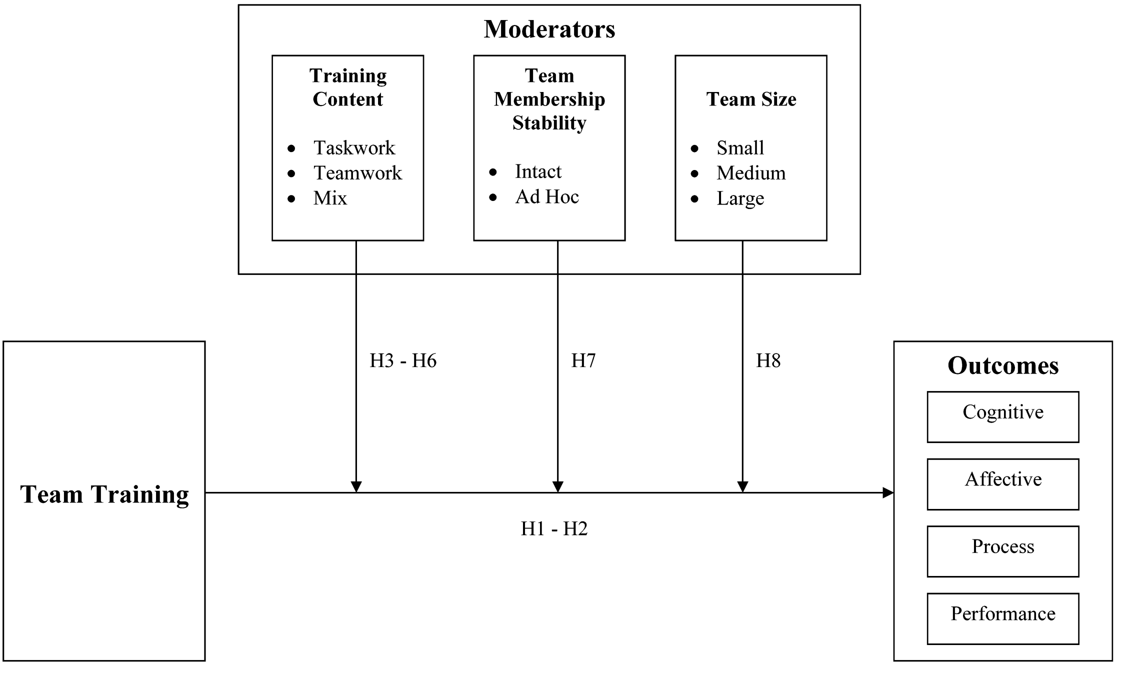

Figure 1. Model depicting study hypotheses (H1–H8).

team training studies (e.g., Ellis, Bell, Ployhart, Hollenbeck, & Ilgen, 2005; E. E. Entin, Serfaty, & Deckert, 1994; Klein, Stagl, Salas, Parker, & Van Eynde, 2007). Given the abundance of theoretical support and empirical evidence speaking directly to the usefulness of team training, the following hypothesis is advanced:

> Hypothesis 1: Team training is positively related to overall team outcomes.

Separating Outcomes

In any meta-analysis, a balance must be struck between casting too wide a net (i.e., including widely disparate independent and dependent variables) and being too restrictive (i.e., reducing the meta-analytic database to a small number of studies). Clearly, results likely to be obtained from the examination of the first research question can be criticized for thoughtlessly combining "apples and oranges" and concluding that "fruit tastes good" (Kraiger, 1985). Thus, there was a need for more finely tuned analyses that examined the relationships between more specific and restrictive definitions of independent and dependent variables. This set of analyses is directed at the next question posed in this study, namely, "Is team training, when conceptualized as a single overarching intervention, an effective technique for producing planned changes in specific types of team-level outcomes?"

This issue is important, as team researchers have offered compelling propositions about (Marks, Mathieu, & Zaccaro, 2001), and meta-analytic evidence for (LePine, Piccolo, Jackson, Mathieu, & Saul, 2008), behavioral and cognitive processes constituting teamwork. Team performance emerges as teams and their members draw upon individual and shared characteristics, capabilities, and cognitive and affective states to enact the core processes that constitute teamwork (Burke, Stagl, Salas, et al., 2006). In turn, the performance outcomes resulting from this unfolding, recursive process are evaluated by organizational stakeholders in terms of their effectiveness (Hackman, 2002). If the utility of team training is to be fully realized, it must be a lever to affect all of the core components of teamwork and produce performance outcomes that are deemed effective in the workplace.

Gagné (1962) astutely argued that the most fundamental design issue in training is the specification of what is to be learned. This is the specification of training or instructional objectives. These objectives specify the knowledge, skills, beliefs, attitudes, or choice behaviors that a trainee or trainees should be capable of accomplishing upon successful completion of the training program (Campbell & Kuncel, 2001; Goldstein & Ford, 2002). Team training is no different from individual training in this regard. Once these objectives are chosen, the training environment is tailored to achieve them in the most efficient way possible. Thus, instructional objectives provide input for the design of training programs as well as the criteria used to judge the programs' adequacy (Goldstein & Ford, 2002).

Because the primary studies included in the present meta-analytic initiative addressed a wide variety of instructional objectives, a meaningful taxonomy for organizing training outcomes was needed to guide this effort. In line with the team theory noted previously and with prior team training meta-analyses (e.g., Salas, Nichols, et al., 2007) – and consistent with Kraiger, Ford, and Salas's (1993) taxonomy – training outcomes can be classified as cognitive, affective, and performance outcomes as well as behavioral processes. Each of the training outcomes investigated in the primary studies included in the present meta-analysis were classified into one of these four taxons.

In the present study, team member cognitive outcomes predominantly consisted of declarative knowledge gains from pre- to posttest measurement. Team member affective outcomes included socialization, trust, confidence in team members' ability, and attitudes concerning the perceived effectiveness of team communication and coordination processes. Team processes examined in this review included behavioral measures of communication, coordination, strategy development, self-correction, assertiveness, decision making, and situation assessment. Finally, the assessment of team performance integrated quantity, quality, accuracy, efficiency, and effectiveness outcomes. To help answer the question posed at the beginning of this section, the following hypotheses are advanced:

> Hypotheses 2a to 2d: Team training is positively related to team-level (a) cognitive outcomes, (b) affective outcomes, (c) behavioral processes, and (d) performance outcomes.

Training Content

It is interesting to know whether team training is useful for improving specific team outcomes, but it is perhaps more important to ask, "What type of team training is most effective for improving team functioning?" To this end, the current research addresses the relative influence of specific training interventions on valued team outcomes. Although particular team training interventions are often referred to by many different names, it can be argued that every team training intervention targets some combination of taskwork or teamwork skills (e.g., Cannon-Bowers et al., 1995; Goldstein & Ford, 2002; McIntyre & Salas, 1995; Salas et al., 1992). Thus, three content categories were coded in the present study: taskwork, teamwork, and interventions that utilized a mix of taskwork and teamwork content. Illuminating the relationships between these three types of team training and each of the four types of outcomes noted previously is important to the objective of the present study.

Team training interventions targeting taskwork KSAs seek to develop the technical competencies of team members (Salas et al., 1992). Cross-training is a common example of a team training intervention that includes a heavy taskwork component. In cross-training, members learn about and sometimes practice the responsibilities of multiple positions and/or roles, including those of their teammates. Having acquired interpositional knowledge and honed teamwork KSAs, individuals who have been cross-trained are better prepared to recognize when their teammates are overburdened and to step in and perform their coworkers' duties when called upon.

In contrast, training interventions targeting teamwork KSAs are focused on improving how individuals work together effectively as a team. For example, teamwork skills targeted by team training may include mutual performance monitoring, feedback, leadership, management, coordination, communication, and decision making (e.g., Cannon-Bowers et al., 1995; Stagl, Salas, & Fiore, 2007). Team coordination and adaptation training is one example of a team training intervention that is focused on providing teamwork skills content.

Hypotheses 3 to 5 expand upon the basic questions addressed in the earlier hypotheses because, assessed in their entirety, they provide preliminary estimates for determining which configurations of team training content are better than others. However, given the absence of a priori expectations for this set of hypotheses, the investigation into the moderating impact of training content is exploratory. Nonetheless, the following hypotheses are advanced to investigate these questions:

> Hypotheses 3a to 3d: Taskwork training is positively related to enhancements in team-level (a) cognitive outcomes, (b) affective outcomes, (c) behavioral processes, and (d) performance outcomes.

> Hypotheses 4a to 4d: Teamwork training is positively related to enhancements in team-level (a) cognitive outcomes, (b) affective outcomes, (c) behavioral processes, and (d) performance outcomes.

> Hypotheses 5a to 5d: Team training that includes a combination of teamwork and taskwork content is positively related to enhancements in team-level (a) cognitive outcomes, (b) affective outcomes, (c) behavioral processes, and (d) performance outcomes.

One truism of collectives is that teams of experts often fail to evolve into expert teams (Salas, Cannon-Bowers, & Johnston, 1997). In other words, the possession of copious amounts of task knowledge is insufficient to overcome the performance decrements that can result from the coordination demands teamwork imposes on team members.

In order to assist teams in fulfilling their potential, it is often necessary to move beyond the provision of mere task-specific content to feature teamwork and team interaction training as well (Hollenbeck, DeRue, & Guzzo, 2004). For example, a team of medical professionals that possesses a high degree of skill in mutual performance monitoring will not only have the task expertise to offer suggestions but will also know when to do so in a timely manner. Therefore, team training that targets both task-specific content and generic teamwork competencies is expected to provide the greatest benefit in terms of enhanced team outcomes. Given this line of thinking, the following hypothesis is advanced:

> Hypothesis 6: Team training that includes a combination of teamwork and taskwork content will be more effective than interventions that target either phenomenon in isolation.

Team Membership Stability

The aforementioned research questions will help illuminate the importance of team training as a lever for facilitating change in a variety of team-level phenomena. However, it is important to move beyond these preliminary questions about the main effects of team training to begin to address other pertinent questions, such as, "Under what conditions is team training most effective as a planned intervention?" Thus, in the remainder of this section, questions are posed and hypotheses are advanced to address additional variables that may moderate the effect of team training on team-level outcomes.

Of primary importance to the present meta-analysis are the potential moderating roles of team membership stability and team size. Membership stability addresses the length of time team members have worked together interdependently to solve joint challenges, and team size indexes the number of team members assigned to a given collective.

In regard to membership stability, studies of team training interventions generally consist of two varieties: (a) studies of training interventions conducted with intact teams, whose members have a shared history as a result of a commonly held assignment to a given collective operating inside an organization (e.g., an energy exploration team); and (b) studies of interventions delivered to a group of ad hoc strangers purposively assembled to conduct either basic or applied research in a contrived setting (e.g., a tank crew composed of randomly selected and assigned students performing in a university lab). In comparison with their ad hoc counterparts, intact teams tend to have relatively stable membership and a shared history of working together (Devine et al., 1999). The total sample of primary studies included in the present meta-analysis is representative of both intact and ad hoc teams.

In their recent meta-analysis on team training, Salas, Nichols, et al. (2007) made a conscious decision to examine the effects only of interventions that were delivered to intact teams. However, these researchers also suggested that future meta-analytic efforts should examine levels of team membership stability as a potential moderator variable. We accept that call and will investigate whether team training works better for intact or ad hoc teams.

Based upon prior research addressing the stages of team development, it seems plausible that team training will be more effective for intact teams. This is because intact collectives presumably have already navigated at least some of the process challenges that characterize the preliminary, and often rocky, stages of group development (e.g., storming, norming [Tuckman, 1965]) that newly formed teams must overcome to develop cohesion. The baseline level of shared expectations, competencies, and emergent states developed from a shared history will likely free up cognitive and interpersonal resources, allowing intact teams to focus more of their resources on skill acquisition during team training. Given these arguments, the following hypothesis is advanced:

> Hypothesis 7: Team training will be more effective for intact teams than for ad hoc teams.

Team Size

This research also examines the relative impact of team training for teams of varying sizes. It has been noted by several researchers that teams are highly effective when they have a sufficient, but not greater than sufficient, number of members to perform team tasks (Guzzo & Shea, 1992; Hackman, 1990). The definition of *sufficient,* however, is somewhat unclear, in part because the research findings on the impact of team size on effectiveness have been mixed.

Some studies have found that larger teams are more effective (Magjuka & Baldwin, 1991; Yetton & Bottger, 1982), whereas other studies have reported that larger teams suffer from coordination and process losses (Gooding & Wagner, 1985; Mullen, Symons, Hu, & Salas, 1989). Kozlowski and Bell (2003) summarized these mixed results, concluding that the benefits of larger teams are dependent on the nature of the task and the team's environment. For example, one study that evaluated team size and its impact on the performance of tasks that required innovativeness found a negative correlation between team size and team processes, $r = -.18$ to $r = -.51$ (Curral, Forrester, Dawson, & West, 2001).

Also bearing on this issue is the seminal work of Steiner (1972), who examined how increasing the size of the team impacted team processes. Steiner's (1972) findings suggested that an increase in team size resulted in increases in productivity but also introduced inefficiencies (e.g., decreased motivation, poorer decision making, poorer coordination, and higher levels of conformity). Other studies have also found additional detrimental effects attributable to large-sized teams – for example, an in-group bias effect (Mullen, Brown, & Smith, 1992), a participation leadership effect (Mullen, Salas, & Driskell, 1989), a decrease in group liking (Indik, 1965), and a decrease in performance (Mullen, 1987).

A common undesirable group phenomenon is social loafing. It has been found that larger groups are more prone to social loafing (Latané, 1981). Because larger groups lose the connection between their inputs and the rewards that are received, individuals in large teams become less motivated to perform (Kidwell & Bennett, 1993). Karau and Williams (1993) discussed what they termed the "dilution" effect. This effect manifests itself in a variety of dysfunctional ways (e.g., free riding, getting lost in a crowd, shirking group work), all of which are associated with an increase in team size.

A study by Hackman and Vidmar (1970) set out to find the perfect size of a team. Based on questions asked of teams large and small, they found the optimal number of team members to be 4.6. Moreover, in a recent study by De Dreu (2007), a weak correlation was found between team size and the measured outcomes of learning ($r = .15$) and team effectiveness ($r = .16$). De Dreu also found a negative correlation between team size and information sharing ($r = -.11$). These results show that the key to enhancing team functioning does not necessarily lie in increasing the size of the team.

The current research categorized teams into three groups; small teams ($n = 2$), medium teams ($2 < n < 5$), and large teams ($n \geq 5$). A recent meta-analysis on team building interventions found that large teams benefited the most from team building interventions, resulting in a mean true score correlation of .66 (Klein et al., in press). Given this finding and past findings suggesting that medium-sized teams will already be performing at a high level, we suggest that team training interventions will show greater effects for dyads and large teams. In other words, team training, whether it is geared toward teamwork or taskwork, will aid small and large teams more than medium-sized teams in improving their inefficiencies. Given this thinking, the following hypothesis is proposed:

> Hypothesis 8: Team training will be more beneficial for small and large teams than for medium-sized teams.

METHOD

This study applied meta-analytical techniques to previously conducted research in an effort to obtain a numerical estimate of the relationships between team training and team outcomes. Analyses were also conducted in order to determine the existence of moderators. The following sections describe in detail the identification, selection, and coding procedures for primary studies. In addition, a description of the method for calculating effect sizes is provided.

Literature Search

More than a half century ago, B. Glass (1955) observed that "no problem facing the individual scientist is more defeating than the effort to cope with the flood of published scientific research, even within one's own narrow specialty" (p. 583). And even within the relatively narrow specialty of team training, there is indeed a rising tide of empirical research. Surfing this tide required a calculated line of attack. In total, four distinct and diverse approaches were used to identify studies for potential inclusion in the meta-analytic database.

First, comprehensive electronic searches of computerized databases were conducted using multiple combinations of appropriate keywords (e.g., *teams*, *cross-training, team performance;* a full list of keywords used can be obtained from the authors). Specifically, the search engines, electronic databases, abstracting services, and proceedings of Google Scholar, ScienceDirect, EBSCOhost, Academic Search Premier, Business Source Premier, PsycINFO, PsycARTICLES, Military and Government Collection, SPORTDiscus, the Interservice/Industry Training, Simulation and Education Conference, and the Defense Technical Information Center were searched for articles published through August 2008.

Second, a targeted electronic search of the following journals was conducted, using the keywords *team training* and *group training,* with no restriction on dates: *Journal of Applied Psychology, Personnel Psychology, Military Psychology, Organization Development Journal, Academy of Management Journal, Small Group Research,* and *Human Factors*. Third, an ancestry approach was employed, whereby the bibliographies and reference sections of relevant studies that had already been retrieved were searched to locate earlier relevant studies. Finally, an informal network of colleagues was queried for potential unpublished articles, manuscripts, and conference papers. The complete study identification and search process resulted in more than 500 articles for possible inclusion.

In the next phase of the literature review, two authors reviewed abstracts and online versions of the identified articles and came to a consensus concerning whether the full text document should be obtained. At this stage, articles containing the following characteristics were excluded from further analyses: (a) master's theses; (b) use of clinical populations; (c) use of children as participants; (d) collectives and groups characterized by a deficient degree of task interdependence; (e) studies failing to report a usable test statistic (e.g., F, t, χ^2, z, d, r) or the raw data necessary to calculate these statistics (e.g., means, standard deviations, sample size); and (f) studies failing to utilize an actual team intervention.

As a result of this effort, 168 relevant published and unpublished papers were identified. These articles, book chapters, dissertations, technical reports, and conference proceedings were subsequently obtained in full text and coded by study authors. A discussion of the specific coding strategy employed for these articles is presented next.

Coding Procedure

Various meta-analytic integrations often use a similar coding scheme to quantify study characteristics and results (Lipsey & Wilson, 2001; Stock, 1994). The coding strategy used in the present research included capturing 20 pieces of information from each study: (a) publication type, (b) study setting, (c) study design type, (d) nature of the organization and participant sample, (e) team type, (f) team membership stability, (g) task interdependence, (h) number of teams, (i) average team size, (j) predictor level of analysis, (k) criterion level of analysis, (l) training content, (m) name and description of training intervention, (n) method of training intervention, (o) criterion report type, (p) criterion description, (q) reliability of predictor measures, (r) reliability of criterion measures, (s) effect size or sizes, and (t) recommendation for inclusion.

For the criterion level of analysis coding, the initial database consisted of 19 effect sizes at the individual level and 93 at the team level. Although it is desirable to include in the subgroup analyses the largest possible number of effect sizes, researchers have strongly recommended against mixing levels of analysis in research integrations (e.g., Gully, Devine, & Whitney, 1995; Hunter & Schmidt, 1990). For the purposes of this research, only team-level outcomes were considered eligible for analyses.

Rater Reliability

One or more study authors independently coded each of the 168 articles originally identified for possible inclusion in the meta-analysis. The coders for this research had previously pilot tested the coding scheme and possessed an adequate level of expertise in the substantive areas being coded. Moreover, before coding began, the authors met to discuss and develop detailed instructions for coding the 20 categories of information from primary studies. More specifically, a resource manual was developed in order to provide the coders with standardized information to code the primary studies. Coders met in order to discuss the coding manual and ensure a shared understanding of the concepts being coded. Furthermore, the coders were trained by coding two articles and then holding a meeting to discuss the information extracted from the articles. Once a satisfactory level of understanding of the categories was reached by all coders, the coding for the present meta-analysis began.

The coding was done in stages, with two checks performed to assess interrater reliability. First, two authors and one of their colleagues (who was blind to the study hypotheses) each independently coded 15 articles from the database. At this point, the interrater reliability for the three coders was assessed by calculating intraclass correlation coefficients (ICCs; Nunnally, 1978; Shrout & Fleiss, 1979). The ICC is a measure of reliability among two or more raters. An ICC can be conceptualized as the ratio of between-group variance to total variance. When an ICC approaches 1 it can be interpreted that the variance between the raters is essentially zero; thus the raters give the same ratings. Furthermore, the variation in the ratings or measurements is attributable solely to the target being rated.

The reliability at this stage was deemed satisfactory (ICCs ranged from .75 to .85) and the full coding procedure continued. The second check on interrater reliability was performed near the end of the coding process, when 10 more articles were chosen to be coded by each of the three coders. The reliability of the coding for these final 10 articles was assessed at the end of the coding process and again was acceptable (ICCs ranged from .85 to 1.0). Taken together, 25 of the 168 articles were coded by two authors and a colleague. These 25 articles represented approximately 15% of the total number of articles.

After the 25 reliability sample articles were coded, the combined codings were assessed for rater reliability. We were particularly interested in the coding for the data that would be contributing to the investigation of the hypotheses. We were also keenly concerned with whether the coders reliably coded the same effect sizes and whether they made the same decisions in regard to including the article in the final database. Upon examination, the coding of the data used for the hypotheses was satisfactory.

Specifically, for the criterion description category, the ICC (3, k) = .89. Similarly, for training content, team membership stability, and team size, the ICCs (3, k) were equal to .77, .94, and 1.00, respectively. Moreover, the coding for the effect size or sizes and inclusion/exclusion decisions were highly reliable (ICCs = .99 and .89, respectively). These estimates of the reliability with which 25 common articles were coded gave us confidence that the coding for this research was performed reliably. The remaining 143 articles were coded independently by one of the two authors who also served as coders for the reliability sample. Any initial ambiguity or uncertainty about whether to include a particular study result or characteristic in the final meta-analytic database was resolved through a consensus discussion by two or more study authors.

Meta-Analysis Procedure

The data analysis for this research was aided by the Hunter-Schmidt Meta-Analysis Programs 1.1 (Schmidt & Le, 2005). Effect sizes were first sorted within their associated subgroups (i.e., outcome type, training content, team membership stability, team size), and then a combined effect size estimate was generated for each subgroup and/or level of each moderator (e.g., small, medium, and large teams).

This software utilizes a random effects model, rather than a fixed effects model, to analyze the data. This model allows the true effect sizes to vary, instead of assuming that the true effect sizes have fixed values. Field (2001) noted that the random effects model is considered to be "more realistic than the fixed-effect model on the majority of occasions" (p. 162). The data output for the program includes (but is not limited to) the mean true score correlation, the standard deviation and variance of true score correlations, credibility intervals, and the percentage of variance attributable to observed correlations after the removal of artifacts. Confidence interval estimates were calculated separately in Excel using formulas provided by Hunter and Schmidt (2004).

Effect Size Calculations

In order to aggregate findings across studies, all effect size estimates culled from primary studies were converted to a common metric, r (correlation). Formulas found in Hunter and Schmidt (2004) were used to transform primary study effect sizes that had been reported as other statistics (e.g., t, F, d, χ^2, or Z). Upon being placed on this common metric of effect size (i.e., r), the primary study results can be pooled and evaluated for fit with the predicted hypotheses in this research.

An important issue that presented itself was that several studies contained more than one effect size estimate. This circumstance occurred for two reasons. First, several studies reported findings for separate samples under investigation. In this instance, it is acceptable to include the findings as separate effect sizes in the database. Alternatively, a large number of studies reported multiple effect sizes from single samples that were either (a) for the same outcomes or (b) for different outcomes. For example, a study might have reported two process outcomes and one performance outcome.

To be clear, however, stochastically dependent effect sizes were not included in the current database (which, if included, would patently violate the assumption of independent effect sizes). In other words, findings generated on the same (or similar) outcomes from the same sample were averaged prior to entry in the database. In contrast, when studies reported findings from the same sample on different outcomes, these findings were included separately in the overall database but were not included in any of the same subgroup analyses. Effect size estimates from primary studies that were combined prior to entry in the meta-analytic database are denoted with an asterisk in Table 1.

Corrections for Unreliability and Range Restriction

Meta-analysis is especially useful for determining whether conflicting results in the literature are attributable to artifactual or actual variation (G. V. Glass, 1976; G. V. Glass, McGaw, & Smith, 1981;

text continues on page 917

TABLE 1: Studies and Effect Sizes Included in the Meta-Analytic Database

Author(s)	Year	r^a	No. of Teams	Average Team Size	Criterion Description	Membership Stability	Training Content
Adelman, Christian, Gualtieri, & Bresnick	1998	.28*	26	2.00	Process	Ad hoc	Mix
	1998	−.17*	28	2.00	Performance	Ad hoc	Mix
Alexander, Kepner, & Tregoe	1962	.81*	4	13.00	Performance	Ad hoc	Mix
Banks, Hardy, Scott, Kress, & Word	1977	.64*	16	9.00	Performance	Intact	Taskwork
Blickensderfer, Cannon-Bowers, & Salas	1997	.45*	40	3.00	Process	Ad hoc	Teamwork
	1997	.03	39	3.00	Performance	Ad hoc	Teamwork
Brannick, Prince, & Salas	2005	.59*	48	2.00	Process	Ad hoc	Mix
	2005	.12*	48	2.00	Performance	Ad hoc	Mix
Brown	2003	.40	41	4.50	Affective	Ad hoc	Teamwork
	2003	.29	41	4.50	Performance	Ad hoc	Teamwork
Buzaglo & Wheelan	1999	.63	2	13.00	Affective	Intact	Teamwork
	1999	.69	2	11.00	Process	Intact	Teamwork
	1999	.85*	2	13.00	Process	Intact	Teamwork
	1999	.78*	2	11.00	Performance	Intact	Teamwork
	1999	.66*	2	13.00	Performance	Intact	Teamwork
Cannon-Bowers, Salas, Blickensderfer, & Bowers	1998	.76	36	3.00	Cognitive	Ad hoc	Mix
	1998	.32*	36	3.00	Process	Ad hoc	Mix
	1998	.39*	36	3.00	Performance	Ad hoc	Mix
Cobb	2000	.25*	36	4.00	Process	Ad hoc	Teamwork
	2000	.28	36	4.00	Performance	Ad hoc	Teamwork
Cooke, Kiekel, & Helm	2001	.82*	10	3.00	Cognitive	Ad hoc	Mix
	2001	.92	9	3.00	Performance	Ad hoc	Mix
Cooke et al. (Cooke et al., 2000)[b]	2003	.48*	33	3.00	Cognitive	Ad hoc	Mix
Cooke et al.	2003	.53	33	3.00	Performance	Ad hoc	Mix
	2003	.43	9	3.00	Performance	Ad hoc	Mix
Duncan et al.	1996	.76	8	5.00	Process	Intact	Teamwork
	1996	.88*	7	5.00	Performance	Intact	Teamwork
Ellis, Bell, Ployhart, Hollenbeck, & Ilgen	2005	.56	65	4.00	Cognitive	Ad hoc	Teamwork
	2005	.31*	65	4.00	Process	Ad hoc	Teamwork
E. B. Entin, Entin, MacMillan, & Serfaty	1993	.58	16	2.00	Process	Intact	Taskwork

[a]Effect sizes (*r*) denoted with an asterisk (*) represent an average for a single sample and a single outcome and have been combined for this meta-analysis. [b]Four effects sizes indexing taskwork knowledge from the Cooke et al. (2000) proceedings paper were added to three teamwork knowledge effect sizes from the Cooke et al. (2003) journal article. These cognitive effect sizes were averaged because they were based upon the same sample of participants.

Continued on next page

TABLE 1 (continued)

Author(s)	Year	r[a]	No. of Teams	Average Team Size	Criterion Description	Membership Stability	Training Content
E. E. Entin & Serfaty	1999	.73*	6	5.00	Process	Ad hoc	Mix
	1999	.63	4	5.00	Performance	Ad hoc	Mix
	1999	.57	4	5.00	Performance	Ad hoc	Mix
E. E. Entin, Serfaty, & Deckert	1994	.53*	4	5.00	Performance	Ad hoc	Teamwork
Freeman, Cohen, & Thompson	1998	.65*	14	2.00	Process	Ad hoc	Mix
	1998	.55	13	2.00	Performance	Ad hoc	Mix
	1998	.75	13	2.00	Performance	Ad hoc	Mix
Ganster, Williams, & Poppler	1991	.05	43	4.70	Performance	Ad hoc	Teamwork
Glickman et al.	1987	.41	13	5.33	Performance	Intact	Taskwork
Green	1994	.811	4	5.00	Process	Adhoc	Teamwork
	1994	.650	4	5.00	Performance	Adhoc	Teamwork
Heffner	1997	.19*	74	2.00	Performance	Ad hoc	Mix
Henry	1997	.28	10	4.70	Performance	Ad hoc	Teamwork
Ikomi, Boehm-Davis, Holt, & Incalcaterra	1999	.54	44	2.50	Performance	Intact	Teamwork
Lampton et al.	2001	.18	49	2.00	Performance	Ad hoc	Taskwork
Lassiter, Vaughn, Smaltz, Morgan, & Salas	1990	.42*	45	2.00	Performance	Ad hoc	Teamwork
Le Blanc, Hox, Schaufeli, Taris, & Peeters	2007	–.01*	29	23.00	Affective	Intact	Teamwork
	2007	.10*	29	23.00	Process	Intact	Teamwork
Leedom & Simon	1995	.53	16	2.00	Affective	Ad hoc	Mix
	1995	.62	16	2.00	Process	Ad hoc	Mix
	1995	.61*	16	2.00	Performance	Ad hoc	Mix
	1995	.10	15	2.00	Affective	Ad hoc	Mix
	1995	.91	8	2.00	Process	Ad hoc	Mix
	1995	.60*	15	2.00	Performance	Ad hoc	Mix
MacMillan, Entin, Entin, & Serfaty	1994	.57	8	2.00	Affective	Intact	Teamwork
	1994	.48*	8	2.00	Process	Intact	Teamwork
Marks, Sabella, Burke, & Zaccaro	2002	.43	45	3.00	Cognitive	Ad hoc	Taskwork
	2002	.36	49	3.00	Cognitive	Ad hoc	Taskwork
Marks, Zaccaro, & Mathieu	2000	.44*	75	3.00	Cognitive	Ad hoc	Mix
Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers	2000	.07*	56	2.00	Cognitive	Ad hoc	Mix
	2000	.35*	56	2.00	Process	Ad hoc	Mix
	2000	.19*	56	2.00	Performance	Ad hoc	Mix

[a]Effect sizes (r) denoted with an asterisk (*) represent an average for a single sample and a single outcome and have been combined for this meta-analysis.

Continued on next page

TABLE 1 (continued)

Author(s)	Year	r[a]	No. of Teams	Average Team Size	Criterion Description	Membership Stability	Training Content
Minionis	1995	.32*	94	3.00	Cognitive	Ad hoc	Teamwork
	1995	.10*	90	3.00	Affective	Ad hoc	Teamwork
	1995	.17*	89	3.00	Process	Ad hoc	Teamwork
	1995	.32*	90	3.00	Performance	Ad hoc	Teamwork
Morey et al.	2002	.84	6	3.00	Affective	Intact	Teamwork
	2002	.85	6	3.00	Process	Intact	Teamwork
	2002	.43	6	3.00	Performance	Intact	Teamwork
Naylor & Briggs	1965	.46*	32	3.00	Performance	Ad hoc	Taskwork
Parsons	1981	.44	12	4.00	Affective	Ad hoc	Teamwork
	1981	.61*	12	4.00	Performance	Ad hoc	Teamwork
Pearsall, Ellis, & West	2004	.00*	54	4.00	Cognitive	Ad hoc	Taskwork
Prichard & Ashleigh	2007	.81	16	3.00	Cognitive	Ad hoc	Mix
	2007	.41	16	3.00	Affective	Ad hoc	Mix
	2007	.54*	16	3.00	Process	Ad hoc	Mix
	2007	.26*	16	3.00	Performance	Ad hoc	Mix
Rapp & Mathieu	2007	.52	15	3.38	Performance	Ad hoc	Teamwork
Rollins & Angelcyk	1995	.43	54	2.50	Affective	Intact	Teamwork
	1995	.38	135	2.50	Affective	Intact	Teamwork
	1995	.26	13	2.50	Affective	Intact	Teamwork
	1995	.62	5	2.50	Affective	Intact	Teamwork
Schreiber et al.	2002	.64*	2	20.00	Affective	Intact	Mix
Serfaty, Entin, & Johnston	1998	.63*	11	5.00	Process	Ad hoc	Teamwork
	1998	.61*	11	5.00	Performance	Ad hoc	Teamwork
Shapiro et al.	2004	.39	4	5.00	Process	Intact	Teamwork
Siegel & Federman	1973	.27	32	6.00	Performance	Intact	Mix
Stout, Salas, & Fowlkes	1997	.41	21	2.00	Cognitive	Ad hoc	Teamwork
	1997	.37	21	2.00	Affective	Ad hoc	Teamwork
	1997	.62*	21	2.00	Process	Ad hoc	Teamwork
	1997	.68*	21	2.00	Performance	Ad hoc	Teamwork
Volpe, Cannon-Bowers, Salas, & Spector	1996	.36*	40	2.00	Process	Ad hoc	Taskwork
	1996	.39*	40	2.00	Performance	Ad hoc	Taskwork

[a]Effect sizes (r) denoted with an asterisk (*) represent an average for a single sample and a single outcome and have been combined for this meta-analysis.

Hunter & Schmidt, 2004; Hunter, Schmidt, & Jackson, 1982). Part of this determination stems from an assessment of the amount of unreliability and range restriction that is present in the data. For example, it is standard practice for authors of meta-analyses to attempt to correct obtained reliability coefficients for unreliability in measures of the predictor and criterion (e.g., Hunter & Schmidt, 2004). For the current research, it was an unfortunate reality that primary studies often failed to report all of the auxiliary information necessary to perform corrections for study artifacts. Criterion reliability estimates were available for only 44% of the effect sizes in the database (41 of 93). Thus, in order to best deal with this issue, we utilized artifact distribution meta-analysis, rather than performing corrections for unreliability on each individual effect size.

In addition to making corrections for the unreliability of predictors and criteria in primary studies, researchers using meta-analytic techniques often attempt to account for the effects of direct or indirect range restriction. In a recent article, Hunter, Schmidt, and Le (2006) argued that correcting for range restriction will generally result in more accurate combined estimates of the relationships between variables. However, the articles included in the current database did not report the necessary information on restricted and unrestricted samples that would be needed to perform corrections for range restriction. As a result, we could not make these corrections in the current research, and our effect size estimates may be conservative. That is, they may underestimate the true nature of the relationships between team training and team functioning if the biasing effects of range restriction could be corrected.

Weighting

It is often recommended that studies with larger sample sizes receive more weight in meta-analyses in order to increase the precision of the estimate of the average effect size across studies (Hunter & Schmidt, 2004; Shadish, Cook, & Campbell, 2002). Although estimates obtained from studies with larger samples are not necessarily more valid, they are presumed to be more stable (i.e., accurate). At the same time, recent simulation studies have found that one can obtain a very accurate combined effect size through procedures employed to weight studies by sample size (e.g., Field, 2005). Taking these considerations into account, the primary study effect sizes that were utilized in the current series of meta-analyses were weighted by their sample sizes.

RESULTS

Description of the Database

The final database for this research consisted of a total of 93 correlations obtained from 45 primary studies. These 93 effect sizes represented 2,650 teams. Of these, 1,660 teams were teams from the lab or classroom setting, 762 teams were from the military sector, 138 teams were aviation teams, 80 teams were medical teams, and 10 teams were from business organizations. It should be noted these 93 effect sizes were not all from independent samples (there were 52 independent samples). However, every subgroup analysis that was performed included only independent samples. Of the 45 studies included in the meta-analytic database, 31 were published (25 journal articles, 4 conference proceedings, and 2 book chapters) and 14 were unpublished (3 conference papers, 6 technical reports, and 5 dissertations). Table 1 summarizes some of the key information that was recorded from each primary study.

Meta-Analytic Results

The meta-analytic results for the four primary areas of investigation are reported in Tables 2 through 5. In each of these tables, key pieces of information from each analysis are displayed. This information includes the number of teams in each analysis (N), the number of independent effect sizes (correlations) in each analysis (k), the mean weighted observed correlation ($\bar{r}$), and the 80% confidence interval for that correlation. In addition, the tables display the estimated true score correlation (ρ), the standard deviation of this true score correlation (SD_ρ), the 80% credibility interval (10% CV and 90% CV), and the percentage of variance accounted for by statistical artifacts.

It is important to explain why both confidence and credibility intervals are reported with this research. In general, they each provide information that contributes to the process of estimating the true nature of the relationships between two variables (Whitener, 1990). However, confidence intervals are applied to observed scores, are centered upon a single mean score, and take into account the effects of sampling error. Alternatively, credibility intervals provide information concerning

text continues on page 922

TABLE 2: Analysis of the Effectiveness of Team Training Based Upon Outcome Type

Outcome Type	N	k	$\bar{r}$	CI_r 10%	CI_r 90%	ρ	SD_ρ	10% CV	90% CV	% Var. Accounted for by Artifacts
Cognitive	554	12	.38	.30	.46	.42	.19	.18	.67	34.87
Affective	465	16	.32	.26	.37	.35	.00	.35	.35	100.00[b]
Process	607	25	.39	.34	.44	.44	.00	.44	.44	100.00[b]
Performance	1,024	40	.33	.29	.37	.39	.09	.27	.50	86.63
All outcomes[a]	1,563	52	.34	.31	.37	.34	.00	.34	.34	100.00[b]

Note. N = number of teams; k = number of correlations coefficients on which each distribution was based; $\bar{r}$ = mean observed correlation; CI_r 10% = lower bound of the confidence interval for observed r; CI_r 90% = upper bound of the confidence interval for observed r; ρ = estimated true correlation between the predictor construct and the relevant criterion (fully corrected for measurement error in both the predictor and criterion); SD_ρ = estimated standard deviation of the true correlation; 10% CV = lower bound of the credibility interval for each distribution; 90% CV = upper bound of the credibility interval for each distribution; % Var. Accounted for by Artifacts = percentage of observed variance accounted for by statistical artifacts.

[a]This entry represents the entire database of 52 independent effect sizes, which represented 93 analyses for distinct outcomes. However, in order to estimate an overall effect of team training on team functioning, we needed to combine multiple effect size estimates that were derived from single samples. [b]These entries represent cases in which the percentage of variance accounted for by artifacts resulted in a number greater than 100. This can be interpreted such that all the variance between effect sizes from primary studies could be accounted for by study artifact. We direct the reader to the Meta-Analytic Results section for a more detailed explanation.

TABLE 3: Analysis of the Effectiveness of Team Training Based Upon Training Content

Training Content	Outcome Type	N	k	$\bar{r}$	CI_r 10%	CI_r 90%	ρ	SD_ρ	10% CV	90% CV	% Var. Accounted for by Artifacts
Taskwork	Cognitive	242	4	.28	.18	.37	.30	.10	.17	.43	62.13
	Affective	90	1	.10	.10	.10	.11	.00	.11	.11	—
	Process	145	3	.27	.16	.37	.28	.02	.26	.31	98.23
	Performance	240	6	.35	.28	.41	.35	.00	.35	.35	100.00[a]
Teamwork	Cognitive	86	2	.52	.46	.58	.52	.00	.52	.52	100.00[a]
	Affective	326	11	.37	.32	.43	.41	.00	.41	.41	100.00[a]
	Process	236	13	.40	.33	.47	.44	.00	.44	.44	100.00[a]
	Performance	374	17	.35	.28	.42	.38	.08	.29	.48	89.54
Mix	Cognitive	226	6	.45	.31	.58	.51	.25	.20	.83	27.18
	Affective	49	4	.36	.25	.48	.36	.00	.36	.36	100.00[a]
	Process	226	9	.47	.40	.54	.56	.00	.56	.56	100.00[a]
	Performance	410	17	.30	.23	.38	.40	.15	.20	.59	75.17

Note. N = number of teams; k = number of correlations coefficients on which each distribution was based; $\bar{r}$ = mean observed correlation; CI_r 10% = lower bound of the confidence interval for observed r; CI_r 90% = upper bound of the confidence interval for observed r; ρ = estimated true correlation between the predictor construct and the relevant criterion (fully corrected for measurement error in both the predictor and criterion); $SD\rho$ = estimated standard deviation of the true correlation; 10% CV = lower bound of the credibility interval for each distribution; 90% CV = upper bound of the credibility interval for each distribution; % Var. Accounted for by Artifacts = percentage of observed variance accounted for by statistical artifacts.

[a]These entries represent cases in which the percentage of variance accounted for by artifacts resulted in a number greater than 100. This can be interpreted such that all the variance between effect sizes from primary studies could be accounted for by study artifact. We direct the reader to the Meta-Analytic Results section for a more detailed explanation.

TABLE 4: Analysis of the Effectiveness of Team Training Based Upon Team Membership Stability

Membership Stability	Outcome Type	N	k	$\bar{r}$	CI_r 10%	CI_r 90%	ρ	SD_ρ	10% CV	90% CV	% Var. Accounted for by Artifacts
Intact	Cognitive	—	—	—	—	—	—	—	—	—	—
	Affective	254	9	.37	.30	.44	.41	.00	.41	.41	100.00[a]
	Process	75	8	.43	.30	.55	.48	.00	.48	.48	100.00[a]
	Performance	122	8	.49	.41	.56	.54	.00	.54	.54	100.00[a]
Ad Hoc	Cognitive	554	12	.38	.30	.46	.42	.19	.18	.67	34.87
	Affective	211	7	.26	.18	.34	.28	.00	.28	.28	100.00[a]
	Process	532	17	.39	.34	.44	.44	.04	.39	.49	95.05
	Performance	902	32	.31	.26	.36	.38	.10	.25	.50	83.66

Note. N = number of teams; k = number of correlations coefficients on which each distribution was based; $\bar{r}$ = mean observed correlation; CI_r 10% = lower bound of the confidence interval for observed r; CI_r 90% = upper bound of the confidence interval for observed r; ρ = estimated true correlation between the predictor construct and the relevant criterion (fully corrected for measurement error in both the predictor and criterion); $SD\rho$ = estimated standard deviation of the true correlation; 10% CV = lower bound of the credibility interval for each distribution; 90% CV = upper bound of the credibility interval for each distribution; % Var. Accounted for by Artifacts = percentage of observed variance accounted for by statistical artifacts.

[a]These entries represent cases in which the percentage of variance accounted for by artifacts resulted in a number greater than 100. This can be interpreted such that all the variance between effect sizes from primary studies could be accounted for by study artifact. We direct the reader to the Meta-Analytic Results section for a more detailed explanation.

TABLE 5: Analysis of the Effectiveness of Team Training Based Upon Team Size

Team Size	Outcome Type	N	k	$\bar{r}$	CI_r 10%	CI_r 90%	ρ	SD_ρ	10% CV	90% CV	% Var. Accounted for by Artifacts
Small	Cognitive	77	2	.16	.03	.30	.16	.00	.16	.16	100.00[a]
	Affective	249	6	.39	.34	.43	.39	.00	.39	.39	100.00[a]
	Process	253	10	.48	.42	.54	.59	.00	.59	.59	100.00[a]
	Performance	418	12	.28	.20	.36	.39	.18	.15	.62	61.63
Medium	Cognitive	477	10	.42	.34	.50	.46	.18	.23	.69	35.29
	Affective	183	7	.27	.18	.36	.28	.00	.28	.28	100.00[a]
	Process	288	7	.31	.24	.37	.33	.00	.33	.33	100.00[a]
	Performance	461	15	.34	.28	.40	.34	.09	.23	.46	77.96
Large	Cognitive	—	—	—	—	—	—	—	—	—	—
	Affective	33	3	.07	–.09	.22	.08	.00	.08	.08	100.00[a]
	Process	66	8	.43	.29	.56	.49	.00	.49	.49	100.00[a]
	Performance	145	13	.44	.37	.51	.50	.00	.50	.50	100.00[a]

Note. N = number of teams; k = number of correlations coefficients on which each distribution was based; $\bar{r}$ = mean observed correlation; CI_r 10% = lower bound of the confidence interval for observed r; CI_r 90% = upper bound of the confidence interval for observed r; ρ = estimated true correlation between the predictor construct and the relevant criterion (fully corrected for measurement error in both the predictor and criterion); SD_ρ = estimated standard deviation of the true correlation; 10% CV = lower bound of the credibility interval for each distribution; 90% CV = upper bound of the credibility interval for each distribution; % Var. Accounted for by Artifacts = percentage of observed variance accounted for by statistical artifacts.

[a]These entries represent cases in which the percentage of variance accounted for by artifacts resulted in a number greater than 100. This can be interpreted such that all the variance between effect sizes from primary studies could be accounted for by study artifact. We direct the reader to the Meta-Analytic Results section for a more detailed explanation.

the distribution of effect sizes after other research artifacts have been removed. That is, credibility intervals use the estimated true correlation between constructs as their basis. Moreover, credibility intervals, along with the estimate for the percentage of variance attributable to statistical artifacts, can provide information on whether moderators may be operating (Whitener, 1990). As the percentage of variance attributable to artifacts increases, the more certain it is that additional moderators are not operating.

A brief discussion explaining the percentage of variance accounted for by artifacts is appropriate. The meta-analytic procedure that we used is based on the hypothesis that much of the variation in study results is often based on statistical and methodological artifacts, rather than true underlying population relationships. These artifacts may distort study findings in many ways. For example, sampling error distorts study findings randomly, whereas the effect of measurement error is to systematically bias study results. Thus, this meta-analytic procedure corrects for the artifactual differences in study results. The amount of variance accounted for by study artifacts is based on calculating the ratio of the variance attributable to artifacts to the total variance.

This index, as calculated by the Hunter-Schmidt Meta-Analysis Program 1.1 (Schmidt & Le, 2005), represents the amount of variance in the observed correlations (those gathered from the primary studies) attributed to artifacts. In other words, this number indicates the amount of variance that is accounted for by sampling error and unreliability. This index allows one to judge whether substantial variation attributable to moderators exists. A number greater than 100% is not uncommon and is an indication that the variation in effect sizes from the primary studies could have been attributable to second-order sampling error and not existing moderators (Hunter & Schmidt, 2004).

If the percentage of variance accounted for by artifacts is calculated to be 100 or greater, whether the number is 150 or 300, the correct conclusion would be that *all* the total variance could be accounted for by artifacts (Hunter & Schmidt, 2004). The magnitude of the number does not reveal anything about the quality of the analysis. Rather, what is meaningful is if the number is greater than 100% or less than 100%. A number under 100% indicates that some of the variance left to account for may be attributable to unhypothesized moderators, whereas a number over 100% indicates that the total variance could be accounted for by artifacts such as sampling error or measurement error.

There is no relationship between this number and the confidence in the estimated true correlation. Rather, this is another piece of information that is used in order to interpret the findings. Moreover, when the percentage of variance accounted for is greater than 100%, the credibility interval becomes a point estimate. The credibility interval is yet another indicator of the existence of moderators. When a credibility interval is a point estimate, it is unlikely that moderators exist based on the analysis.

In the following sections, the results derived from our investigations into the four primary areas of interest in this research are presented. In meta-analysis there is no criterion that illustrates the minimum number of effect sizes to include in an analysis. We do acknowledge and understand that conducting an analysis with a small number of effect sizes does increase the possibility of second-order sampling error (Arthur, Bennett, & Huffcutt, 2001; Hunter & Schmidt, 2004). In order to be complete and to fully illustrate the database used in this meta-analysis, we conducted all possible analyses with one or more effect sizes. However, we do emphasize, as did Arthur, Bennett, Edens, and Bell (2003, see p. 238), that any meta-analyzed effect size based on fewer than five effect sizes should be interpreted with caution.

Does team training work? The overall results support our hypothesis that team training does work. For the combined (and averaged, when necessary) set of independent outcomes, team training was shown to have a moderate, positive effect on team functioning ($\rho = .34$; 10% CV = .34; 90% CV = .34). This analysis was based on 52 effect sizes representing 1,563 teams and provided support for Hypothesis 1.

Criterion description. The results of our analyses revealed team training to have a positive effect on each of the four outcomes under investigation (see Table 2). For the 12 available effect sizes for team-level cognitive outcomes, the estimated true score correlation (ρ) was .42. The 80% credibility interval (10% CV and 90% CV) ranged from .18 to .67. For affective outcomes, the analysis of 16 effect sizes (representing 465 teams) revealed an estimated true score correlation of .35. Next, team training appeared to work best for process outcomes ($\rho = .44$; 10% CV = .44; 90% CV = .44);

this analysis was based on 25 independent effect sizes representing 607 teams. However, given the overlapping confidence and credibility intervals, this assertion is merely suggestive and not definitive. Finally, team training was also shown to be effective for performance outcomes (ρ = .39; 10% CV = .27; 90% CV = .50). Overall, these findings provide support for Hypotheses 2a through 2d.

Training content. The content of training interventions was also investigated as a potential moderator variable (see Table 3). For this analysis, the content of study interventions was coded as primarily taskwork, teamwork, or both. Overall, the analyses provided support for Hypotheses 3 through 5 but not necessarily Hypothesis 6. That is, although each type of intervention was shown to be useful, there is insufficient evidence to conclude that mixed content interventions are superior to taskwork or teamwork interventions alone.

Moreover, there was very little difference among taskwork (ρ = .35), teamwork (ρ = .38), or mixed content (ρ = .40) interventions for the improvement of performance. However, for process outcomes, taskwork training (ρ = .28) did not work as well as teamwork (ρ = .44) or mixed content training (ρ = .56); caution should be exercised in the interpretation of the taskwork effect size, as it is based on three effect sizes. Similarly, teamwork-focused training appeared to result in enhanced affective outcomes in comparison with taskwork training (ρ = .41 and .11, respectively). We again caution the reader in the interpretation made on the estimated correlation for taskwork content on affective outcomes, as it was based on one point estimate.

Team membership stability. The results for the analyses on team membership stability provided partial support for Hypothesis 7 (see Table 4). Specifically, team training appeared to work just as well (if not better) for intact teams. This was especially the case for performance outcomes, for which ρ = .54 and .38 for intact and ad hoc teams, respectively. At the same time, the findings for process outcomes were highly similar across the two team membership configurations (ρ = .48 and .44 for intact and ad hoc teams, respectively).

Team size. For the analyses addressing the potential moderating effect of team size, we divided studies into three categories based on the average size of the teams under investigation: small, medium, and large. The most informative comparison here concerns performance outcomes. Our results suggest that team size impacts the benefit from team training in terms of improved performance (ρ = .39, .34, and .50, respectively, for small, medium, and large teams; see Table 5). These results are based upon having 12, 15, and 13 effect sizes in each group, respectively. The findings from the other three outcomes with team size as a potential moderator are not as clear. For example, team training appeared to influence cognitive outcomes to a larger extent in teams of medium size (ρ = .46). Conversely, small teams benefited the most in terms of improvements in affective (ρ = .39) and process (ρ = .59) outcomes. This diverse set of findings does not provide any appreciable level of support for Hypothesis 8.

Post hoc analyses. We conducted some post-hoc meta-analyses on effect sizes based on domain and strategy (see Tables 6 and 7). The five primary domains that are represented in our sample are studies conducted in the aviation, military, medical, and business sectors and those conducted in laboratories with student samples. The training strategies that were represented in our database consisted of coordination training/crew resource management, cross-training, communication training, critical thinking training, self-guided training, self-correction training, team adaptation and coordination training (TACT), stress training, tactical training, and team knowledge training. These results should be interpreted with caution but nonetheless can be used to paint a picture of not only team training in each of these domains but also where the research and quantifiable data on training effectiveness are lacking.

Most of the primary studies could be categorized into two domains: studies conducted in the military and those conducted in the lab with student samples. These analyses were based on 39 and 37 effect sizes, representing 762 teams in the military domain and 1,660 student teams. Team training appeared more effective in military settings (ρ = .59; 10% CV = .59; 90% CV = .59) than in laboratory settings (ρ = .33; 10% CV = .22; 90% CV = .43).

The other three domains represented by the primary studies in our analyses were from the medical, aviation, and business settings. These three domains did not render a significant number of primary studies that reported the necessary quantifiable data to use in meta-analysis. Six independent effect sizes were utilized to conduct the analysis for medical and aviation studies, and five independent

text continues on page 926

TABLE 6: Post Hoc Analysis of the Effectiveness of Team Training Based Upon Domain

Domain	N	k	$\bar{r}$	CI_r 10%	CI_r 90%	ρ	SD_ρ	10% CV	90% CV	% Var. Accounted for by Artifacts
Lab/university	1,660	37	.30	.27	.33	.33	.08	.22	.43	77.42
Military	762	39	.49	.45	.53	.59	.00	.59	.59	100.00[a]
Business	10	5	.72	.68	.77	.85	.00	.85	.85	100.00[a]
Aviation	138	6	.29	.15	.43	.33	.20	.08	.58	56.83
Medical	80	6	.21	.06	.37	.23	.10	.11	.35	91.17

Note. N = number of teams; k = number of correlations coefficients on which each distribution was based; $\bar{r}$ = mean observed correlation; CI_r 10% = lower bound of the confidence interval for observed r; CI_r 90% = upper bound of the confidence interval for observed r; ρ = estimated true correlation between the predictor construct and the relevant criterion (fully corrected for measurement error in both the predictor and criterion); SD_ρ = estimated standard deviation of the true correlation; 10% CV = lower bound of the credibility interval for each distribution; 90% CV = upper bound of the credibility interval for each distribution; % Var. Accounted for by Artifacts = percentage of observed variance accounted for by statistical artifacts.

[a]These entries represent cases in which the percentage of variance accounted for by artifacts resulted in a number greater than 100. This can be interpreted such that all the variance between effect sizes from primary studies could be accounted for by study artifact. We direct the reader to the Meta-Analytic Results section for a more detailed explanation.

TABLE 7: Post Hoc Analysis of the Effectiveness of Team Training Based Upon Strategy

Strategy	N	k	$\bar{r}$	CI_r 10%	CI_r 90%	ρ	SD_ρ	10% CV	90% CV	% Var. Accounted for by Artifacts
Coordination/CRM training	744	33	.43	.39	.46	.47	.00	.47	.47	100.00[a]
Cross-training	432	14	.40	.33	.47	.44	.14	.26	.62	59.05
Communication training	86	3	.13	–.02	.28	.13	.08	.02	.23	83.91
Critical thinking training	83	4	.34	.14	.53	.60	.40	.09	1.12	43.58
Self-guided training	82	2	.35	.30	.39	.36	.00	.36	.36	100.00[a]
Self-correction training	79	2	.24	.05	.43	.27	.16	.06	.48	51.46
TACT	67	6	.50	.43	.57	.56	.00	.56	.56	100.00[a]
Stress training	58	2	.05	–.01	.10	.05	.00	.05	.05	100.00[a]
Tactical training	20	2	.67	.61	.73	.67	.00	.67	.67	100.00[a]
Team knowledge training	15	2	.81	.76	.87	.81	.00	.81	.81	100.00[a]

Note. N = number of teams; k = number of correlations coefficients on which each distribution was based; $\bar{r}$ = mean observed correlation; CI_r 10% = lower bound of the confidence interval for observed r; CI_r 90% = upper bound of the confidence interval for observed r; ρ = estimated true correlation between the predictor construct and the relevant criterion (fully corrected for measurement error in both the predictor and criterion); SD_ρ = estimated standard deviation of the true correlation; 10% CV = lower bound of the credibility interval for each distribution; 90% CV = upper bound of the credibility interval for each distribution; % Var. Accounted for by Artifacts = percentage of observed variance accounted for by statistical artifacts.

[a]These entries represent cases in which the percentage of variance accounted for by artifacts resulted in a number greater than 100. This can be interpreted such that all the variance between effect sizes from primary studies could be accounted for by study artifact. We direct the reader to the Meta-Analytic Results section for a more detailed explanation.

effect sizes were found from studies conducted in business settings. The results indicate that team training conducted with business groups ($\rho = .85$) resulted in greater effect sizes than did studies conducted with medical teams ($\rho = .23$; 10% CV = .11; 90% CV = .35) and aviation teams ($\rho = .33$; 10% CV = .08; 90% CV = .58). Again, we caution the reader in interpreting these analyses.

Our last post hoc analysis was based on the training strategy used. The two most popular strategies represented in our database were coordination training and cross-training. Thirty-three effect sizes were meta-analyzed in order to assess the impact that coordination training had on team outcomes. The results indicated that this type of training had a moderate positive effect on outcomes ($\rho = .47$; 10% CV = .47; 90% CV = .47). Cross-training was represented in our database by 14 effect sizes and 432 teams. These results suggest that cross-training does have a moderate positive effect on outcomes ($\rho = .44$; 10% CV = .26; 90% CV = .62). We direct the reader to Table 7, which includes all the analyses based upon strategy. And again, we strongly caution the interpretation of these analyses and can state only that more research is needed based on strategy.

DISCUSSION

Meta-analysis is a valuable tool for navigating the research literature when conflicting findings are the norm rather than the exception. Most areas of research fall prey to this issue, and the team training literature is certainly not immune. A meta-analysis of this area is beneficial so that the findings across studies can be integrated and reveal some simpler patterns of relationships. These preliminary patterns can then provide a basis for theory development. In addition to this benefit, meta-analysis can be particularly advantageous for research on teams because of the small-sample studies that are typical of this area. The current investigation therefore provides an overdue empirical accounting of the team training literature and, in turn, identifies important patterns of results.

What Works and Why?

The findings from the present study provide insight into the extent to which team training interventions relate to team outcomes. That is, across a wide variety of settings, tasks, and team types, team training efforts were successful. Although "a positive relationship between training and performance in team settings is and has been clear for some time" (Hollenbeck et al., 2004, p. 360), this research is, to the best of our knowledge, the first to comprehensively and quantitatively evaluate the relationships between these interventions and outcomes.

These findings suggest that team training interventions are a viable approach for organizations to take in order to enhance team outcomes. They are useful for improving cognitive outcomes, affective outcomes, teamwork processes, and performance outcomes. That is, team training accounted for approximately 12% to 19% of the variance in the examined outcomes. This percentage, though statistically correct, is an underestimate of the practical and theoretical significance of the relationships between the variables investigated in the present meta-analysis (e.g., Hunter & Schmidt, 1990). Thus, the utility of team development techniques is likely understated by the present findings. Further, the large number of studies in the database and the large overall observed effect size make these findings robust to disconfirmation, excluding those analyses that included fewer than five effect sizes.

Across all outcomes, team training interventions were more effective for team processes than for the other outcome types. As a primary explanation, in dynamic input-throughput-output models it is generally understood that team-member and team-level processes such as communication and coordination are determinants of team performance and effectiveness. Although the training interventions reviewed in the present study did appear to have a nearly equal impact on outcomes as processes, it stands to reason that improving processes will also positively influence performance outcomes.

For example, training teams to communicate, coordinate, and make decisions should ultimately increase the quantity of products produced. Finally, it is also recognized that team performance is shaped by additional environmental and/or organizational characteristics and contingencies that are out of the control of team members; team members actively execute team processes but are often at the mercy of a larger organizational system to produce more distal performance outputs. Of course, team processes may also be influenced by the larger organizational system, but probably not to the same degree as team performance outcomes.

In contrast to our hypothesis, team training interventions with a mixed training content (i.e., teamwork and taskwork) were not superior to those with a taskwork or teamwork focus in isolation. Team training was deemed useful for improving team performance outcomes, regardless of the type of training content. This finding could be interpreted such that team training is helpful but that combining teamwork and taskwork content does not necessarily make the training incrementally more beneficial. The problem could potentially be that too much information was included in the training given to teams. As such, the teams exposed to mixed-content training may have been overwhelmed and could not successfully grasp the intended concepts and information that were presented to them.

An additional explanation is available, but it merits more research. Specifically, Salas, Burke, and Cannon-Bowers (2002) identified the importance of both taskwork and teamwork training. They argued that it is best to develop a team's taskwork knowledge prior to the team's acquisition of teamwork skills. Alternatively, the results from a recent study by Ellis et al. (2005) suggested that enhancing teamwork skills prior to taskwork knowledge leads to greater performance improvements. Although these explanations do not combine to provide a clear explanation for the lack of support for Hypothesis 6, they do serve as a call to researchers to further examine the impact of temporal factors when designing training that contains both teamwork and taskwork components.

Our research also showed that membership stability moderated the relationship between team training and team outcomes, such that intact teams that underwent training improved the most on process and performance outcomes. Our hypothesis that intact teams would benefit most from team training interventions was only partially supported, however. There was clear support, given the nonoverlapping confidence intervals, for our assertion that intact teams would benefit the most from training interventions when the focus was on performance outcomes. This finding was consistent with our belief that those teams that had an opportunity to work with one another would benefit the most from team training.

Intact teams should have already overcome some of the maturational challenges that newly formed teams would not have had the opportunity to navigate. Ad hoc teams are likely forced to attend to many more factors during their training session than intact teams, and thus intact teams should benefit the most from team training. An additional explanation could be that intact teams may have seen the benefits of training more so than ad hoc teams and, thus, were more likely to attend to the important aspects of training, whether taskwork or teamwork focused in nature. In other words, intact teams may have perceived the training as being more instrumental to their team's performance improvement and, ultimately, to the attainment of valued outcomes in the workplace.

The moderator analysis on team size helped to further clarify when team training is most helpful for collectives. Team performance, as expected, improved the most for large teams. Moreover, team processes improved the most for small teams. These findings partially supported our hypotheses. Intuitively, if the members of a small team are behaving in a way that impedes their performance, once they are trained how to properly communicate and coordinate with one another, they may have an easier time maintaining those process enhancements.

A surprising finding was the extremely low effect size ($\rho = .08$) for affective outcomes in large teams. However, it is important to interpret this finding with caution because there were only three effect sizes to combine for this calculation. Additionally, it speaks to the need for more research on these phenomena in order to appropriately determine the effect of team training on affective outcomes for large teams. Moreover, as there were no effect sizes available to meta-analyze for cognitive outcomes for large teams, more research is necessary in this area as well.

Implications for Practitioners

We believe that much progress has been made with respect to team training, both in research and in practice. The meta-analytic findings are encouraging in that they show team training to be effective. However, these findings highlight some practical issues for organizational change agents to consider when designing and implementing training. First, practitioners should relish the fact that team training is effective. The implications of this are critical. To know that team training can explain 12% to 19% of the variance of a team's performance (i.e., better-performing teams will exhibit more effective team behaviors than their counterparts) can mean reducing medical errors (in health care), saving an aircraft (in aviation), increasing the bottom line (in business), or saving lives (in the military).

Second, given the heightened interest in team training in health care, change agents in health care institutions should utilize this information to bolster their argument for implementing such training.

Third, in deciding on the specific type and focus of team training, practitioners need to be cognizant of the ultimate outcome they are interested in, as our findings indicated differential effects. Sometimes these effects were rather small, but at other times they were rather large, depending on the specific outcome targeted (i.e., cognition, behavior/process, affect, performance).

In conclusion, the data we present, as we have stated several times, are encouraging. Of most importance, however, is what individuals do to improve the training and development initiatives that they manage. We hope that practitioners and researchers alike find useful and practical advice to follow based on the findings presented.

Study Limitations

In general, a large number of methodological tools are available to researchers who attempt to discover the meaning of organizational phenomena (McCall & Bobko, 1990); meta-analysis is just one such tool. The validity of inferences concerning any relationship examined through meta-analysis is a function of several factors, including (a) the representativeness of the sample of primary studies considered by the meta-analyst, (b) the degree of validity of each of the primary studies, and (c) the number of primary studies (Bobko & Stone-Romero, 1998).

Despite the strengths and contributions of the current study, several potential limitations should be noted. First, although extensive efforts were made to be inclusive and a large number of primary studies were included, definitive conclusions regarding the validity of each included study cannot be made. To the extent that the primary studies lacked proper control mechanisms, either via design features or statistically, it is not possible to unequivocally conclude that team training interventions cause improvements in team functioning.

Second, as Rosenthal and DiMatteo (2001) suggested, every meta-analysis has some inherent bias attributable to the selection of inclusion/exclusion criteria and the methods chosen to review the literature. To the extent that the studies in the database come from populations that are more homogenous and nonrepresentative than is typical, the more likely it is that the external validity of these findings may be threatened. For example, the overwhelming majority of the studies included herein were conducted using college students acting as action and performing teams in simulated environments (e.g., aviation, military). Clearly, this focus has been influenced by years of generous funding by the Department of Defense and various military institutions (e.g., U.S. Army, U.S. Navy). Although these studies have obvious merit in advancing the science of this area, they often do not capture the unique dynamics and context in which other teams perform.

Finally, in our attempt to drill down and gauge relationships at more finite levels, we unfortunately left many of our subgroups lacking a robust number of effect sizes to analyze. This was both a necessary evil and a cogent assessment of the current state of research on teams. It was necessary because we deemed it important to provide more prescriptive findings than simply saying, "team training works to improve team outcomes." The goal of addressing when and where it works best is an important undertaking. This research takes an initial step in that direction, but it is worth repeating that more and better evaluations of the impact of team training in organizations are needed.

In addition, we ask researchers and practitioners, in any domain, to consider the data that they present when reporting their findings. Whereas *p* values and percentage improvement provide an indication of the impact of team training, they leave the reader with no true measure of the magnitude with which team training improved desired outcomes. We call for researchers (and editors of journals) to report (and to demand) informative statistics, such as the group means and standard deviations, rather than solely *p* values or percentages. Researchers need to examine diverse samples and settings and to elicit more published data, which would allow for additional meta-analyses in the future.

Recommendations for Future Research

The results of the present meta-analysis suggest several avenues for future research. First, most of the articles contained in the database focused on teams working jointly on a task in the same location. However, many organizations have operations in multiple countries. In multinational organizations, globally distributed teams are often responsible for making and implementing important decisions. Research is needed that examines whether team training improves team outcomes in distributed teams. Distributed teams have to resolve the same problems that other teams encounter; however, these challenges are often more pronounced when team members are distributed across temporal and geographic boundaries (Stagl, Salas, Rosen, et al., 2007). This type of research would provide useful information to managers and human resource professionals regarding the increasing use of globally distributed teams.

More than a decade ago, Stout, Salas, and Fowlkes (1997) noted, "there are few systematic evaluations of the effectiveness of particular team training efforts" (p. 169). Unfortunately, the large amount of published literature that had to be excluded from the meta-analytic database bears evidence that this critique still merits attention. That is, although the study of work teams has increased in the past 2 decades (Cohen & Bailey, 1997; Nielsen, Sundstrom, & Halfhill, 2005; Sundstrom, McIntyre, Halfhill, & Richards, 2000), relatively few researchers properly report their findings. In short, there is a plethora of team training but little systematic evaluation and dissemination of results. Practitioners who facilitate team training interventions and researchers who study teams and team training should make increased efforts to collaborate and empirically evaluate and publish their findings in peer-reviewed scientific outlets.

Another area meriting further attention involves the extent to which the team leader is incorporated as a critical component of team training design. Research has demonstrated the beneficial influence of team leader support, by way of leader briefings, for posttraining recall and the performance of trainees (e.g., Mathieu et al., 2000; Smith-Jentsch, Salas, & Brannick, 2001). Following this further, a recent meta-analysis found that person-focused and task-focused behaviors by the leader were related to team effectiveness and team productivity (Burke, Stagl, Klein, et al., 2006). However, it was an unfortunate fact that very few studies in the database included any discussion of the role of the leader in team training. This is in contrast to the vast literature on team leadership in general. Given the identification of a sufficient number of empirical articles on the topic, a moderator analysis investigating the presence or absence of team leader support during training may be an informative avenue of future research.

Also highlighted within the current findings was the potentially important role of time in intervention delivery. This area has been recently highlighted as one that needs more attention in the team's arena by several researchers. For example, with regard to developmental interventions, Hackman and Wageman (2005) noted the importance of timing. Specifically, they discussed how team coaching interventions should be made: "at times when the team is ready for them and able to deal with them" (p. 283). They suggested that coaching be done at the beginning of the team's task cycle for effort-related interventions, near the midpoint for strategy-related interventions, and at the end of a task cycle for interventions that address knowledge and skill (Hackman & Wageman, 2005). We join with those authors and others in arguing for research that examines issues related to time and developmental interventions – in our specific case, team training interventions.

Last, our meta-analysis did not consider the impact of team composition on the effectiveness of team training. Most of the primary studies in the database did not consider team composition. Thus, considering the dearth of reported data on team composition variables in team training studies, we make a call to researchers. Specifically, although much research has been done to evaluate the moderating effects of individual characteristics (e.g., goal orientation, motivation to learn, conscientiousness) on individual learning, how does a team's composition impact team learning?

Concluding Remarks

Team training works! The findings from this research are encouraging. In general, our results also suggest that training content, team membership stability, and team size moderate the effectiveness of team training interventions. This study was the first to thoroughly and quantitatively summarize the team training literature. We hope our findings provide valuable information to those stakeholders charged with making informed decisions regarding the planning, design, development, and delivery of team training interventions. We also hope it inspires more and richer team training evaluations from a variety of sectors.

ACKNOWLEDGMENTS

This work was supported by funding from the U.S. Army Research Institute for the Behavioral and Social Sciences (Contract W74V8H-04-C-0025). The views, opinions, and/or findings contained in this paper are those of the authors and should not be construed as an official position, policy, or decision of the Department of the Army or the University of Central Florida.

We would like to thank Huy Le, John Mathieu, and Florian Jentsch for their insightful guidance during the course of this project. We would also like to acknowledge Shannon Scielzo and Samuel Wooten for their assistance during the early coding portions of the project.

REFERENCES

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

*Adelman, L., Christian, M., Gualtieri, J., & Bresnick, T. A. (1998). Examining the effects of communication training and team composition on the decision making of patriot air defense teams. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 28,* 729–741.

*Alexander, L. T., Kepner, C. H., & Tregoe, B. B. (1962). The effectiveness of knowledge of results in a military system-training program. *Journal of Applied Psychology, 46,* 202–211.

Arthur, W., Jr., Bennett, W., Jr., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology, 88,* 234–245.

Arthur, W., Jr., Bennett, W., Jr., & Huffcutt, A. I. (2001). *Conducting meta-analysis using SAS.* Mahwah, NJ: Erlbaum.

Baker, D. P., Salas, E., Barach, P., Battles, J., & King, H. (2007). The relation between teamwork and patient safety. In P. Carayon (Ed.), *Handbook of human factors and ergonomics in health care and patient safety* (pp. 259–271). Mahwah, NJ: Erlbaum.

Banker, R. D., Field, J. M., Schroeder, R. G., & Sinha, K. K. (1996). Impact of work teams on manufacturing performance: A longitudinal field study. *Academy of Management Journal, 39,* 867–890.

*Banks, J. H., Jr., Hardy, G. D., Scott, T. D., Kress, G., & Word, L. E. (1977). *Realtrain validation for rifle squads: Mission accomplishment* (Research Rep. 1192, Army Project No. 2Q763743A773 and 2Q763743A775, Tactical Team Performance). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

*Blickensderfer, E., Cannon-Bowers, J. A., & Salas, E. (1997, April). *Training teams to self-correct: An empirical investigation*. Paper presented at the 12th Annual Meeting of the Society for Industrial and Organizational Psychology, St. Louis, MO.

Bobko, P., & Stone-Romero, E. L. (1998). Meta-analysis may be another useful tool, but it is not a panacea. *Research in Personnel and Human Resources Management, 16,* 359–397.

Boehm-Davis, D. A., Holt, R. W., & Seamster, T. L. (2001). Airline resource management programs. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations* (pp. 191–215). Mahwah, NJ: Erlbaum.

*Brannick, M. T., Prince, C., & Salas, E. (2005). Can PC-based systems enhance teamwork in the cockpit? *International Journal of Aviation Psychology, 15,* 173–187.

*Brown, T. C. (2003). The effect of verbal self-guidance training on collective efficacy and team performance. *Personnel Psychology, 56,* 935–964.

Burke, C. S., Stagl, K. C., Klein, C., Goodwin, G. F., Salas, E., & Halpin, S. M. (2006). What type of leadership behaviors are functional in teams? A meta-analysis. *Leadership Quarterly, 17,* 288–307.

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

*Buzaglo, G., & Wheelan, S. A. (1999). Facilitating work team effectiveness: Case studies from Central America. *Small Group Research, 30,* 108–129.

Campbell, J. P., & Kuncel, N. R. (2001). Individual and team training. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of work and organizational psychology* (pp. 278–312). London: Blackwell.

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

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team competencies and establishing team training requirements. In R. Guzzo, E. Salas, & associates (Eds.), *Team effectiveness and decision making in organizations* (pp. 333–380). San Francisco: Jossey-Bass.

*Cobb, M. G. (2000). The impact of environmental complexity and team training on team processes and performance in multiteam environments. *Dissertation Abstracts International, 60,* 4281–4411.

Cohen, S. G., & Bailey, D. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite? *Journal of Management, 23,* 239–290.

Cohen, S. G., & Ledford, G. E. (1994). The effectiveness of self-managing teams: A quasi-experiment. *Human Relations, 47,* 13–43.

*Cooke, N. J., Cannon-Bowers, J. A., Kiekel, P. A., Rivera, K., Stout, R. J., & Salas, E. (2000). Improving teams' interpersonal knowledge through cross-training. In *Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 2.390–2.393). Santa Monica, CA: Human Factors and Ergonomics Society.

*Cooke, N. J., Kiekel, P. A., & Helm, E. E. (2001). Measuring team knowledge during skill acquisition of a complex task. *International Journal of Cognitive Ergonomics, 5,* 297–315.

*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.

Curral, L. A., Forrester, R. H., Dawson, J. F., & West, M. A. (2001). It's what you do and the way that you do it: Team task, team size, and innovation-related group processes. *European Journal of Work and Organizational Psychology, 10,* 187–204.

Davies, J. M. (2001). Medical applications of crew resource management. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations* (pp. 265–281). Mahwah, NJ: Erlbaum.

De Dreu, C. K. W. (2007). Cooperative outcome interdependence, task reflexivity, and team effectiveness: A motivated information processing perspective. *Journal of Applied Psychology, 92,* 628–638.

Denson, R. W. (1981). *Team training: Literature review and annotated bibliography* (Final Rep. AFHRL-TR-80-40). Brooks Air Force Base, TX: HQ Air Force Human Resources Laboratory.

Devine, D. J., Clayton, L. D., Philips, J. L., Dunford, B. B., & Melner, S. B. (1999). Teams in organizations: Prevalence, characteristics, and effectiveness. *Small Group Research, 30,* 678–711.

Driskell, J. E., Salas, E., & Johnston, J. (1995, April). *Is stress training generalizable to novel settings?* Paper presented at the 10th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

*Duncan, P. C., Rouse, W. B., Johnston, J. H., Cannon-Bowers, J. A., Salas, E., & Burns, J. J. (1996). Training teams working in complex systems: A mental model–based approach. In W. B. Rouse (Ed.), *Human/technology interaction in complex systems* (Vol. 8, pp. 173–231). Greenwich, CT: JAI Press.

Dyer, J. L. (1984). Team research and team training: A state-of-the-art review. In F. A. Muckler (Ed.), *Human factors review* (pp. 285–323). Santa Monica, CA: Human Factors and Ergonomics Society.

*Ellis, A. P. J., Bell, B. S., Ployhart, R. E., Hollenbeck, J. R., & Ilgen, D. R. (2005). An evaluation of generic teamwork skills training with action teams: Effects on cognitive and skill-based outcomes. *Personnel Psychology, 58,* 641–672.

*Entin, E. B., Entin, E. E., MacMillan, J., & Serfaty, D. (1993). *Structuring and training high-reliability teams* (AD No. ADA302385). Burlington, MA: Alphatech.

*Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors, 41,* 312–325.
*Entin, E. E., Serfaty, D., & Deckert, J. C. (1994). *Team adaptation and coordination training* (TR-648-1). Burlington, MA: Alphatech.
Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods, 6,* 161–180.
Field, A. P. (2005). Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological Methods, 10,* 444–467.
Flin, R. H. (1995). Crew resource management for teams in the offshore oil industry. *Journal of European Industrial Training, 19,* 23–27.
Flin, R. H., & O'Connor, P. (2001). Applying crew resource management on offshore oil platforms. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations* (pp. 217–233). Mahwah, NJ: Erlbaum.
Foushee, H. C. (1984). Dyads and triads at 35,000 feet: Factors affecting group process and aircrew performance. *American Psychologist, 39,* 885–893.
*Freeman, J. T., Cohen, M. S., & Thompson, B. (1998). Effects of decision support technology and training on tactical decision making. In *Proceedings of the 1998 Command and Control Research & Technology Symposium* [CD-ROM]. McLean, VA: Command and Control Research Program.
Gagné R. M. (1962). Military training and principles of learning. *American Psychologist, 17,* 83–91.
*Ganster, D. C., Williams, S., & Poppler, P. (1991). Does training in problem solving improve the quality of group decisions? *Journal of Applied Psychology, 76,* 479–483.
George, G. C., Hannibal, M. J., & Hirsch, W. L. (2004). Building the brand through people. *WorldatWork Journal, 13*(1), 39–45.
Glass, B. (1955). News and notes. *Science, 21,* 583–596.
Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5,* 3–8.
Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.
*Glickman, A. S., Zimmer, S., Montero, R. C., Guerette, P. J., Campbell, W. J., Morgan, B. B., et al. (1987). *The evolution of teamwork skills: An empirical assessment with implications for training* (Tech. Rep. Contract No. N00014-86-K-0732). Ft. Belvoir, VA: Defense Technical Information Center.
Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations: Needs assessment, development, and evaluation* (4th ed.). Belmont, CA: Wadsworth.
Gooding, R. Z., & Wagner, J. A., III. (1985). A meta-analytic review of the relationship between size and performance: The productivity and efficiency of organizations and their subunits. *Administrative Science Quarterly, 30,* 462–481.
*Green, L. R. (1994). *The effectiveness of tactical adaptation and coordination training on team performance in tactical scenarios.* Unpublished master's thesis, Naval Postgraduate School, Monterey, California.
Gully, S. M., Devine, D. J., & Whitney, D. J. (1995). A meta-analysis of cohesion and performance: Effects of levels of analysis and task interdependence. *Small Group Research, 26,* 497–520.
Guzzo, R. A., & Shea, G. P. (1992). Group performance and intergroup relations in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, pp. 269–313). Palo Alto, CA: Consulting Psychologists Press.
Hackman, J. R. (1983). *A normative model of work team effectiveness* (Tech. Rep. No. 2). New Haven, CT: Yale University.
Hackman, J. R. (1987). The design of work teams. In J. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 215–342). New York: Prentice Hall.
Hackman, J. R. (Ed.). (1990). *Groups that work (and those that don't): Creating conditions for effective teamwork.* San Francisco: Jossey-Bass.
Hackman, J. R. (2002, August). New rules for team building. *Optimize,* pp. 50–62.
Hackman, J. R., & Vidmar, N. (1970). Effects of size and task type on group performance and member reactions. *Sociometry, 33,* 37–54.
Hackman, J. R., & Wageman, R. (2005). A theory of team coaching. *Academy of Management Review, 30,* 269–287.
*Heffner, T. S. (1997). The influence of team training methods on simulator performance. *IEEE Transactions on Systems, Man, and Cybernetics, 27,* 4165–4170.
Helmreich, R. L., & Foushee, H. C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E. L. Weiner, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. 3–45). San Diego, CA: Academic Press.
*Henry, L. J. (1997). *The effect of group process training on team effectiveness.* Unpublished doctoral dissertation, Texas Women's University, College of Health Sciences, Denton, TX.
Hollenbeck, J. R., DeRue, D. S., & Guzzo, R. (2004). Bridging the gap between I/O research and HR practice: Improving team composition, team training, and team task design. *Human Resource Management, 43,* 353–366.
Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Beverly Hills, CA: Sage.
Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Newbury Park: Sage.
Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.
Hunter, J. E., Schmidt, F. L., & Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis method and findings. *Journal of Applied Psychology, 91,* 594–612.
*Ikomi, P. A., Boehm-Davis, D. A., Holt, R. W., & Incalcaterra, K. A. (1999). Jump seat observations of advance crew resource management (ACRM) effectiveness. In R. S. Jensen, B. Cox, J. D. Callister, & R. Lavis (Eds.), *Proceedings of the 10th International Symposium on Aviation Psychology* (pp. 292–297). Columbus: Ohio State University.
Ilgen, D. R. (1999). Teams embedded in organizations: Some implications. *American Psychologist, 54,* 129–139.
Indik, B. P. (1965). Operational size and member participation: Some empirical tests of alternative explanations. *Human Relations, 18,* 339–350.
Judge, T. A., Cable, D. M., Colbert, A. E., & Rynes, S. L. (2007). What causes a management article to be cited—Article, author, or journal? *Academy of Management Journal, 50,* 491–506.
Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65,* 681–706.
Kidwell, R. E., & Bennett, N. (1993). Employee propensity to withhold effort: A conceptual model to intersect three avenues of research. *Academy of Management Review, 18,* 429–456.
Klein, C., DiazGranados, D., Salas, E., Le, H., Burke, C. S., & Lyons, R. (in press). The effect of team building interventions on team outcomes: An update and extension. *Small Group Research.*
Klein, C., Stagl, K. C., Salas, E., Parker, C., & Van Eynde, D. F. (2007). Returning to flight: Simulation-based training for the U.S. National Aeronautics and Space Administration's Space Shuttle Mission Management Team. *International Journal of Training and Development, 11,* 132–138.
Kozlowski, S. W. J., & Bell, B. S. (2003). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Comprehensive handbook of psychology: Industrial and organizational psychology* (Vol. 12, pp. 333–375). New York: Wiley.
Kozlowski, S. W. J., & Klein, J. K. (2000). A multi-level approach to theory and research in organizations: Contextual, temporal and emergent processes. In J. K. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). San Francisco: Jossey-Bass.
Kraiger, K. (1985). On learning from the past: A meta-analytic fable. *Personnel Psychology, 38,* 799–801.
Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78,* 311–328.
*Lampton, D. R., McDonald, D. P., Rodriguez, M. E., Cotton, J. E., Morris, C. S., Parsons, J., et al. (2001). *Instructional strategies for training teams in virtual environments* (Tech. Rep. No. 1110). Alexandria, VA: U.S. Army Research Institute for the Behavioral Sciences.
*Lassiter, D. L., Vaughn, J. S., Smaltz, V. E., Morgan, B. B., Jr., & Salas, E. (1990). A comparison of two types of training interventions on team communication performance. In *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 1372–1376). Santa Monica, CA: Human Factors and Ergonomics Society.

Latané, B. (1981). The psychology of social impact. *American Psychologist, 36,* 343–356.

*Le Blanc, P. M., Hox, J. J., Schaufeli, W. B., Taris, T. W., & Peeters, M. C. W. (2007). Take care! The evaluation of a team-based burnout intervention program for oncology care providers. *Journal of Applied Psychology, 92,* 213–227.

*Leedom, D. K., & Simon, R. (1995). Improving team coordination: A case for behavior-based training. *Military Psychology, 7,* 109–122.

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

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

*MacMillian, J., Entin, E. E., Entin, E. B., & Serfaty, D. (1994). *Structuring and training high reliability teams* (Year 2 Interim Tech. Rep. TR-632). Fort Rucker, AL: U.S. Army Research Institute/Aviation Research and Development Activity.

Magjuka, R. J., & Baldwin, T. T (1991). Team-based employee involvement programs: Effects of design and administration. *Personnel Psychology, 44,* 793–812.

Marks, M. A. (2006). The science of team effectiveness. *Psychological Science in the Public Interest, 7*(3), i.

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

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

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

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

Mathieu, J. E., & Leonard, R. L. (1987). Applying utility concepts to a training program in supervisory skills: A time-based approach. *Academy of Management Journal, 30,* 316–335.

McCall, M., & Bobko, P. (1990). Research methods and discovery in industrial/organizational psychology. In M. Dunnette & L. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 381–418). Palo Alto, CA: Consulting Psychologists Press.

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

McIntyre, R. M., & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 149–203). San Francisco: Jossey-Bass.

*Minionis, D. P. (1995). Enhancing team performance in adverse conditions: The role of shared team mental models and team training on an interdependent task (Doctoral dissertation, George Mason University, 1995). *Dissertation Abstracts International, 56,* 1139.

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

*Morey, J. C., Simon, R., Jay, G. D., Wears, R. L., Salisbury, M., Dukes, K. A., et al. (2002). Error reduction and performance improvement in the emergency department through formal teamwork training: Evaluation results of the MedTeams project (quality of care). *Health Services Research, 37,* 1553–1582.

Mullen, B. (1987). Self-attention theory. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 125–146). New York: Springer-Verlag.

Mullen, B. (1989). *Advanced BASIC meta-analysis.* Hillsdale, NJ: Erlbaum.

Mullen, B., Brown, R. J., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22,* 103–122.

Mullen, B., & Rosenthal, R. (1985). *BASIC meta-analysis.* Hillsdale, NJ: Erlbaum.

Mullen, B., Salas, E., & Driskell, J. E. (1989). Salience, motivation, and artifact as contributions to the relation between participation rate and leadership. *Journal of Experimental Social Psychology, 25,* 545–559.

Mullen, B., Symons, C., Hu, L., & Salas, E. (1989). Group size, leadership behavior, and subordinate satisfaction. *Journal of General Psychology, 116,* 155–170.

*Naylor, J. C., & Briggs, G. E. (1965). Team-training effectiveness under various conditions. *Journal of Applied Psychology, 49,* 223–229.

Nielsen, J. M., Sundstrom, E. D., & Halfhill, T. R. (2005). Group dynamics and effectiveness: Five years of applied research. In S. A. Wheelan (Ed.), *The handbook of group research and practice* (pp. 285–311). Thousand Oaks, CA: Sage.

Nieva, V. F., Fleishman, E. A., & Rieck, A. (1978). *Team dimensions: Their identity, their measurement and their relationships* (Contract No. DAHC 19-78-C-0001). Washington, DC: Advanced Research Resources Organization.

Noe, R. A. (2002). *Employee training and development.* Boston: McGraw-Hill.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Orasanu, J., & Fischer, U. (1997). Finding decisions in natural environments: The view from the cockpit. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making expertise: Research and applications* (pp. 343–357). Mahwah, NJ: Erlbaum.

Orasanu, J., & Salas, E. (1993). Team decision making in complex environment. In G. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 327–345). Norwood, NJ: Ablex.

Oser, R. L., Salas, E., Merket, D. C., & Bowers, C. A. (2001). Applying resource management training in naval aviation: A methodology and lessons learned. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations* (pp. 283–304). Mahwah, NJ: Erlbaum.

*Parsons, K. L. (1981). The effects of load sharing system training upon team performance (Doctoral dissertation, Michigan State University, 1980). *Dissertation Abstracts International, 42*(6-A), 2876–2877.

*Pearsall, M. J., Ellis, A. P. J., & West, B. J. (2004, August). *Reducing the negative effects of stress in team contexts: The impact of cross-training*. Paper presented at the Annual Conference of the Academy of Management, New Orleans, LA.

Pfeffer, J., & Sutton, R. I. (2000). *The knowing-doing gap: How smart companies turn knowledge into action.* Boston: Harvard Business School Press.

*Prichard, J. S., & Ashleigh, M. J. (2007). The effects of team-skills training on transactive memory and performance. *Small Group Research, 38,* 696–726.

*Rapp, T. L., & Mathieu, J. E. (2007). Evaluating an individually self-administered generic teamwork skills training program across time and levels. *Small Group Research, 38,* 532–555.

*Rollins, M. L., & Angelcyk, J. J. (1995). A descriptive study of crew resource management attitude change. In N. Johnston, R. Fuller, & N. McDonald (Eds.), *Aviation psychology: Training and selection: Proceedings of the 21st Conference of the European Association for Aviation Psychology* (Vol. 2, pp. 45–50). Aldershot, UK: Avebury Aviation.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52,* 59–82.

Saavedra, R., Earley, P. C., & Van Dyne, L. (1993). Complex interdependence in task-performing groups. *Journal of Applied Psychology, 78,* 61–72.

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors, 43,* 641–674.

Salas, E., Burke, C. S., & Cannon-Bowers, J. A. (2002). What we know about designing and delivering team training: Tips and guidelines. In K. Kraiger (Ed.), *Creating, implementing, and managing effective training and development: State-of-the-art lessons for practice* (pp. 234–259). San Francisco: Jossey-Bass.

Salas, E., & Cannon-Bowers, J. A. (1997). Methods, tools, and strategies for team training. In M. A. Quinones & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace* (pp. 249–279). Washington, DC: American Psychological Association.

Salas, E., & Cannon-Bowers, J. A. (2000). The anatomy of team training. In S. Tobias & J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military* (pp. 312–335). New York: Macmillan Reference.

Salas, E., Cannon-Bowers, J. A., & Johnston, J. H. (1997). How can you turn a team of experts into an expert team? Emerging training

strategies. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 359–370). Mahwah, NJ: Erlbaum.
Salas, E., DiazGranados, D., Weaver, S. J., & King, H. (2008). Does team training work? Principles for health care. *Academic Emergency Medicine, 15,* 1002–1009.
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.
Salas, E., Nichols, D. R., & Driskell, J. E. (2007). Testing three team training strategies in intact teams: A meta-analysis. *Small Group Research, 38*, 471–488.
Salas, E., Rozell, D., Mullen, B., & Driskell, J. E. (1999). The effect of team building on performance: An integration. *Small Group Research, 30,* 309–329.
Salas, E., Stagl, K. C., & Burke, C. S. (2004). 25 years of team effectiveness in organizations: Research themes and emerging needs. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 47–91). New York: Wiley.
Salas, E., Stagl, K. C., Burke, C. S., & Goodwin, G. F. (2007). Fostering team effectiveness in organizations: Toward an integrative theoretical framework of team performance. In R. A. Dienstbier (Series Ed.) & B. Shuart, W. Spaulding, & J. Poland (Vol. Eds.), *Modeling complex systems. Vol. 52: Current theory and research in motivation* (pp. 185–243). Lincoln: University of Nebraska Press.
Salas, E., Wilson, K. A., Burke, C. S., & Wightman, D. (2006). Does crew resource management training work? An update, an extension, and some critical needs. *Human Factors, 48,* 392–412.
Schmidt, F. L., & Le, H. (2005). *Hunter-Schmidt Meta-Analysis Programs* (Version 1.1) [Computer software]. Iowa City: University of Iowa, Department of Management & Organizations.
*Schreiber, M. A., Holcomb, J. B., Conaway, C. W., Campbell, K. D., Wall, M., & Mattox, K. L. (2002). Military trauma training performed in a civilian trauma center. *Journal of Surgical Research, 104,* 8–14.
*Serfaty, D., Entin, E. E., & Johnston, J. H. (1998). Team coordination training. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 221–245). Washington, DC: American Psychological Association.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.
*Shapiro, J. J., Morey, J. C., Small, S. D., Langford, V., Kaylor, C. J., Jagminas, L., et al. (2004). Simulation based teamwork training for emergency department staff: Does it improve clinical team performance when added to an existing didactic teamwork curriculum? *Quality and Safety in Health Care, 13,* 417–421.
Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.
*Siegel, A. I., & Federman, P. J. (1973). Communications content training as an ingredient in effective team performance. *Ergonomics, 16,* 403–416.
Smith-Jentsch, K. A., Salas, E., & Brannick, M. T. (2001). To transfer or not to transfer? Investigating the combined effects of trainee characteristics, team leader support, and team climate. *Journal of Applied Psychology, 86,* 279–292.
Stagl, K. C., Salas, E., & Fiore, S. M. (2007). Best practices in cross training teams. In D. A. Nembhard (Ed.), *Workforce cross training handbook* (pp. 156–175). Boca Raton, FL: CRC Press.
Stagl, K. C., Salas, E., Rosen, M. A., Priest, H. A., Burke, C. S., Goodwin, G. F., et al. (2007). Distributed team performance: A multilevel review of distribution, demography, and decision making. *Research in Multi-Level Issues, 6,* 11–58.
Steiner, I. D. (1972). *Group process and productivity.* Orlando, FL: Academic Press.
Stock, W. A. (1994). Systematic coding for research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 97–110). New York: Russell Sage Foundation.
*Stout, R. J., Salas, E., & Fowlkes, J. E. (1997). Enhancing teamwork in complex environments through team training. *Group Dynamics: Theory, Research, and Practice, 1,* 169–182.
Sundstrom, E., McIntyre, M., Halfhill, T., & Richards, H. (2000). Work groups: From the Hawthorne studies to work teams of the 1990s and beyond. *Group Dynamics, 4,* 44–67.
Tannenbaum, S. (2002). A strategic view of organizational training and learning. In K. Kraiger (Ed.), *Creating, implementing, and managing effective training and development* (pp. 10–52). San Francisco: Jossey-Bass.
Tannenbaum, S. I., Beard, R. L., & Salas, E. (1992). Team building and its influence on team effectiveness: An examination of conceptual and empirical developments. In K. Kelley (Ed.), *Issues, theory, and research in industrial/organizational psychology* (pp. 117–153). Amsterdam: Elsevier.
Tuckman, Bruce W. (1965). Developmental sequence in small groups. *Psychological Bulletin*, 63, 384–399.
*Volpe, C. E., Cannon-Bowers, J. A., Salas, E., & Spector, P. E. (1996). The impact of cross-training on team functioning: An empirical investigation. *Human Factors, 38,* 87–100.
Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology, 75,* 315–321.
Wiener, E. K., Kanki, B. G., & Helmreich, R. L. (Eds.). (1993). *Cockpit resource management* (pp. 479–501). San Diego, CA: Academic Press.
Yetton, P. W., & Bottger, P. C. (1982). Individual versus group problem solving: An empirical test of best-member strategy. *Organizational Behavior and Human Performance, 29,* 307–321.

Eduardo Salas is a Pegasus Professor and university trustee chair in the Department of Psychology and Institute for Simulation & Training at the University of Central Florida, Orlando. He received a Ph.D. in industrial/organizational psychology from Old Dominion University in 1984.

Deborah DiazGranados is a doctoral candidate and research associate in the Department of Psychology and Institute for Simulation & Training at the University of Central Florida, Orlando, where she received her M.S. in industrial/organizational psychology in 2005.

Cameron Klein is a doctoral candidate in the Department of Psychology at the University of Central Florida, Orlando, and a consultant at Kenexa in Lincoln, Nebraska. He received his M.S. in industrial/organization psychology from the University of Central Florida in 2004.

C. Shawn Burke is a research scientist in the Institute for Simulation & Training, University of Central Florida, Orlando. She received her Ph.D. in industrial/organizational psychology from George Mason University in 2000.

Kevin C. Stagl is a senior talent advisor at Talent Threshold LLC, Orlando, and an adjunct professor at the University of Central Florida, where he received his Ph.D. in industrial and organizational psychology in 2006.

Gerald F. Goodwin is a research psychologist in the Basic Research Unit at the Army Research Institute, Arlington, Virginia. He received his Ph.D. in industrial/organizational psychology from Pennsylvania State University in 1999.

Stanley M. Halpin is chief of the Fort Leavenworth Research Unit, the leader development unit of the Army Research Institute, Fort Leavenworth, Kansas. He received his Ph.D. in social psychology from Purdue University in 1970.

Date received: June 18, 2008
Date accepted: November 25, 2008