The Pennsylvania State University

The Graduate School

College of the Liberal Arts

THE EFFECTS OF STORYTELLING AND REFLEXIVITY ON TEAM MENTAL MODELS AND PERFORMANCE IN DISTRIBUTED DECISION-MAKING TEAMS

A Thesis in

Psychology

by

Rachel Tesler

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science

May 2011

The thesis of Rachel Tesler was reviewed and approved* by the following:

Susan Mohammed
Associate Professor of Psychology
Director of Graduate Training
Thesis Advisor

Richard A. Carlson
Professor of Psychology
Associate Head

Alicia Grandey
Associate Professor of Psychology

Melvin Mark
Professor of Psychology
Department Head

Michael McNeese
Professor of Information Sciences and Technology
Associate Dean for Research and Graduate Programs

*Signatures are on file in the Graduate School

Abstract

With the increasing number of virtual teams in the workforce, it is imperative to determine how to facilitate their performance (Cascio & Aguinis, 2008; Martins, Gilson, & Maynard, 2004). While team training and team mental models have been found to have positive effects on team performance, there are numerous unexplored team training tools and potential antecedents of team mental models that could further facilitate successful team outcomes. This study addressed these needs by investigating the effects of two team-level interventions, storytelling and guided team reflexivity, on team mental model similarity and performance outcomes. One hundred seven 3-person teams participated in NeoCITIES, a scaled-world simulation designed to mimic emergency crisis management situations in a distributed team environment. The presence of both storytelling and guided team reflexivity was manipulated. Results indicated that storytelling, when combined with reflexivity, had a positive effect on team mental model similarity and subsequent team performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I'd like to thank my advisor, Susan Mohammed, for her assistance in preparing me to reach this milestone. I'd also like to thank our entire research lab, who made it possible for this project to come to fruition. Lastly, I am deeply grateful to my husband Kevin and the rest of my family for their continued support of me in all of my academic endeavors.

Introduction

Structural and technological changes are becoming increasingly prevalent in the workplace. Two particularly noticeable trends are the flattening of organizational hierarchies and the increase of sophisticated technology (Bell & Kozlowski, 2002b; Devine, Clayton, Phillips, Dunford, & Melner, 1999; Lawler, Mohrman, & Ledford, 1992, 1995; Marks, Mathieu, & Zaccaro, 2001). Without a clear hierarchical power structure to direct employees in their responsibilities, teams are becoming a prevalent method of accomplishing tasks in the workplace (Kozlowski & Ilgen, 2006). Likewise, technology is enabling numerous employees to work from remote locations (telecommuting). In fact, in 2003, over 15 million people were reported to be telecommuting, with an average annual growth rate of 22% (U.S. Department of Transportation, 2003). When telecommuters, or any workers distributed across different organizations, locations, and/or time zones use technology to work together as a team, this can be considered a "virtual team" (Martins, Gilson, & Maynard, 2004; Townsend, DeMarie, & Hendrickson, as cited in Bell & Kozlowski, 2002b).

While there has been an explosion of team research over the past decade, Cascio and Aguinis (2008) concluded that there are still areas requiring further study to bridge the gap between science and practice. Based on their content analysis of all *Journal of Applied Psychology* and *Personnel Psychology* articles published between 1963 and 2007, they recommended that "strategies for…managing the performance of members of global virtual teams" be investigated in future research (p. 1077). Because over 60% of professional employees are now involved in virtual teams (Kanawattanachai & Yoo, 2002), it is important to understand what facilitates their performance. This study looked at storytelling and guided team reflexivity as two potential contributors to team mental models and virtual team performance. In

addition, team mental models (i.e., the overlapping of team members' organized mental representations of task-relevant information; Mohammed & Dumville, 2001; TMMs) were examined as a potential mediator between the storytelling and guided team reflexivity interventions and virtual team performance.

Storytelling is one of the oldest methods of transmitting knowledge across people and generations (e.g., Bal, 1997; Denning, 2001; Rosen, Fiore, McDaniel, & Salas, in press). With specific regard to teams, a story (also referred to as a "narrative") can be defined as a "structured expression of a given team member, or team's experiences" (Fiore, McDaniel, & Jentsch, 2009, p. 29). Storytelling has been described anecdotally as improving team learning and performance (e.g. Bartel & Garud, 2009; Denning, 2001; Fiore et al., 2009; Rosen et al., in press), but to my knowledge this has not been empirically tested as a planned team intervention. Team reflexivity refers to a team's overt reflection on its performance and goals (West, 1996), and guided team reflexivity specifically refers to a planned intervention to elicit team reflection (Gurtner, Tschan, Semmer, & Nägele, 2007). Reflexivity has been empirically investigated, but mainly as a naturally occurring phenomenon (e.g., Carter & West, 1998; De Dreu, 2007; Schippers, Den Hartog, Koopman, & van Knippenberg, 2008), as opposed to a guided intervention (cf. Gurtner et al., 2007). Therefore, the purpose of this study was to empirically investigate the impact of storytelling, as well as guided team reflexivity, on two outcomes in virtual teams: TMM similarity and team performance.

Four main contributions can be derived from this research, which are enumerated in the paragraphs below. First, storytelling theory was extended to teams in the empirical domain. A small amount of anecdotal and theory-based literature has described the positive effects of storytelling on outcomes such as team innovation and performance (e.g., Bartel & Garud, 2009;

Denning, 2001; Fiore et al., 2009), but no empirical research (to my knowledge) has been conducted to test these existing propositions. Similarly, although there have been numerous research efforts addressing tools and strategies for training teams (e.g., Blickensderfer, Cannon-Bowers, & Salas, 1997; Marks, Sabella, Burke, & Zaccaro, 2002; Volpe, Cannon-Bowers, Salas, & Spector, 1996), storytelling has not been examined as a formal intervention for transmitting knowledge, but rather as a naturally occurring, informal event. In addition, to my knowledge, there are no empirical investigations of the effects of any kind of storytelling. Thus, this research not only expands the storytelling construct to the empirical domain, but it also leverages the potential power of storytelling as a team training intervention.

The second contribution of this research was identifying a new potential antecedent of TMMs. In their review of existing TMM research to date, Mohammed, Ferzandi, and Hamilton (2010) stress that the outcomes of TMMs have been the main focal point of TMM research. Indeed, TMMs have been shown to be important determinants of team performance (e.g., DeChurch & Mesmer-Magnus, 2010a; DeChurch & Mesmer-Magnus, 2010b; Ellis, 2006; Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000; Smith-Jentsch, Mathieu, & Kraiger, 2005), but Mohammed and colleagues (2010) recommend that more attention should be given to potential antecedents of/tools for improving TMMs as well. For example, Marks and colleagues (2000) found that leader briefings and team interaction training had a positive effect on TMM similarity and accuracy. In addition, Fiore et al. (2009) suggested in a theoretical piece that storytelling may have a positive impact on cognitive structures such as TMMs by fostering a shared understanding of the story's meaning. However, storytelling has not been empirically investigated as an antecedent of TMMs. With Mohammed and colleagues' (2010)

recommendation in mind, there was value in examining storytelling as an additional predictor of TMMs.

A third contribution was the expansion of research on guided team reflexivity. Reflexivity is most commonly defined as "the extent to which group members overtly reflect upon the group's objectives, strategies, and processes, and adapt them to current or anticipated endogenous or environmental circumstances" (West, 1996, p. 559). Similar to the occurrence of storytelling, reflexivity can either naturally arise, or it can be used as a formal intervention to encourage growth and awareness of a team's circumstances and strategies (i.e., guided team reflexivity; Gurtner et al., 2007; guided team self-correction training; Blickensderfer et al., 2007). While there has been some research on the effects of naturally occurring team reflexivity on outcomes such as mental health and team performance (e.g., Carter & West, 1998; De Dreu, 2007; Schippers, Den Hartog, & Koopman, 2007), it appears that only one empirical article has addressed the use of guided team reflexivity (Gurtner et al., 2007). However, much of the reflexivity responsibility was placed upon the team leader as opposed to team members in this study. Therefore, the current research aimed to strengthen conclusions on the effectiveness of using reflexivity as a planned team intervention in situations where all team members have potentially valuable information and insights to provide.

The final contribution of this research is that it addressed the call from multiple researchers for more studies on virtual teams (e.g., Cascio & Aguinis, 2008; Martins et al., 2004). Martins and colleagues (2004) even specifically noted that "higher level cognitive outcomes…[in virtual teams] have not been examined extensively" (p. 822). Therefore, the three prior contributions that the study has provided can all fall under the larger contribution of expanding research on virtual teams. In other words, the effects of storytelling and guided team

reflexivity were specifically investigated in a distributed context, and the examination of TMM development in such a context directly addressed the concerns of Martins and colleagues (2004).

This thesis has been organized to facilitate a background understanding of the relevant constructs before describing the study's hypotheses, methodology, and results. Specifically, a brief overview of virtual teams and related research will be presented. Then, a review of TMMs will be provided, followed by literature reviews on storytelling and guided team reflexivity. In providing rationale for the hypotheses, the latter two constructs' proposed relationship with TMMs and team performance will be explored. Finally, an experiment designed to manipulate storytelling and guided team reflexivity in a virtual team context will be introduced, followed by results and discussion.

Virtual Teams

Virtual teams (VTs) are defined as "groups of geographically and/or organizationally dispersed coworkers that are assembled using a combination of telecommunications and information technology to accomplish a variety of critical tasks" (Townsend, DeMarie, & Hendrickson, 1998). While there are variations on this definition, Bell and Kozlowski (2002b) stress that spatial distance and communication method are the critical components that determine "virtualness." More specifically, VT members are not co-located, and regardless of whether they are only a few miles or countries apart, this affects the nature of how they communicate with each other. While less virtual (e.g., co-located) teams rely heavily on face-to-face interactions, teams high on virtualness mainly rely on technology-mediated communication such as e-mail, chat, and videoconferencing to communicate with each other.

Martins and colleagues (2004) note that virtual teams often do not differ from conventional teams with respect to their end goals, but rather in how they define and approach their goals. Specifically, VTs may have more difficulty establishing shared goals due to decreased interactions; therefore, Bell and Kozlowski (2002b) recommend that leaders provide formalized goals for virtual teams. In addition, some studies suggest that VTs communicate less frequently, take longer to reach decisions, and have lower performance outcomes (e.g., Andres, 2002; Bhappu, Zellmer-Bruhn, & Anand,1997; Graetz, Boyle, Kimble, Thompson, & Garloch, 1998). However, these findings are not conclusive, as other studies have found that VTs communicate just as frequently and make decisions of equal or greater quality as conventional teams (e.g., Hiltz, Johnson, & Turoff, 1986; Jarvenpaa, Rao, & Huber, 1988; Schmidt, Montoya-Weiss & Massey, 2001).

In light of many inconsistent findings when comparing virtual and conventional teams, Martins and colleagues (2004) recommend further empirical investigation of the effects of team virtualness on team outcomes. In particular, they state that inconsistent findings may be explained by a lack of research on mediating and moderating variables. Furthermore, they note that although team performance has often been studied as an outcome in the VT literature, cognitive outcomes have not. Therefore, this study incorporated these recommendations by examining the TMM construct as both an outcome in itself and as a mediator. TMMs are further described below.

Team Mental Models

TMMs represent "team members' shared, organized understanding and mental representation of knowledge about key elements of the team's relevant environment" (Mohammed & Dumville, 2001, p. 90). This allows team members to have a similar

interpretation of what is occurring and to better adapt and coordinate their actions to address a given task (Cannon-Bowers, Salas, & Converse, 1993). This can be especially difficult to achieve in virtual teams due to *team opacity*, or the absence of normally salient visual, auditory, and social team member cues (Fiore et al., 2009). Because these cues can be important facilitators of group processes, team opacity could have a detrimental effect on the amount of knowledge that is shared across team members (i.e., TMM similarity). Thus, it is especially important to determine how to facilitate TMM development in virtual teams.

TMM Content

Two main types of TMM content have been examined in the literature: task-related (i.e., equipment and task requirements) and team-related (i.e., team member roles and team interaction patterns) (e.g., Cannon-Bowers et al., 1993; Mathieu et al., 2000; Mohammed & Dumville, 2001). Furthermore, TMMs can represent different types of knowledge such as declarative (what), procedural (how), and strategic (context and application; Rouse, Cannon-Bowers, & Salas, 1992). Thus, team members often possess multiple TMMs simultaneously (Mohammed et al., 2010).

TMM Similarity and Accuracy

The TMM construct can represent similarity/overlap among various team members' models, and it can also represent the accuracy of those models. TMM similarity is also often referred to as "shared." While members' TMMs need not be identical, some overlap is useful for team performance (e.g., DeChurch & Mesmer-Magnus, 2010a, 2010b). The ideal amount of overlap may depend on various factors such as the type of task and the content of the TMM, but

has not been thoroughly investigated (Mohammed et al., 2010). TMM accuracy has not received as much attention when compared to TMM similarity. If a team's mental model is inaccurate, then having a shared awareness of that information may not enhance performance. Thus, TMM similarity in tandem with TMM accuracy may have the most positive impact on team outcomes (Mathieu et al., 2005; Mohammed et al., 2010). However, research on TMM accuracy has produced more mixed results as compared to TMM similarity (e.g., Marks et al., 2000; Mathieu et al., 2005; Webber et al., 2000). Therefore, for the purposes of this study, I focused on TMM similarity.

Outcomes of TMMs

Team performance has been the most extensively investigated outcome of TMM similarity. Various studies and meta-analyses have suggested that TMM similarity positively predicts team performance (e.g., DeChurch & Mesmer-Magnus, 2010a, 2010b; Ellis, 2006; Mathieu et al., 2000; Smith-Jentsch, Mathieu, & Kraiger, 2005), although there are inconsistent findings regarding the relative strengths of different content areas. For example, some research has found that taskwork mental model similarity has a stronger effect on team performance than teamwork models (e.g., Cooke et al., 2001; Mathieu et al., 2005), while other research has found that only teamwork mental model similarity has a direct effect on team performance (Mathieu et al., 2000). Still other research has found that the interaction of the two content areas has a significant effect on performance (e.g., Smith-Jentsch et al., 2005). Finally, team processes such as communication and coordination have been found to mediate the relationship between TMM similarity and team performance (e.g., Gurtner et al., 2007; Mathieu et al., 2000).

Antecedents of TMMs

Outcomes of TMMs have clearly been the main focus of TMM research, and Mohammed and colleagues (2010) recommend that more attention should be given to predictors and tools that foster TMM development. Existing research on antecedents of TMMs has found that team member characteristics, context, and team interventions influence the development of TMMs (e.g., Ellis, 2006; Marks et al., 2000; Smith-Jentsch, Campbell, Milanovich, & Reynolds, 2001). For example, job tenure and experience were positively related to shared TMMs (Smith-Jentsch et al., 2001), while stress was negatively related to shared TMMs (Ellis, 2006). Team-level interventions such as training (Marks et al., 2000) and reflexivity (Gurtner et al., 2007) have also been shown to be positively related to TMMs, as will be discussed in subsequent sections.

Storytelling

Stories, or narratives, may be one of the oldest methods of retaining and transmitting knowledge across generations (Denning, 2001; Fiore et al., 2009). The terms "narrative" and "storytelling" have been used interchangeably, as they both reflect the concept of expressing one's experiences (Fiore & McDaniel, 2006; Fiore et al., 2009). They will continue to be used interchangeably in this paper.

A story's purpose is to clearly structure and convey complicated ideas in a simple way (Carter, 1993; Denning, 2001; Doyle & Carter, 2003; Klein, 1998). A story may also facilitate problem solving when a situation is ambiguous with no clearly defined correct solution (Bartel & Garud, 2009; Rosen et al., in press). The story's content should be long enough and detailed enough to contain specific features such as a protagonist, a setting, a central theme, and an ending (Klein, 1998; Fiore & McDaniel, 2006). In addition, to be considered a story, there must

be various narrative components present (Bruner, 1991). For example, *context sensitivity* refers to how a story's recipient may interpret the story based on his/her own background, while *canonicity and breach* describes how a story must be interesting, with a surprising deviation from the expected sequence of events (Bruner, 1991).

How Stories Convey Meaning

Stories are intended to be engaging, drawing the audience in and allowing it to become connected to the protagonist via empathy, thus elucidating and enhancing the deeper meaning of the story (Denning, 2001; Klein, 1998). More specifically, a good story will allow the listener to see the application to his/her own context, thus allowing learning to occur. This bridging of ideas is termed "translation" and the story is the mechanism, or "boundary object," by which translation can occur (Bartel & Garud, 2009).

Some research on the application of narratives has focused on the individual level. For example, narratives have been explored as mechanisms for problem solving, effectively teaching students new concepts, creating educational computer programs, and comprehending blogs (Abrahamson, 2005; Ang & Rao, 2008; Gick & Holyoak, 1983; Pachler & Daly, 2008). However, the use of storytelling may be particularly important for teams, who must foster a sense of communication and collaboration in order to perform effectively. For example, Bartel and Garud (2009) theorize how storytelling can effectively allow teammate collaboration in innovative organizations. Specifically, teammates from different disciplines may find it difficult to convey their solutions for the team's deep level issues without causing confusion as a result of the more discipline-specific, surface-level details. Thus, finding an effective story that can

"translate" across disciplines can help multi-disciplinary teams and organizations achieve their goals.

Fiore and colleagues (2009) further theorize how storytelling may have positive implications for distributed teams by connecting team members socially, cognitively, and affectively. For example, storytelling may enhance social interaction by overcoming the less rich and perhaps asynchronous communication dynamics accompanying virtual teams. A more cognitive-oriented question would be how distributed decision making and TMMs are affected by storytelling. Finally, storytelling may have an affective impact, such as improving cohesion and trust in virtual teams. However, Fiore and colleagues (2009) stress that empirical research needs to be conducted. Therefore, the proposed study specifically measures the use of narratives in virtual teams.

Content of Stories

Fiore and colleagues (2009) note that the literature on storytelling tends to stem from business and organizational science, where stories typically reflect actual experiences of employees or group members. Similarly, Bartel and Garud (2009) discussed how employees would help others understand and incorporate the knowledge underlying their innovations by describing the "set of events and the contextual details surrounding their occurrence" (p. 108). Furthermore, Fiore and colleagues (2009) and Rosen and colleagues (in press) describe how stories can convey each team member's first-hand experiences, cumulating to an overall better picture of the situation. While these examples certainly convey the possibility of imparting knowledge via storytelling, they do not acknowledge whether the content of the story (i.e., first-hand experience versus more individually removed knowledge) differentially affects learning.

Denning (2001) takes the position that stories should be "close to home, but not too close" (p. 128). That is, stories will have more of an impact when they involve a topic that a person can relate to, but that has not been experienced personally or does not evoke pre-existing strong personal opinions. In that line of thinking, it may be that stories do not have to be based in truth or first-hand experience at all in order to facilitate understanding, as long as they are relatable. This may be particularly important in the context of team training. In other words, some of the existing literature has focused on storytelling as an ad-hoc reaction to circumstances that facilitates team member understanding and problem-solving (e.g., Bartel & Garud, 2009; Fiore et al., 2009). However, if a team encounters circumstances with which no member is personally familiar, then reactive storytelling is not useful. Rather, if any story with relevant content can be proactively shared in the form of a training intervention, then this could have a positive effect on team members' understanding and subsequent performance.

Storytelling as a Team Training Intervention

Team training can be defined as "a set of strategies that create a context in which team skills can be practiced, assessed, and learned" (Salas & Cannon-Bowers, 1997). There have been numerous research efforts addressing diverse approaches to training teams. For example, lectures, group activities, simulations, and feedback have all been suggested as useful tools for improving team performance (Goldstein & Ford, 2001). Specific types of team training include cross-training (i.e., training each team member on other team members' tasks, duties, and responsibilities; Blickensderfer, Cannon-Bowers, & Salas, 1997; Marks et al., 2002; Volpe, Cannon-Bowers, Salas, & Spector, 1996), team coordination and adaptation training (i.e., improving teamwork and team decision-making abilities; Kozlowski & Ilgen, 2006; Salas &

Cannon-Bowers, 2000), and team interaction training (i.e., "the training of task information embedded in the necessary teamwork skills for effective team task execution"; Marks et al., 2000, p. 974). However, little research has been conducted on the use of narratives as a tool for team training.

As mentioned previously, teams may encounter novel situations, making it very difficult for personal narratives to be effective. However, because it is important for teams to be able to adapt to varying situations and respond accordingly (Coovert, Craiger, & Cannon-Bowers, 1996; Marks, Zaccaro, & Mathieu, 2000), then not only should storytelling be purposefully introduced as a proactive training intervention, but it should address key teamwork skills that could be more effective across a wide range of conditions than specific task training (Marks et al., 2000; Stevens & Campion, 1994).

Storytelling and Team Performance

Although there is not much empirical research upon which to base hypotheses regarding storytelling as a training intervention and team performance, prior studies have found that training in general has a positive relationship with performance (e.g., Salas, Nichols, and Driskell, 2007; Rapp & Mathieu, 2007; Volpe et al., 1996). For example, Volpe and colleagues (1996) found that cross-training improved team performance on a flight simulation task, while Rapp and Mathieu (2007) discovered that technology-based generic skills training increased MBA student team performance. Additionally, Salas and colleagues (2007) conducted a meta-analysis of seven studies involving training strategies and found a small to moderate effect size on team performance.

With respect to storytelling and performance, Fiore and colleagues (2009) suggested that narratives "can be used to facilitate the transfer, comprehension, and retention of team-related information…to better improve…overall team performance" (p. 35). For example, storytelling can facilitate understanding of complex ideas, thus enabling team members to engage in collaborative problem solving and achieve performance goals (Bartel & Garud, 2009; Denning, 2001: Rosen et al., in press). Therefore, I propose the following:

Hypothesis 1: Storytelling is positively related to team performance.

Storytelling and Team Mental Models

According to Salas, Cannon-Bowers, and Johnston (1997), team training is intended "to foster in team members an accurate and sufficient mental representation of the team task structure, team role, and the process by which the two interact" (p. 362). Therefore, training has been investigated as a mechanism for TMM development (Mohammed et al., 2010). For example, Marks and colleagues (2000) demonstrated that team interaction training had a positive effect on TMM similarity, while Smith-Jentsch et al. (2001) found their computer-based training on generic teamwork skills positively affected teamwork mental models. Furthermore, Marks et al. (2002) showed that cross-training elicited similar team-interaction mental models. These experiments support the idea that training can have a positive impact on TMMs, but no prior empirical research appears to have addressed the use of storytelling as a training tool despite anecdotal and theoretical support, as well as the call for more research on antecedents of TMMs (Mohammed et al., 2010).

There has been conceptual literature discussing how narratives can enhance team cognition by allowing the "translation" of ideas across disciplines (e.g., Bartel & Garud, 2009; Fiore et al., 2009). If translation allows different parties to have a shared understanding of a narrative's deeper meaning, then this is presumably a result of developing shared mental models. Fiore and colleagues (2009) specifically indicate that narratives should be investigated for their potential as a tool in cognitive learning tasks. They propose that when information is presented "through the lens of…narrative perspective…[this] may strengthen a team's shared mental model associated with their task and teammates" (p. 34). Thus:

Hypothesis 2: Teams whose members receive a story have more similar team mental models than teams whose members do not receive a story.

Hypothesis 3: The relationship between storytelling and team performance is mediated by team mental model similarity.

Reflexivity

Reflexivity is most often defined as "the extent to which group members overtly reflect upon the group's objectives, strategies, and processes, and adapt them to current or anticipated endogenous or environmental circumstances" (West, 1996, p. 559). Reflexivity involves an analysis of what the group has accomplished, what it needs to accomplish, and how it can do so. This can either be an individual process, whereby team members independently reflect on their performance and do not necessarily communicate with each other (e.g., Barge, 2004; Barry, Britten, Barber, Bradley, & Stevenson, 1999; Gurtner et al., 2007) or an interactive team process,

which involves explicit communication in the form of a group discussion (e.g., De Dreu, 2007; Gurtner et al., 2007; Schippers et al, 2007). However, Schippers and colleagues (2007) point out that team reflexivity is most often conceptualized and measured as the latter: an overt group process.

West (1996) proposed that reflexivity is most useful in complex decision-making teams (CDMT). This is because these teams operate in environments that can change and present new challenges/tasks over time, tasks are often complex, there is high team member interdependence, and team members have autonomy over their work. In contrast, a simple decision-making team may perform the same tasks repeatedly without the autonomy to alter how the task is approached; once the task is understood, no further discussion may be necessary. A CDMT may need to reflect on its objectives more frequently because there are often no pre-defined correct solutions to tasks that could often change. By reflecting on whether goals are being met, and collaboratively identifying strategies and processes to better achieve those goals, a reflexive CDMT can more readily adapt to an ambiguous or changing environment. Thus, reflexivity can represent the idea of "double-loop learning" (Argyris, 1993), whereby groups may reflect on objectives, plan strategies to accomplish them, and then enact those plans, which can subsequently lead back to additional reflection on the plans' effectiveness. The relationship between reflexivity and learning is further detailed below.

Reflexivity and Team Learning

Although team reflexivity and team learning have stemmed from different literatures, the constructs are closely linked, and their distinction is not clear (Edmondson, Dillon, & Roloff, 2007; Kozlowski & Ilgen 2006). For example, Wilson, Goodman, and Cronin (2007)

acknowledged that the sharing stage of learning is "closely related" to reflexivity (p. 1046) because group members discuss information that needs to be learned and create a common strategy for using that knowledge in the future. Furthermore, Edmondson (1999) explicitly defined group learning as an "ongoing process of reflection and action" (p. 353), whereby group members discuss what they've done and how to improve. In fact, in their development of a team learning behavior instrument, Savelsbergh and colleagues (2009) utilized numerous items from reflexivity measures.

Team learning itself has not been defined in an entirely consistent manner. For example, Argote, Gruenfeld, and Naquin (2001) define it as when individuals collect, share, and combine information with each other, while Argyris and Schön (1995) use a more simply definition of error detection and correction. Team learning has been categorized in terms of outcome improvement and group process (Edmondson et al., 2007). Outcome improvement views learning as performance improvement, usually via efficiency. This stream of research focuses primarily on objective performance outcomes such as time and cost reduction. Group process, the most similar to the team reflexivity construct, refers to team members' learning behaviors, which include seeking feedback, discussing errors, and openly evaluating the team's work (e.g., Edmondson, 1999, 2002; Van der Vegt & Bunderson, 2005).

Edmondson and colleagues (2007), in their review of the extant literature on team learning, agree that team learning is a very broad term that "should remain an encompassing rubric" (p. 300). Likewise, Kozlowski and Ilgen (2006) concluded that "the research base to specify the meaning of team learning as a construct distinct from other team cognitive constructs…[is] not yet sufficiently developed" (p. 87). Therefore, while recognizing the

parallels between the team learning and team reflexivity constructs, I draw more heavily from the reflexivity literature in deriving hypotheses.

Reflexivity and Team Outcomes

Various studies have found that reflexivity has a positive relationship with team outcomes (e.g., Carter & West, 1998; De Dreu, 2007; Gurtner et al., 2007; Schippers et al., 2008; Schippers et al., 2007). For example, Gurtner and colleagues (2007) found that team reflexivity led to better performance on a military air surveillance simulation, while Schippers and colleagues (2008) discovered that team reflexivity partially mediated the positive relationship between transformational leadership and team performance. In addition to team performance outcomes, team reflexivity has been linked to information sharing and learning (De Dreu, 2007), satisfaction and commitment with one's team (Schippers et al., 2003), and communication and strategy implementation (Gurtner et al., 2007).

Guided Team Reflexivity

While there has been some research on team reflexivity and its effect on team outcomes, there appears to have been little empirical research on using reflexivity as an active intervention tool. In other words, most studies have measured team reflexivity in correlational studies, rather than investigating the causal effects of "an intervention to induce reflection in groups" (Gurtner et al., 2007, p. 128), which has been termed guided reflexivity. Guided reflexivity appears to have only been empirically investigated once in the published literature, by Gurtner and colleagues (2007). Using participants in a laboratory setting, they assigned teams of three (one leader and two subordinates) to participate in seven simulated tactical aviation scenarios,

assessing the threat levels of various planes. Halfway through the seven scenarios, team members in the reflexivity condition, who could only communicate via e-mail, were given three questions guiding them on how to engage in group reflexivity, with those questions representing reflecting, planning, and implementing/adapting.

Gurtner and colleagues (2007) found that teams in the reflexivity condition performed significantly better than teams in the control condition, even when controlling for prior team performance. These results suggest that not only may reflexivity be useful when it naturally occurs in teams (cf. Carter & West, 1998; De Dreu, 2007; Gurtner et al., 2007; Schippers et al., 2008; Schippers et al., 2007), but that actively introducing reflexivity can be beneficial as well. In fact, some researchers have suggested that spontaneous reflexivity is less likely to occur when it is most necessary. For example, Weingart (1992) found that teams facing a complex task engaged in less planning than those facing a simple task, and West (1996) proposed that the longer a team is together, the less its members will be aware of environmental changes and the less reflexive they will become. Even before the concept of reflexivity was introduced, Hackman, Brousseau, and Weiss (1976) suggested that teams would not strategize without prompting. Therefore, there may be circumstances where it would be particularly beneficial to formally intervene with an initiative that directly encourages team reflexivity.

Furthermore, guided reflexivity may be useful for ad-hoc teams and for training purposes. First, if teams are more likely to be receptive to creating strategies and plans early in their tenure (e.g., Hackman & Wageman, 2005; West, 1996), then making certain that they are exposed to reflexivity opportunities at that time is crucial. In addition, ad-hoc teams are new, and members may therefore feel uncomfortable initiating group reflexivity on their own due to a lack of

psychological safety (Edmondson, 1999). A formal intervention may compel team members to engage in more productive reflexive discussions.

Second, guided reflexivity could prove useful in situations where team members need to be trained to work together in complex or changing environments. Researchers have suggested a concept similar to team reflexivity termed "team self-correction," whereby team members provide each other with feedback on their task performance (Blickensderfer et al., 1997; Smith-Jentsch, Cannon-Bowers, Tannenbaum, & Salas, 2008). Specifically, the process begins by objectively reviewing what occurred during the task to form an overall group picture, followed by identifying errors and what should be done to avoid those errors in the future. This concept is very similar to guided team reflexivity, and Smith-Jentsch, Zeisig, McPherson, and Acton (1998) even recommended that team self-correction should be structured (i.e., "guided"; Smith-Jentsch et al., 2008) to avoid focusing efforts and strategies too narrowly to particular tasks, as opposed to more generally applicable contexts. Thus, guided reflexivity/self-correction training would prove particularly useful in an environment where tasks are complex or often changing.

Although the two concepts are similar, one manner in which guided team self-correction differs from guided team reflexivity is that it revolves around a "pre-specified expert model of teamwork" (Smith-Jenstch et al., 2008, p. 312). Gurtner and colleagues (2007) argue that these expert models cause guided team self-correction to be more task-specific than guided team reflexivity. Smith-Jentsch and colleagues (2008), on the other hand, explain that the expert model consists more of general teamwork behaviors, such as information exchange and supporting behavior. Another distinction between the two concepts is that guided team self-correction training involves using a trained facilitator to review performance and guide team members in creating strategies using an expert model as a frame of reference, while guided team

reflexivity is not quite so structured and allows team members more flexibility in their discussion. However, despite these distinctions, the two constructs appear very similar, and Gurtner and colleagues (2007) assert that if an expert model or formal facilitator is not readily available, then reflexivity is a viable, less expensive alternative. Therefore, while acknowledging the distinctions, this study draws from both sets of literature in developing hypotheses.

Although existing research on team reflexivity has revealed a relationship with team performance and meeting deadlines (e.g., Gevers et al., 2009; Schippers et al., 2007), the one published study on guided team reflexivity (i.e., Gurtner et al., 2007) utilized teams with an asymmetrical distribution of information and power. That is, reflexivity mainly depended on the leader's making strategy suggestions to the two subordinate members based on information that only he/she had. Because of this imbalance of power, it is unclear whether teams whose members have equal potential to contribute constructively to guided team reflexivity will demonstrate similar increases in performance. However, Smith-Jentsch and colleagues (2008) did find that guided team-self correction training was positively related to the performance of Navy lieutenants, so I expect similar results for guided team reflexivity, and hypothesize the following:

Hypothesis 4: Teams that engage in guided team reflexivity demonstrate better team performance than teams that do not engage in guided team reflexivity.

Guided Team Reflexivity and Team Mental Models

If team members have the opportunity to discuss their experiences and collaborate to

create strategies for future performance episodes, then this should enhance TMM similarity, thus allowing teammates to approach tasks in similar ways (Blickensderfer et al., 2007; Gurtner et al., 2007; Smith-Jentsch et al., 2008). For example, Mathieu and colleagues (2000) found that without any form of after-action review, team members participating in a flight combat simulation did not develop greater levels of mental model similarity over time, concluding that "teams need guided experiences…if we expect them to learn" (p. 280). Supporting this argument, Gurtner et al. (2007) found that guided team reflexivity increased TMM similarity, and Smith-Jentsch et al. (2008) found results in the same direction for guided team self-correction training, though they were not significant for the 25 teams studied. Together, the evidence would suggest that guided team reflexivity is positively related to TMM similarity.

Hypothesis 5: Team members who engage in guided team reflexivity have more similar team mental models than those that do not engage in guided team reflexivity.

If guided team reflexivity is positively related to TMM similarity, and many researchers have also found that TMM similarity is positively related to team performance (e.g., DeChurch & Mesmer-Magnus, 2010a, 2010b; Ellis, 2006; Mathieu et al., 2000; Smith-Jentsch, Mathieu, & Kraiger, 2005), this could suggest that TMM similarity mediates the relationship between guided team reflexivity and team performance. In fact, Blickensderfer and colleagues (2007) specifically proposed that team self-correction would improve team performance via TMM similarity, and a mediated relationship was supported by Gurtner et al.'s (2007) findings. However, since direct relationships between reflexivity and performance have also been found

(e.g., De Dreu, 2007; Schippers et al., 2008), it is likely that TMM similarity serves as a partial mediator between guided team reflexivity and team performance. In other words:

Hypothesis 6: The relationship between guided team reflexivity and team performance is partially mediated by team mental model similarity.

Storytelling and Guided Team Reflexivity

Although it has been theorized (Fiore et al., 2009), the use of narratives for improving TMMs and/or team performance does not appear to have been empirically tested, and certainly not in combination with guided team reflexivity. If on their own, narratives and guided team reflexivity are hypothesized to improve TMM similarity and team performance, then it is possible that when used in tandem, their effects are enhanced. For example, if teams are exposed to a story and then subsequently engage in guided team reflection, the story could provide a structured framework (as recommended by Smith-Jentsch et al., 1998) that could help to focus the team's discussions (Fiore et al., 2009). That is, instead of discussion strategies that are too general (e.g., "We need to work together better") or too task-specific (e.g., "Remember that next time you see a plane flying at that exact altitude, it's an enemy"), a storytelling intervention could encourage discussion on particular teamwork skills that are useful across various tasks. This may be especially salient when there are fewer social cues available to facilitate reflexivity, such as is the case with virtual teams (Martins et al., 2004). Therefore, a moderated relationship is hypothesized, potentially impacting team performance or TMM similarity:

Hypothesis 7a: The relationship between storytelling and team performance is moderated by guided team reflexivity, such that storytelling has a stronger effect on team performance when there is also a reflexivity intervention, and a weaker effect when there is no reflexivity intervention.

Hypothesis 7b: The relationship between storytelling and team mental model similarity is moderated by guided team reflexivity, such that storytelling results in more similar mental models for teammates who have also received a reflexivity intervention, and less similar mental models when there is no reflexivity intervention.

If Hypothesis 7a or 7b is supported, and there is also support for a mediated relationship as described in Hypotheses 3 and 6, then there would be a possibility of mediated moderation as well. All seven hypotheses are summarized in an overall research model, presented in Figure 1.

Method

Participants

Overall, this study recruited 413 participants (203 male, 201 female, 9 unknown) in 145 teams. Participants were drawn from the undergraduate psychology research subject pool, as well as recruited from Information Sciences and Technology (IST), Psychology, and Human Development and Family Studies (HDFS) undergraduate classes at the Pennsylvania State University. A total of 38 teams were dropped before conducting analyses. Specifically, 29 teams were dropped due to procedural or technical errors, four teams were dropped due to incomplete data, and an additional four teams were dropped due to the participants not following directions during the reflexivity sessions. Finally, one team was dropped for failing to answer the manipulation check questions correctly. Therefore, final analyses were conducted on 321 students, comprising 107 three-person teams.

The NeoCITIES Team Simulation

This study consisted of a laboratory experiment using a computer simulation called NeoCITIES. NeoCITIES uses emergency management events to allow an examination of the behaviors and performance of spatially distributed decision-making teams (McNeese, Bains, Brewer, Brown, Connors, & Jefferson, 2005). The most recent version, NeoCITIES 3.1, was developed by the Information Sciences and Technology Multidisciplinary Initiatives in Naturalistic Decision Systems (MINDS) lab through funding by the Office of Naval Research. It is an outgrowth of several prior versions, dating back to Wellens and Ergener's (1988) original CITIES (Command, Control, and Communication (C^3) Interactive Task for Identifying Emerging Situations) task (Hellar, 2009). NeoCITIES 3.1 is a web-based application that uses instant chat technology and three team members each assigned to a unique role: Police, Fire, or Hazardous Materials (Hazmat). Within each role are three types of resources. For example, within the Police role, there are Investigators, Squad Cars, and SWAT Teams (for a complete listing of the resources within roles, please refer to Appendix C).

The objective of the simulation is to dispatch the relevant type and number of resources to emergency management events that appear on the screen (see Appendix D for a screenshot of the simulation console). These events are scripted, appearing at pre-designated times and all taking place on the University Park campus. Specifically, each participant is tasked with "determining the severity of incoming events, deciding on [the] appropriate measure of response, and coordinating [his/her] actions with the managers from the two other [roles]" (Hellar, 2009, p. 56).

Events can vary in timing, duration, and intensity. A small routine event can usually be solved by a single team member and represents typical events that could occur for emergency

response teams, while other events are more complex, requiring resources from multiple teammates. For example, a small-scale independent event could be the Fire team member putting out a trash can fire. A larger, interdependent event could be a terrorist who has released an airborne chemical on campus, which would require the resources of Police, Fire, *and* Hazmat. If such events are not addressed quickly and/or correctly, the complexity of the event could escalate further, requiring even more resources and ultimately leading to failure of the event if not properly addressed (e.g., the fire spreads to nearby buildings, the airborne chemical spreads across the whole campus; Hellar, 2009). Thus, not only is teamwork emphasized in this simulation by requiring participants to collectively solve many of the events, but the importance of temporal dynamics (i.e., deadlines, pacing, and sequencing) has been infused into the simulation as well.

The NeoCITIES simulation represents a key strength of the study. It is set up to accommodate teamwork and is flexible enough to vary the intensity of the user's experience (Hellar, 2009). Its portrayal of realistic events in a familiar setting (i.e., the Penn State campus) creates a more salient and meaningful user encounter. Participants consistently report that the simulation is engaging and holds their attention. Because events can be custom-made, the experimenter can create events that reflect not only time pressure (e.g., Police needs to disarm a bomb in the next 60 seconds, three separate events appear at the same time), but that reflect a range of magnitudes and types of potential loss (e.g., destruction of a building from a fire, loss of money from a bank robbery, or loss of life from a car bomb). Furthermore, because teams are working in a self-contained system and can carry out tasks without outside (i.e., leader) intervention, as well as engaging in technology-mediated communication, this simulation setup is also very representative of self-managing virtual teams (e.g., Manz & Sims, 1987).

Laboratory Setting

There are two laboratories, and each is set up so that two teams can be run at once. One lab is located in the Psychology Department, and the other in the College of Information Sciences and Technology. The two labs are set up in a similar fashion, with three computer stations in a row on one side, separated by dividers, and another three computer stations directly across from the first row, also separated by dividers. There are also two server computers used by the experimenters at the front of the two rows (see Figures 3 and 4 for diagrams of the two lab setups). These labs are supported by a grant from the Office of Naval Research, and the study was part of the output from this grant.

Experiment Design

The study design consisted of a fully-crossed 2 (guided team reflexivity vs. none) x 2 (storytelling vs. none) factorial design. There were four conditions, with the number of participants/teams (randomly assigned) in each condition shown in Figure 2.

Preliminary Study

A preliminary study was conducted to investigate the effect of storytelling on team performance in the NeoCITIES simulation (Mohammed & McNeese are the principal and co-principal investigators, respectively). The current study was an extension of this preliminary study and used a similar experimental design, which will be described in the Experimental Procedure sections below. There are only three notable differences between the two studies.

First, the preliminary study did not address the construct of reflexivity, which was a formal intervention in the current study. Second, there were four storytelling conditions in the preliminary study, with each story containing varying teamwork training content (see Appendix E for an example of a story script). Specifically, there were two stories that dealt with a patient with post-operative complications in a hospital setting, and two stories that dealt with a victim of chemical burns in a NeoCITIES setting (i.e., an event on the Penn State campus involving Police, Fire, and Hazmat). Within each of these pairs, one story described the negative consequences when teammates do not communicate and collaborate with each other, and the other story described the negative consequences when teammates do not meet their deadlines or address their tasks in the correct order.

In the current study, I chose to investigate the effects of only one of these stories, intending for the results of the preliminary study to inform my decision. However, few differences were found between story types as they related to TMMs or performance. Therefore, the choice of which story to utilize was based on previous literature (e.g., Gentner, 1983; Ross, 1989) and reactions from participants. In the cognitive science literature, structure mapping theory (Gentner, 1983) describes how an analog (base) that has deep-level similarities to the problem (target) at hand can help with problem solving, even if the surface-level details are not similar. However, in the absence of extensive training, a person tends to recall an analog with a similar surface structure more easily (e.g., Gentner & Landers, 1985; Ross, 1989). Therefore, in this experiment, the story that was most analogous to the experimental setting in both surface and deep level structure (i.e., the story within the NeoCITIES setting) was selected. Furthermore, because the story involving deadlines and sequencing was reported to contain a greater amount

of applicable lessons, as opposed to the general message of "Communication is good," the NeoCITIES story with timing content was chosen.

The third difference between the preliminary and current study was that the storytelling intervention in the preliminary study was consistently introduced after all of the basic experiment training, but prior to any performance scenarios. However, if a newly formed team facing a series of tasks is bombarded with both task- and team-related training information all at once, then team members may not retain as much as if the information had been presented in a more sequential manner over time, allowing a more gradual accumulation of knowledge. Furthermore, Hackman and Wageman (2005) raised the point that certain task-related strategies should be introduced after a team has already gained some experience with the task. In other words, if team members already have a solid grasp of the task through first-hand experience, then they will be able to better identify how additional team interventions can be applied to future tasks of a similar nature. Following this logic and the past example of Gurtner and colleagues (2007) with their team intervention, in the current study the storytelling intervention was introduced in between the two performance scenarios.

Experimental Procedure

Participants were directed to computer cubicles separated by dividers. The participants signed a consent form, also keeping a copy for themselves. The experimenters (typically two per session) then introduced themselves to the group and instructed the participants to put on headphones, which remained on for the duration of the experiment to minimize distractions from outside noises, including auditory nonverbal cues from teammates, such as sighs. The participants then filled out a short online questionnaire assessing demographic information such

as gender, age, familiarity with teammates, school major, knowledge of emergency response protocol, and experience working in virtual/distributed teams. Some of these data were for descriptive purposes, and others served as control variables because they could potentially affect team performance.

The participants then viewed an 8-minute training video on how to play the NeoCITIES simulation (see Appendix F for a sample of the slides used to make the video). Within this training, it was stressed that participants should not try to look around the dividers at each other, and that all communication should be restricted to the chat function within the simulation. Participants were then directed to a unique set of slides that informed them of the role to which they had been assigned (Police, Fire, or Hazmat) and the typical duties of each of the resources within that role. In addition, they reviewed a brief summary of the duties of the other team members' roles. It should also be noted that in order to prevent unintended effects of demographics such as gender and age, participants were not be aware of who was assigned to which role.

After completing the video training session on NeoCITIES, participants engaged in two sets of interactive practice sessions, lasting approximately five minutes each. The purpose of these practice sessions was to allow participants to become experienced at solving routine scenarios in NeoCITIES. The first practice session consisted of simple, spaced out events that required no interdependence (i.e., no resources from multiple roles; see Appendix G for a examples of events from the sessions). At the end of the session, participants were able to view their team performance score as well as the solutions to the events. They then viewed a brief one-minute video explaining how to work together with their teammates. Specifically, the use of the chat function and team resource monitor (which allows a participant to see what resources

his/her teammate has sent to an event) was explained. The participants then engaged in a second practice scenario so they could better understand and correct their previous errors. The second set of events mirrored the first, but were slightly more difficult (though not taxing, as determined by pilot testing) in terms of time pressure and role interdependence.

At this point, the participants were told that the first official performance scenario was about to begin. Participants were reminded to only communicate via chat and to keep their eyes on their own computer. This first performance scenario lasted approximately 15 minutes. Events were timed at a slightly faster pace than the training sessions, and there was a mix of independent and interdependent events varying in magnitude (e.g., breaking up a bar fight, containing a riot on the football field). Specifically, there were a total of 19 events all revolving around the theme of a football game against Ohio State University. Thirteen of these events were independent (i.e., only requiring resources from one participant), and 6 were interdependent (i.e., requiring resources from at least two participants). As determined by pilot testing, participants viewed the events as interesting and engaging, maintaining their attention throughout each scenario.

During the first scenario, one information briefing was distributed to participants. This briefing contained information critical to correctly responding to certain events, particularly with regard to temporal dynamics. However, the information was distributed in a hidden profile format, whereby each team member received one unique piece of information. For example, Police was told what the time limit is for responding to the event (e.g., bank-related threats must be addressed within 45 seconds), while Fire was told the order in which the team should respond to the event (e.g., Fire should arrive first to treat injuries, then Police should search the area for evidence, and Hazmat can then clean up any chemical substances on scene). Therefore, the

participants needed to share the information in their briefings to successfully complete such events (see Appendix H for an example of a briefing).

At the conclusion of the first performance scenario, the participants viewed their team performance scores. They then filled out online questionnaires assessing their reactions to and perceptions of the scenario, followed by a storytelling and/or reflexivity manipulation, or neither, depending on the condition.

At this point, the second performance scenario began. It revolved around the theme of final exams week, and also lasted about 15 minutes. There again was a mix of independent (11) and interdependent (6) events, with two sets of information briefings being distributed. Each set contained unique information for each role, again allowing a hidden profile task. Following the scenario, reactions were measured online, followed by the participants' filling out paper-based grids designed to measure TMMs (see Measures section). This concluded the experiment, and the participants were debriefed and dismissed. The study took approximately 2.5 hours to complete.

Storytelling Manipulation

After viewing their performance in the first performance scenario, participants watched a 4.5 minute video telling the story of a graduate student who had serious health injuries incurred due to a lack of coordination and timing by his emergency response team (see Appendix E for the complete script). The surface-level structure of the story contained features analogous to the NeoCITIES simulation such as involving the roles of Police, Fire, and Hazmat teams, and events taking place on a university campus. The deep-level structure (i.e., the deeper meaning of the story) also reflected parallels to the NeoCITIES simulation by illustrating the importance of

meeting deadlines (e.g., a student did not receive assistance in time to prevent nerve damage from a spilled chemical in the lab), communication (e.g., Hazmat did not tell Fire what the identity of the chemical substance was), and sequencing (e.g., Police needed to clear the roads of snow so Hazmat and Fire could reach the student, followed by Hazmat identifying the chemical substance, which would then allow Fire to administer the appropriate treatment.).

In the storytelling intervention, to establish credibility, the participants were first told that past participants had found the upcoming story useful in improving their performance, followed by the claim that it was based on a true story. A narrator's voice was heard while slides showing pictures and other visual summaries of the key points of the story were displayed. The objective of the story was to convey the message that team members must collaborate with each other and time their responses well (i.e., meet deadlines, arrive at events in the proper order) if they were to succeed in their scenario. As described in the Preliminary Study section, the NeoCITIES analog story was selected for the storytelling manipulation in this study, based on the fact that people tend to recall an analog with a similar surface structure more easily than one with only a similar deep level structure (e.g., Gentner & Landers, 1985; Ross, 1989).

For the storytelling control groups, the participants viewed a 1-minute video summarizing the same deep structure points as the story, but without the context of a story. To avoid any possible implications of this intervention lasting less time than the other, such as "empty time" that could allow for additional processing of the meaning of the video's message, the video was preceded by a 3.5 minute filler survey. The survey was sufficiently long enough to preclude anyone finishing within 3.5 minutes, and the survey did not affect the participants' mood or experiment engagement, as determined during pilot testing.

Guided Team Reflexivity Manipulation

Between the first and second performance scenarios, and after the storytelling manipulation, participants in the guided team reflexivity condition were given six minutes to reflect on their performance via chat. First, the experimenter distributed the instructions and three discussion questions in paper form. He/she then read the instructions from a script, informing the participants that when teams can reflect upon and discuss their task performance and strategies, they can improve their scores. The participants were told that this was an opportunity to discuss their performance and how they wanted to improve it, but that they could only communicate via chat. They were also reminded to take into account their training, the video, and performance from scenario 1 during their discussion. The experimenter then briefly explained the three discussion questions, emphasizing that participants should not discuss a given question until instructed to do so. The teams were given all of the questions in advance so that they could have an overall sense of the purpose of their discussion session, and to help them understand how their discussion should be targeted for each question. Pilot testing revealed that this was the most favorable way to present the questions.

The participants were logged into a private group chat room, and at predetermined times the experimenter typed each of three questions into the chat to help guide the discussion. However, at no point did the experimenters actually participate in the chat discussion. In fact, they maintained "invisible" status throughout the discussion, with their user id not being listed in the list of members in the chat room.

The three questions that were used to guide the discussion are based on the questions used by Gurtner and colleagues (2007) in their reflexivity study, as well as West's (1996) definition of reflexivity. They are:

1. What were the main points you learned from the video you just watched that could apply to playing NeoCITIES?

2. How well do you think you and your team performed in Scenario 1? What went right? What went wrong?

3. You will be playing another performance scenario shortly. As a group please come up with strategies for how to improve your performance on this upcoming scenario.

Teams were given 1.5, 1.5, and 3 minutes, respectively, to answer the three questions. Pilot tests revealed that this gave participants adequate time to answer each question.

In the control group, the participants were given a topic to discuss that was unrelated to the NeoCITIES experiment or reflexivity. Specifically, they were given six minutes to discuss the following question:

Is the use of technology a benefit or detriment to society and interpersonal relationships?

For the complete experimenter script and all discussion questions, please refer to Appendix I.

Measures

Team Performance

A team's overall performance consisted of scores from independent events (i.e., those requiring the resources of only one team member) and interdependent events (i.e., those requiring the resources of two or three team members). Because the storytelling and reflexivity interventions were aimed at improving *team* performance via TMMs, it stands to reason that the events requiring teammates to work together would be most affected. Therefore, only performance on interdependent events was examined. Specifically, team performance was operationalized as the number of interdependent events successfully solved within a scenario.

TMM Similarity

Because TMMs need to be tailored to the specific context, there is no consistent method in which they have been measured (Mohammed et al., 2010). However, both content and structure are considered crucial components of any TMM measurement approach, as they reveal the pattern of relationships between TMM elements (DeChurch & Mesmer-Magnus, 2010; Mohammed et al., 2010). Meeting both of these criteria, TMMs were evaluated using concept mapping, where "participants place researcher-generated concepts in a pre-specified hierarchical structure depicting the sequence of activities required to perform the team task" (Mohammed et al., 2010). Concept maps are a popular TMM measurement tool in the literature (e.g., Ellis, 2006; Marks et al., 2000, 2002).

In this study, a concept map consisted of three boxes indicating the order in which units were supposed to respond to a given interdependent event. In each performance session, there were three interdependent events that included order requirements. Thus, a complete concept map for a performance session contained three sets of three boxes each to fill in. For a visual example of a concept maps, please refer to Appendix J.

Similarity in participants' responses to the concept maps indicated that teammates were communicating and collaborating with each other, particularly since for some events information about order requirements was only given to one teammate in the form of an intelligence briefing, as described in the experiment procedure above. If that team member did not share his/her information, then the other teammates would most likely not have the same answers. To measure similarity, corresponding boxes were compared across pairs of teammates for the same answer, regardless of whether the answers were correct or not. A pair with the same answer

received a 1 for that box, and a pair with different answers received a 0. Because there were three possible pair-wise comparisons for a team of three, and a total of nine boxes in the concept map measure, this resulted in 27 total values of 0 or 1. These values were averaged to represent the final concept map similarity score: the percentage of correct matches between teammates with regard to the response order for interdependent events.

Manipulation Checks

Storytelling. The manipulation check for storytelling was evaluated using objective measures. After viewing the storytelling video, participants were given two categorical questions asking whether they heard a story during the experiment and what the content of the story was.

Reflexivity. After the reflexivity session, participants were given questions to both objectively and subjectively evaluate the reflexivity manipulation. Specifically, they were given categorical and scaled questions. First, participants answered an objective multiple choice question asking what their discussion session was about, followed by Likert-style questions intended to assess whether they felt they were given the opportunity to discuss their performance scores and devise strategies for improvement. Eight items from Carter and West's (1998) task reflexivity scale were adapted and used for this purpose. Sample questions include, "The team reviewed its performance so far" and "The methods used by the team to dispatch resources to events were discussed." All of these items were measured on a 5-point scale, ranging from 1= "strongly disagree" to 5= "strongly agree." The adapted reflexivity scale had a reliability of .94 in the current study. Finally, I reviewed the reflexivity chat logs to confirm that the teams stayed on task and engaged in a reflexive discussion as they were instructed.

Affective responses. Because I was specifically interested in the cognitive mechanisms underpinning any relationships found between the team interventions and team outcomes, it was necessary to rule out other possible explanations, such as affective changes. Therefore, two items measuring team members' evaluation of bonding with their teammates, as well as the PANAS scale (Positive Affect α =.92, Negative Affect α=.81), were administered after the team interventions.

For a complete list of all the manipulation check items, please refer to Appendix K.

Covariates

As previous experience has been found to affect TMMs (e.g., Rentsch & Klimoski, 2001; Smith-Jentsch et al., 2001), I controlled for video game experience as well as knowledge of and experience in emergency/hospital response settings. Specifically, I wanted to ensure that team outcomes were not due to experienced gamers learning the simulation more quickly, or an unfair advantage for participants with greater insight into how emergency events should be solved. These variables were assessed in the demographics survey given at the start of the experiment (e.g., "Please indicate the amount of time you have worked in an emergency response setting"). In addition, when analyzing any relationships with team performance on Scenario 2 as the outcome variable, team performance on Scenario 1 was entered as a control variable.

Results

Manipulation Checks

Storytelling. To assess whether participants answered the storytelling manipulation check questions correctly, the frequencies of responses were examined by condition. The results

indicated that the storytelling manipulation check was effective at the individual and team levels. Overall, only two (.6%) participants incorrectly answered the question on whether they heard a story, and only sixteen (4.9%) incorrectly answered the question about what the story was about, with the majority of these being attributed to a typing mistake in the question that was fixed after the first few days of the experiment. Only one team had two members who answered the storytelling manipulation check questions incorrectly (no team had three incorrect answers), and that team was subsequently dropped from analyses.

Reflexivity. There were two types of questions used to check the reflexivity manipulation: categorical and scale. For the one categorical item asking what the discussion session was about, frequencies of each response option were examined by condition. This manipulation check was effective at both the individual and team levels, as only 7 (2.2%) participants answered the question incorrectly, and no teams had two or more members who answered this reflexivity manipulation check question incorrectly.

Although the reflexivity scale items were measured at the individual level, they contained a team referent, so it was necessary to determine whether the results could be aggregated to the team level. To make this determination, intraclass correlation coefficient (ICC) analyses were conducted to confirm that there was sufficient within-group agreement in the ratings and that an adequate amount of individual-level variance could be explained at the group-level (Bliese, 2000). ICC analyses revealed that 79% of the variability in participants' reflexivity scores was explained by team membership (ICC(1)=.79). Furthermore, the group means on this scale had reliable internal consistency (ICC(2)=.92). Therefore, there was justification for aggregating the reflexivity scale scores to the team level.

After aggregating the reflexivity scales to the team level, an independent samples t-test was performed on the reflexivity scale to ensure that team members in the reflexivity condition had a greater opportunity to collectively review their performance and set strategies for improvement. The results revealed that there were significant differences between reflexivity and non-reflexivity participants in their reflexivity scale scores, with those in the reflexivity (M=4.04, SD=.40) group rating higher on the Reflexivity scale than those in the control group (M=1.76, SD=.47; $t(105) = -26.91$, $p < .01$; see Table 1) .

Finally, only four teams were dropped due to failure to follow directions/stay on task during the reflexivity discussion, as assessed through reading the chat logs from the reflexivity conditions.

Affective responses. As expected, an independent samples t-test revealed that there were no significant differences in team member bonding or positive/negative affect across conditions (see Table 2 for corresponding t values). Because the interventions were not found to have differential effects on affective outcomes, affect could not explain any relations between the interventions and team performance.

Correlations

Means, standard deviations, and correlations of the study variables at the team level are presented in Table 3. There were no significant correlations between the covariates of interest and team outcomes aside from knowledge of emergency hospital protocols with TMM similarity ($r= -.26$, $p<.01$) and with team performance ($r= -.29$, $p<.01$). Interestingly, this means that the more such emergency knowledge team members had, the less similar their TMMs and the worse their performance. Although developed to be as realistic as possible, simulation events were also scripted to be relevant to undergraduate students. As such, NeoCITIES scenarios may have

differed from hospital procedures, thereby resulting in negative transfer for those who were more familiar with emergency protocols in that environment.

Despite these being the only significant correlations among the original control variables, I retained them all in my analyses to provide a more conservative test of the hypotheses. In addition, gender composition (measured by using a proportion of women index, which is currently the least susceptible to Type II errors; Williams & Mean, 2004) had a significant relationship with team performance ($r= -.32$, $p< .01$), with groups with more males performing better than groups with more females. Since gender composition was not of primary interest in this study, it was also included as a covariate.

Hypothesis Testing

Hypotheses were tested at the team-level using hierarchical regression. The reflexivity and storytelling conditions were dummy coded and centered. Mediation was tested using MacKinnon, Lockwood, Hoffman, West, and Sheets (2002) joint significance test for mediation. Results of all hypothesis tests can be seen in Tables 4 through 6.

Hypothesis 1 predicted that those who received the storytelling manipulation would perform better in the NeoCITIES simulation. However, as indicated in Table 4, there was no significant difference between those in the storytelling condition and those in the control condition ($\beta=.003$, ns).

Hypothesis 2 predicted that those in the storytelling condition would have more similar TMMs than those in the control condition. This was supported by the data, with storytelling leading to greater Performance 2 concept map similarity ($r=.24$, $p<.05$; $\beta=.25$, $p<.05$; $\Delta R^2=.06$) after controlling for demographic variables (see Table 5). Therefore, teams who received a story demonstrating the importance of collaboration and timing were more similar in their view on

how events should be solved than teams who were given the same instructions in non-story form. Furthermore, by regressing Performance 2 scores onto Performance 2 concept map similarity (after controlling for demographics and Performance 1 scores), I found support for a positive effect of TMM similarity on team performance ($r=.26$, $p<.05$; $\beta=.21$, $p<.05$; $\Delta R^2=.04$; refer to Table 6). Thus, based on the joint significance test for mediation described by MacKinnon and colleagues (2002), the latter two findings confirm that TMM similarity mediated the effect of storytelling on team performance. As such, Hypothesis 3 was supported.

Moving on to a similar set of hypothesis tests using the reflexivity intervention, I first regressed Performance 2 scores on reflexivity, using the same control variables, and found no significant results ($\beta= -.11$, ns). Thus, the presence or absence of guided team reflexivity did not directly impact team performance, contrary to the prediction of Hypothesis 4. Likewise, there was no significant effect of reflexivity on TMM, such that after controlling for the demographic variables described previously, those who were in the reflexivity condition had the same levels of concept map similarity as those in the control condition ($\beta=.00$, ns). Because Hypothesis 5 was not supported, this precluded any test of mediation as predicted in Hypothesis 6.

Hypothesis 7a posited that the presence of guided team reflexivity and storytelling together would have the greatest positive impact on team performance. I did not find support for this relationship ($\beta= -.08$, ns; refer to Table 4). However, in support of Hypothesis 7b, I did find support for a moderated effect on TMM similarity (see Table 5). That is, the interaction of reflexivity and storytelling significantly predicted team members' concept map similarity during Performance 2, beyond storytelling or reflexivity alone ($\beta=.19$, $p< .05$; $\Delta R^2=.04$). Figure 5 displays the nature of the interaction.

The pattern of the interaction suggests that when participants received a story and were able to subsequently engage in guided team reflexivity, they had the most similar mental models, and when they did not receive a story but did have guided team reflexivity, they had the least similar mental models. Those who had no reflexivity session, regardless of storytelling condition, had similar levels of TMM similarity, which were a little worse than having storytelling with reflexivity and a little better than having no storytelling with reflexivity. Formal post-hoc contrast tests revealed a significant difference between teams who received a story and were able to have a discussion (M=.47, SD=.15) and teams who did not receive a story but were still allowed to have a discussion (M=.37, SD=.15; $t(50)= -2.35$, $p< .05$). Although no other post-hoc contrasts were significant, perhaps due to power limitations, the pattern of the interaction suggests that receiving both interventions was better than receiving neither or just the story, which in turn were better than just participating in guided team reflexivity with no story.

While there was no evidence of a direct effect of the storytelling-reflexivity interaction on team performance, the significant effect on TMM similarity, along with the significant positive relationship of TMM similarity with team performance (as described for Hypothesis 3), suggests a mediated moderation relationship (MacKinnon et al., 2002). That is, the interaction of storytelling and reflexivity appears to have exerted its influence on team performance via TMM similarity.

Ancillary Analyses

Although reflexivity was found to be useful in conjunction with storytelling, additional analyses were conducted to help determine why it did not independently predict TMM similarity or team performance. It is interesting to note that the reflexivity condition did have an immediate significant effect on perceived helpfulness ($t(105)= -13.65$, $p<.01$). That is, those is

the reflexivity condition believed that their discussion session would be more helpful for improving their performance in the next scenario (M=3.77, SD=.61) than those in the control group (M=1.99, SD=.75). Furthermore, after completing the final performance scenario, team members in the reflexivity condition indicated that they had indeed thought about their discussion while playing and that they still believed it had helped them to perform better (M=4.01, SD=.51), all to a greater extent than those in the control condition (M=2.56, SD=.70; $t(105)=-12.28$, $p<.01$).

In addition to garnering favorable reactions from team members, the reflexivity intervention also had a significant effect on whether teams believed they came up with certain strategies for improving their performance. Specifically, those in the reflexivity condition indicated that they more often strategized to communicate, share briefing information, dispatch in the correct order, and meet deadlines more often than those in the control group (see Table 7 for means and t values). However, it is crucial to note that there were no significant differences across conditions with respect to the extent which teams actually implemented these strategies (see Table 7). In other words, although team members in the reflexivity condition reported coming up with specific strategies useful for improving performance, they did not rate themselves as using these strategies to a greater extent than those in the control condition. This could potentially explain why these participants rated the reflexivity session as helpful but did not exhibit significantly different team outcomes. All of the survey items used for these ancillary analyses can be found in Appendix L.

As previously mentioned, research on TMM accuracy has produced more mixed results for team performance than TMM similarity (e.g., Marks et al., 2000; Mathieu et al., 2005; Webber et al., 2000). In addition to ancillary analyses on reactions to the reflexivity

intervention, I also chose to investigate whether concept map accuracy had any relationship with the interventions or team performance. Accuracy was measured by looking at each box in the concept map from each team member (27 boxes total), and assigning a value of 0 if the answer was incorrect and a value of 1 if the answer was correct. This involved no subjective judgment, as the correct answer was always indisputably Police, Fire, or Hazmat. These values were averaged to create a final TMM accuracy score.

Interestingly, TMM accuracy was not significantly predicted by storytelling or guided team reflexivity (β= -.05, -.07, ns; see Table 8). However, it did predict team performance (β=.24, $p<.05$; see Table 9). Thus, although in this case TMM accuracy did prove to be related to team performance, its antecedents did not include the interventions of interest, thus supporting my decision to investigate TMM similarity instead.

Discussion

While there has been strong support for the effects of TMM similarity on team performance outcomes (DeChurch & Mesmer-Magnus, 2010a, 2010b), there has not been equal focus on the antecedents of TMMs (Mohammed et al., 2010). Likewise, the effects of storytelling and guided team reflexivity as formal team interventions have not been well-studied in an empirical context. Furthermore, all of these issues have been studied even less in virtual teams, where it is especially important for teams to develop interventions to facilitate shared understanding in the absence of media-rich cues. Therefore, this study investigated the effects of storytelling and guided team reflexivity on TMM similarity, as well as the subsequent effect of TMM similarity on team performance, all within a distributed team context.

Storytelling was found to have a positive effect on TMM similarity. That is, team members who learned the importance of collaboration and timing for successfully playing NeoCITIES via a story had more similar views on how to solve interdependent events than team members who were taught the same lessons through a more straightforward, non-story approach. In fact, even though both methods contained the same underlying message/deep structure, those in the storytelling condition rated the video as more helpful ($t(314)=4.36$, $p<.01$). Thus, I found support for Denning's (2001) proposition that a story allows someone to better relate to and internalize a message. In addition, although there was no support for a direct relationship between storytelling and team performance, the positive effect of storytelling on TMM similarity, and the subsequent effect of TMM similarity on team performance, suggests a mediated relationship. That is, storytelling had an indirect effect on performance via TMM similarity.

It is noteworthy that although storytelling was found to positively impact team outcomes, the manipulation utilized a very passive method of learning. That is, instead of participants actively being engaged in the act of storytelling itself, they were the passive recipients of a video with an unseen narrator's voice. While I attempted to increase the credibility of the story by prefacing it with a statement that past participants had found the story helpful for improving their performance, this may not have adequately substituted for the dynamism and fidelity of a live storytelling experience. Yet despite this passive and potentially weaker storytelling manipulation, a significant effect on TMM similarity was still found. Thus, the effect found may actually be a conservative estimate of the true potential of storytelling interventions. Future studies could investigate how different types of storytelling interventions could enhance these positive results.

By itself, guided team reflexivity did not have a significant relationship with TMM similarity or team performance in this study. That is, simply being given the opportunity to discuss their performance and strategize for the future did not improve team outcomes. However, this does not mean that reflexivity is unimportant. First, the content of team members' strategies must be taken into consideration. Every team given the opportunity to engage in reflexivity did not necessarily come up with the same quality of strategies. Some strategies may have been useful, while others may have been too general (e.g., we should communicate more) or too complex (e.g., *every* time someone dispatches to an event he/she should tell the team). Although not assessed in this study, it may be that the quality of the ideas generated during the reflexivity intervention does have a relationship with team outcomes.

A second indication that reflexivity may be important despite this study's findings is that those who had the opportunity to engage in guided team reflexivity had positive reactions, rating it as useful for improving subsequent performance in the simulation. The problem may lie in that although reflexivity participants indicated they came up with specific performance improvement strategies to a greater extent than non-reflexivity participants, they did not actually use these strategies to a greater extent. Therefore, perhaps additional measures should be taken to ensure that team members are able to put their plans into action. In other words, this study's reflexivity intervention allowed strategy generation, but did not facilitate strategy implementation.

The third, and perhaps most telling, indication that reflexivity is indeed influential on team outcomes is that analyses suggest that a planned reflexivity session actually enhanced the positive effects of storytelling, as indicated by their significant interaction on TMM similarity. In other words, storytelling may have provided the content, but reflexivity provided the

mechanism by which the content could be more fully processed. However, having a discussion to create strategies and "get on the same page" was detrimental if teammates did not have common knowledge on which to base their discussion. Specifically, they needed a clear understanding of the importance of timing and collaboration, as emphasized in the story, to have a productive discussion session that would improve, rather than hinder, team outcomes. In fact, if a team was not exposed to a storytelling intervention, it was better if there was no reflexivity session either. Based on the patterns indicated by the interaction, while a more promising option would be to have both storytelling and guided team reflexivity, in the absence of storytelling, guided team reflexivity should not be used.

One final noteworthy point about this study's results is that the positive relationship of the storytelling-reflexivity interaction with TMM similarity, and the subsequent positive relationship between TMM similarity and team performance, suggests a mediated moderation relationship. Thus, this study supports the notion that storytelling paired with reflexivity will maximize the potential for favorable team performance via TMM similarity.

Limitations and Future Studies

As a 2.5 hour experiment, I only examined short-term learning effects that could be demonstrated minutes after an intervention was introduced. While I did find that storytelling, in combination with reflexivity, can affect TMM similarity, it is unclear whether this effect is long-lasting. Future studies should therefore investigate the longitudinal impact of these interventions on TMM similarity and subsequent team performance.

As mentioned previously, this study utilized a fairly passive method of storytelling, but still found significant results. It would be interesting to utilize different types of storytelling

interventions in future studies to determine how they may enhance the positive effects on team outcomes found here, both directly and through their interaction with guided team reflexivity. Some possibilities could involve participants sharing their own applicable experiences with their teammates, or having a confederate share a pre-scripted story at a pre-determined time. This could potentially lead to greater affective commitment and a better internalization of the story's lessons, which could in turn enhance the positive effects on TMM similarity and subsequent performance, and perhaps even enhance or alter the role of reflexivity.

Finally, due to the nature of the TMM that was used in this study (concept mapping), it was impossible to administer the measure at an optimal time to make conclusions about mediation. That is, the concept map measure ideally should have been administered after the storytelling and/or reflexivity intervention, but prior to the second performance scenario. However, because the contents of the concept map were inextricably entwined with the content of Scenario 2, it was necessary to complete the measure after the performance task was completed. Other studies have also measured team cognition after the occurrence of team outcomes (e.g., Ellis, 2006; Gurtner et al., 2007; Mohammed & Ringseis, 2001). Rather than interrupting the performance task to administer team cognition measures, which could potentially have artificial negative effects on performance, I followed the example of such prior studies and chose to administer the TMM measure immediately following the second performance session. Therefore, even though the results support mediation, the placement of the measures does not allow me to make firm statements about causality.

Conclusion

This study ventured into largely unexplored territory: the effect of planned storytelling

and reflexivity interventions on TMMs and team performance. In addition to possibly being the first empirical study of the effects of storytelling on team performance, it has expanded limited empirical research on guided team reflexivity, and it has also explored new antecedents of TMMs, all within the context of virtual teams. The results provide encouraging evidence that these interventions may help overcome the collaborative obstacles faced by team members in distributed environments, particularly when used in tandem. This study will hopefully provide a stepping-stone for further exploration of interventions that can enhance TMM similarity and subsequent team performance.

References

Abrahamson, C. E. (2006). Motivating student through personal connections: Storytelling as pedagogy in Introductory Psychology. In D. S. Dunn & S. L. Chew (Eds.), *Best practices for teaching introduction to psychology* (pp. 245-258). Mahwah, New Jersey: Erlbaum.

Andres, H. P. (2002). A comparison of face-to-face and virtual software development teams. *Team Performance Management, 8,* 39–48.

Ang, C. S., & Rao, G. S. V. R. K. (2008). Computer game theories for designing motivating educational software: A survey study. *International Journal on E-Learning, 7*(2), 181-199.

Argote, L., Gruenfeld, D., & Naquin, C. 1999. Group learning in organizations. In M. E. Turner (Eds.), *Groups at work: Advances in theory and research* (pp. 369-412). Mahwah, NJ: Lawrence Erlbaum.

Argyris, C. (1993). *On organizational learning*. Cambridge, MA, US: Blackwell Business/Blackwell Publishers.

Bal, M. (1997). *Narratology: Introduction to the theory of narrative.* Toronto: University of Toronto Press.

Barge, K. J. (2004). Reflexivity and managerial practice. *Communication Monographs, 71,* 70–96.

Barry, C. A., Britten, N., Barber, N., Bradley, C., & Stevenson, F. (1999). Using reflexivity to optimize teamwork in qualitative research. *Qualitative Health Research, 9*(1), 26-44.

Bartel, C. A., & Garud, R. (2009). The role of narratives in sustaining organizational innovation. *Organization Science, 20*(1), 107-117.

Bell, B. S., & Kozlowski, S. J. (2002b). A typology of virtual teams: Implications for effective leadership. *Group & Organization Management, 27*(1), 14-49.

Bhappu, A. D., Zellmer-Bruhn, M., & Anand, V. (2001). The effects of demographic diversity and virtual work environments on knowledge processing in teams. *Advances in the Interdisciplinary Study of Work Teams, 8,* 149–165.

Blickensderfer, E. L., Cannon-Bowers, J. A., & Salas, E. (2007). Theoretical bases for team self-correction: Fostering shared mental models. In M. Beyerlein, D. Johnson, & S. Beyerlein (Eds.), *Advances in interdisciplinary studies in work team series* (Vol. 4, pp. 249-279). Greenwich, CT: JAI Press.

Blickensderfer, E., Cannon-Bowers, J. A., & Salas, E. (1997a). *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.

Blickensderfer, E., Cannon-Bowers, J. A., & Salas, E. (1998). Cross-training and team performance. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 299-311). Washington, DC: American Psychological Association.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349-381). San Francisco: Jossey-Bass.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*, 1–21.

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

Carter, S., & West, M. A. (1998). Reflexivity, effectiveness, and mental health in BBC-TV production teams. *Small Group Research, 29,* 583–601.

Carter, K. (1993). The place of story in the study of teaching and teacher education. *Educational Researcher, 22*(1), 5-12,18

Cascio, W. F., & Aguinis, H. (2008). Research in industrial and organizational psychology from 1963 to 2007: Changes, choices and trends. *Journal of Applied Psychology, 93*(5), 1062-1081.

Coovert, M. D., Craiger, J. P., and Cannon-Bowers, J. A. (1995). Innovations in Modeling and Simulating Team Performance: Implications for Decision Making. In R. A. Guzzo & E. Salsa (Eds.), *Team Effectiveness and Decision Making in Organizations* (pp. 149-203). Jossey-Bass Publishers: San Francisco.

De Dreu, C. K. W. (2007). Cooperative outcome interdependence, task reflexivity, and team effectiveness: A motivated information processing perspective. *Journal of Applied Psychology, 92*(3), 628-638.

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

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

Denning, S. (2001). The springboard: How storytelling ignites action in knowledge-era organizations. Boston: Butterworth Heinemann.

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*(6), 678-711.

Doyle, W., & Carter, K. (2003). Narrative and learning to teach: Implications for teacher-education curriculum. *Journal of Curriculum Studies, 35*(2), 129-137.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly, 44*, 350-383.

Edmondson, A. C. (2002). The local and variegated nature of learning in organizations. *Organization Science, 13*(2), 128–146.

Edmondson, A. C., Dillon, J. R., & Roloff, K. S. (2007). Three perspectives on team learning: Outcome improvement, task mastery, and group process. *The Academy of Management Annals, 1*, 269-314.

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

Fiore, S. M., & McDaniel, R. (2006). Building Bridges: Connecting Virtual Teams Using Narrative and Technology. *THEN: Journal of Technology, Humanities, Education and Narrative*. Published Online January 02, 2006 http://thenjournal.org/commentary/95/. Lead Article Inviting Commentary.

Fiore, S. M., Johnston, J., & McDaniel, R. (2007). Narratology and Distributed Training: Using the Narrative form for Debriefing Distributed Simulation-based Exercises. In S. M. Fiore and E. Salas (Eds.), *Towards a science of distributed learning* (pp. 119 – 145). Washington, DC: American Psychological Association.

Fiore, S. M., McDaniel, R., & Jentsch, F. (2009). Narrative-Based Collaboration Systems for Distributed Teams: Nine Research Questions for Information Managers. *Information Systems Management, 26*(1), 28 – 38.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

Gentner, D., & Landers, R. (1985). Analogical reminding: A good match is hard to find. Proceedings *of the International Conference on Systems, Man and Cybernetics* (pp. 607-613). New York: IEEE.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive*

Psychology, 15, 1-38.

Goldstein, I. L., & Ford, J. K. (2001). *Training in organizations: Needs assessment, development, and evaluation* (4th ed.). Belmont, CA: Wadsworth.

Graetz, K. A., Boyle, E., Kimble, C., Thompson, P., & Garloch, J. (1998). Information sharing in face-to-face, teleconferencing and electronic chat groups. *Small Group Research, 29,* 714–743.

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

Hackman, J. R., & Wageman, R. (2005). A theory of team coaching. *Academy of Management Review,* 30(2), 269-287.

Hackman, J. R., Brousseau, K. R., & Weiss, J. A. (1976). The interaction of task design and group performance strategies in determining group effectiveness. *Organizational Behavior and Human Performance, 16,* 350–365.

Hellar, B. (2009). *An investigation of data overload in team-based distributed cognition systems.* Unpublished doctoral dissertation, The Pennsylvania State University, University Park, PA.

Hiltz, S. R., Johnson, K., & Turoff, M. (1986). Experiments in group decision making: Communication process and outcome in face-to-face versus computerized conferences. *Human Communication Research, 13,* 225–252.

Jarvenpaa, S. L., Rao,V. S.,& Huber,G. P. (1988). Computer support for meetings of groups working on unstructured problems: A field experiment. *MIS Quarterly, 12,* 645–666.

Kanawattanachai, P., & Yoo, Y. (2002). Dynamic nature of trust in virtual teams. *Journal of Strategic Information Systems, 11,* 187–213.

Klein, G. (1998). *Sources of power.* Cambridge, MA: MIT Press.

Kozlowski, S.W.J., & Ilgen, D. R. 2006. Enhancing the effectiveness of work groups and teams (monograph). *Psychological Science in the Public Interest, 7,* 77-124.

Lawler, E. E., Mohrman, S. A. and Ledford, G. E. (1992). *Employee involvement and total quality management: Practices and results in fortune 1000 companies*. San Francisco: Jossey-Bass.

Lawler, E. E., Mohrman, S. A., and Ledford, G. E. (1995). *Creating high performance organizations: Practices and results of employee involvement and quality management in fortune 1000 companies.* San Francisco: Jossey-Bass.

Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology, 88,* 587-604.

Manz, C. C., & Sims, H. P. (1987). Leading workers to lead themselves: The external leadership of self-managing work teams. *Administrative Science Quarterly, 32*(1), 106-129.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review, 26*(3), 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*(1), 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*(6), 971–986.

Martins, L. L., Gilson, L. L., & Maynard, M. T. (2004). Virtual teams: What do we know and where do we go from here? *Journal of Management, 30*(6), 805-835.

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*(2), 273–283.

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

McNeese, M. D., Bains, P., Brewer, I., Brown, C., Connors, E. S., Jefferson, T., et al. (2005). The NeoCITIES simulation: Understanding the design and experimental methodology used to develop a team emergency management simulation. *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society,* 591-594.

Mohammed, S., & Dumville, B. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior, 22,* 89-106.

Mohammed, S., Ferzandi, L., and Hamilton, K. (2010). Metaphor no more: A 15-year Review of the team mental model construct. *Journal of Management, 36*(4), 876-910.

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.

Pachler, N., & Daly, C. (2009). Narrative and learning with web 2.0 technologies: Towards a research agenda. Journal of Computer Assisted Learning, *25*(1), 6-18.

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*(4), 532-555.

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

Rosen, M. A., Fiore, S.M., McDaniel, R., & Salas, E. (in press). Adapting Narrative Theory to Improve the Implementation of Story in Problem-based Learning. In D. H. Andrews & T. D. Hull (Eds.), *Storytelling as an Instructional Method: Research Perspectives*. Rotterdam: Sense Publishers.

Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 456-468.

Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions on Systems, Man, & Cybernetics, 22*, 1296-1308.

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

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). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Salas, E., Nichols, D. R., & Driskell, J. E. (2007). Testing three team training strategies in intact teams: A meta-analysis. *Small Group Research, 38*(4), 471-488.

Schippers, M. C., Den Hartog, D., Koopman, P. L., & Wienk, J. A. (2003). Diversity and team outcomes: The moderating effects of outcome interdependence and group longevity and the mediating effect of reflexivity. *Journal of Organizational Behavior, 24,* 779–802.

Schippers, M. C., Den Hartog, D. N., & Koopman, P. L. (2007). Reflexivity in teams: A measure and correlates. *Applied Psychology: An International Review*, *56*(2), 189-211.

Schippers, M. C., Den Hartog, D. N., Koopman, P. L., & van Knippenberg, D. (2008). The role of transformational leadership in enhancing team reflexivity. Human Relations, 61(11), 1593-1616.

Schmidt, J. B., Montoya-Weiss, M. M., & Massey, A. P. (2001). New product development decision-making effectiveness: Comparing individuals, face-to-face teams, and virtual teams. *Decision Sciences, 32,* 575–600.

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

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

Smith-Jentsch, K. A., Mathieu, J. E., & Kraiger, K. (2005). Investigating linear and interactive effects of shared mental models and safety and efficiency in a field setting. *Journal of Applied Psychology, 90*(3), 532–535.

Smith-Jentsch, K. A., Zeisig, R. L., Acton, B., & McPherson, J. A. (1998). Team Dimensional Training: A strategy for guided team self-correction. In J.A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 271-297). Washington, DC: American Psychological Association.

Townsend, A. M., DeMarie, S. M., & Hendrickson, A. R. (1998). Virtual teams: Technology and the workplace of the future. *Academy of Management Executive, 12*(3), 17-29.

U.S. Department of Transportation. (2003). *Transportation Implications of Telecommuting.* Retrieved November 8, 2009, from http://ntl.bts.gov/DOCS/telecommute.html

Van der Vegt, G. S., & Bunderson, J. S. (2005). Learning and performance in multidisciplinary teams: The importance of collective team identification. *Academy of Management Journal, 48*(3), 532–547.

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.

Weingart, L. R. (1992). Impact of group goals, task component complexity, effort, and planning on group performance. *Journal of Applied Psychology, 77*(5), 682–693.

Wellens, A. R., & Ergener, D. (1988). The C.I.T.I.E.S. game: A computer-based situation assessment task for studying distributed decision making. *Simulation & Games, 19*, 304-327.

West, M. A. (1996). Reflexivity and work group effectiveness: A conceptual integration. In M. A. West (Ed.), *Handbook of work group psychology* (pp. 555–579). Chichester, England: Wiley.

Williams, H. M., & Mean, L. J. (2004). Measuring gender composition in work groups: A comparison of existing methods. *Organizational Research Methods, 7,* 456-475.

Wilson, J. M., Goodman, P. S., & Cronin, M. A. (2007). Group learning. *Academy of Management Review, 32*(4), 1041-1059.

Appendix A

Tables

Table 1

Summary of Independent Samples T-tests on Reflexivity Manipulation Check

Manipulation Check Item	M	SD	df	t	p
Reflexivity Scale			105	-26.91	.00
Reflexivity group	4.04	.40			
Control group	1.76	.47			

Table 2

Summary of Independent Samples T-tests on Affective Response Manipulation Check Items

Manipulation Check Item	M	SD	df	t	p
Affective Response					
This group discussion session allowed me to bond with my teammates.			105	-1.17	.24
Reflexivity group	3.58	.60			
Control group	3.43	.74			
Positive Affect Scale					
			105	-1.04	.30
Reflexivity group	24.72	5.16			
Control group	23.68	5.20			
			105	1.12	.27
Storytelling group	23.66	5.61			
Control group	24.78	4.70			
Negative Affect Scale					
			105	-.49	.63
Reflexivity group	13.86	2.61			
Control group	13.62	2.36			
			105	.29	.77
Storytelling group	13.67	2.11			
Control group	13.81	2.82			

Table 3

Descriptive Statistics and Correlations of Study Variables

	Variable	Mean	SD	1	2	3	4	5	6	7	8
1.	Gender [a]	1.49	0.33								
2.	Average hrs/wk playing video games	10.77	12.60	-.30**							
3.	Experience in emergency response settings	0.13	0.45	.09	.05						
4.	Knowledge of hospital emergency response protocols	1.59	0.50	.27**	-.16	.12					
5.	Reflexivity	.05[b]	1.00	.01	.07	.00	.04				
6.	Storytelling	-.01[b]	1.00	.07	.03	.12	.04	.01			
7.	Time 2 concept map similarity	0.42	0.13	.02	-.09	.21	-.26**	-.03	.24*		
8.	Time 1 number of interdependent events completed	0.77	0.14	-.16	.04	.01	-.10	-.07	.04	.193	
9.	Time 2 number of interdependent events completed	0.45	0.24	-.32**	.15	-.19	-.29**	-.11	-.01	.26*	.27**

[a] Composition of gender within the team, ranging from 1 (all males) to 2 (all females).
[b] Dummy-coded variable: control= -1; reflexivity/storytelling= 1.
* p <.05
** p < .01

Table 4

Summary of Hierarchical Regression Analysis for Storytelling and Reflexivity Predicting Team Performance in Scenario 2

Variable	B	SE	β	ΔR^2
Step 1				.22**
Gender	-.15	.07	-.21*	
Average time playing video games (hrs/wk)	.00	.00	.06	
Experience in emergency response settings	-.07	.05	-.14	
Knowledge of emergency hospital response protocols/procedures	-.09	.05	-.20	
Team Performance in Scenario 1	.38	.16	.23*	
Step 2				.01
Gender	-.16	.07	-.21*	
Average time playing video games (hrs/wk)	.00	.00	.07	
Experience in emergency response settings	-.07	.05	-.15	
Knowledge of emergency hospital response protocols/procedures	-.09	.05	-.19	
Team Performance in Scenario 1	.37	.16	.22*	
Storytelling	.01	.02	.03	
Reflexivity	-.03	.02	-.11	
Step 3				.01
Gender	-.15	.078	-.21*	
Average time playing video games (hrs/wk)	.00	.00	.09	
Experience in emergency response settings	-.08	.05	-.15	
Knowledge of emergency hospital response protocols/procedures	-.09	.05	-.19	
Team Performance in Scenario 1	.39	.16	.23*	
Storytelling	.01	.02	.03	
Reflexivity	-.03	.02	-.11	
Storytelling x Reflexivity	-.02	.02	-.078	

N= 107 teams

*p<.05

**p<.01

Table 5

Summary of Hierarchical Regression Analysis for Predicting TMM Similarity

Variable	B	SE	β	ΔR^2
Step 1				.10*
Gender	.019	.041	.048	
Average time playing video games (hrs/wk)	-.001	.001	-.130	
Experience in emergency response settings	.040	.031	.127	
Knowledge of emergency hospital response protocols/procedures	-.083	.027	-.319**	
Step 2				.06*
Gender	.014	.040	.037	
Average time playing video games (hrs/wk)	-.001	.001	-.144	
Experience in emergency response settings	.028	.031	.089	
Knowledge of emergency hospital response protocols/procedures	-.083	.026	-.320**	
Storytelling manipulation	.032	.013	.246*	
Reflexivity manipulation	.000	.013	-.002	
Step 3				.04*
Gender	.009	.040	.023	
Average time playing video games (hrs/wk)	-.002	.001	-.180	
Experience in emergency response settings	.035	.031	.110	
Knowledge of emergency hospital response protocols/procedures	-.089	.026	-.344**	
Storytelling manipulation	.031	.013	.236*	
Reflexivity manipulation	.001	.012	.011	
Storytelling x Reflexivity	.025	.013	.193*	

N= 107 teams
*p<.05
**p<.01

Table 6

Summary of Hierarchical Regression Analysis for TMM Similarity Predicting Team Performance in Scenario 2

Variable	B	SE	β	ΔR^2
Step 1				.23**
Gender	-.17	.08	-.24*	
Average time playing video games (hrs/wk)	.00	.00	.06	
Experience in emergency response settings	-.06	.06	-.12	
Knowledge of emergency hospital response protocols/procedures	-.09	.06	-.19	
Number of completed team events in Scenario 1	.36	.16	.22*	
Step 2				.04*
Gender	-.19	.08	-.26*	
Average time playing video games (hrs/wk)	.00	.00	.08	
Experience in emergency response settings	-.08	.05	-.14	
Knowledge of hospital emergency protocols/procedures	-.06	.05	-.13	
Number of completed team events in Scenario 1	.29	.16	.17	
TMM Similarity (Concept Map)	.39	.19	.21*	

N= 107 teams
*p<.05
**p<.01

Table 7

Summary of Independent Samples T-tests on Ancillary Analysis Items

Manipulation Check Item	M	SD	df	t	p
Reflexivity Helpfulness- Immediately after intervention			105	-13.65	.00
Reflexivity group	3.77	.61			
Control group	1.99	.75			
Reflexivity Helpfulness- End of experiment			105	-12.28	.00
Reflexivity group	4.01	.51			
Control group	2.56	.70			
Strategy Generation					
Communication			105	20.81	.00
Reflexivity group	1.04	0.11			
Control group	1.78	0.24			
Share briefings			105	11.49	.00
Reflexivity group	1.32	0.31			
Control group	1.88	0.17			
Send units in correct order			105	18.38	.00
Reflexivity group	1.12	0.20			
Control group	1.85	0.21			
Arrive at events in time			105	12.04	.00
Reflexivity group	1.29	0.29			
Control group	1.86	0.18			
Strategy Implementation					
Communication			105	-1.95	.05
Reflexivity group	4.32	0.53			
Control group	4.10	0.63			
Share briefings			105	.79	.43
Reflexivity group	3.77	0.91			
Control group	3.90	0.80			
Send units in correct order			105	-.67	.50
Reflexivity group	3.77	0.69			
Control group	3.67	0.82			
Arrive at events in time			105	1.43	.16
Reflexivity group	3.39	0.68			
Control group	3.59	0.78			

Table 8

Summary of Hierarchical Regression Analysis for Predicting TMM Accuracy

Variable	B	SE	β	ΔR^2
Step 1				.05
Gender	-.06	.04	-.15	
Average time playing video games (hrs/wk)	.00	.00	-.03	
Experience in emergency response settings	.01	.03	.03	
Knowledge of emergency hospital response protocols/procedures	-.03	.03	-.13	
Step 2				.01
Gender	-.06	.04	-.15	
Average time playing video games (hrs/wk)	.00	.00	-.02	
Experience in emergency response settings	.01	.03	.04	
Knowledge of emergency hospital response protocols/procedures	-.03	.03	-.13	
Storytelling manipulation	-.01	.01	-.05	
Reflexivity manipulation	-.01	.01	-.07	

N= 107 teams
*p<.05
**p<.01

Table 9

Summary of Hierarchical Regression Analysis for TMM Accuracy Predicting Team Performance in Scenario 2

Variable	B	SE	β	ΔR^2
Step 1				.23**
Gender	-.17	.08	-.23*	
Average time playing video games (hrs/wk)	.00	.00	.06	
Experience in emergency response settings	-.06	.05	-.12	
Knowledge of emergency hospital response protocols/procedures	-.09	.05	-.19	
Number of completed team events in Scenario 1	.36	.16	.22*	
Step 2				.06*
Gender	-.15	.07	-.21*	
Average time playing video games (hrs/wk)	.00	.00	.06	
Experience in emergency response settings	-.07	.05	-.13	
Knowledge of hospital emergency protocols/procedures	-.07	.05	-.15	
Number of completed team events in Scenario 1	.33	.16	.20*	
TMM Accuracy (Concept Map)	.44	.17	.24*	

N= 107 teams
*p<.05
**p<.01

Appendix B

Figures

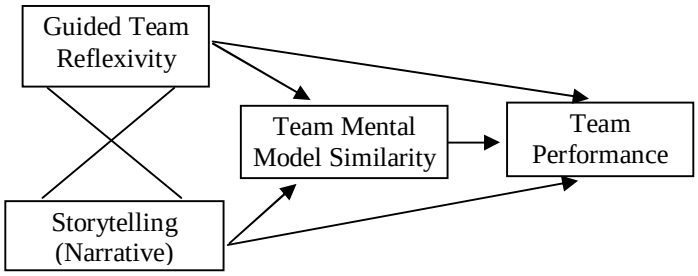


Figure 1. Summary of the research model.

		Reflexivity: Present	Reflexivity: None
Storytelling	Present	84 (28)	75 (25)
	None	84 (28)	78 (26)

Figure 2. The experiment design. The numbers in each cell indicate the number of participants (teams) in each condition.

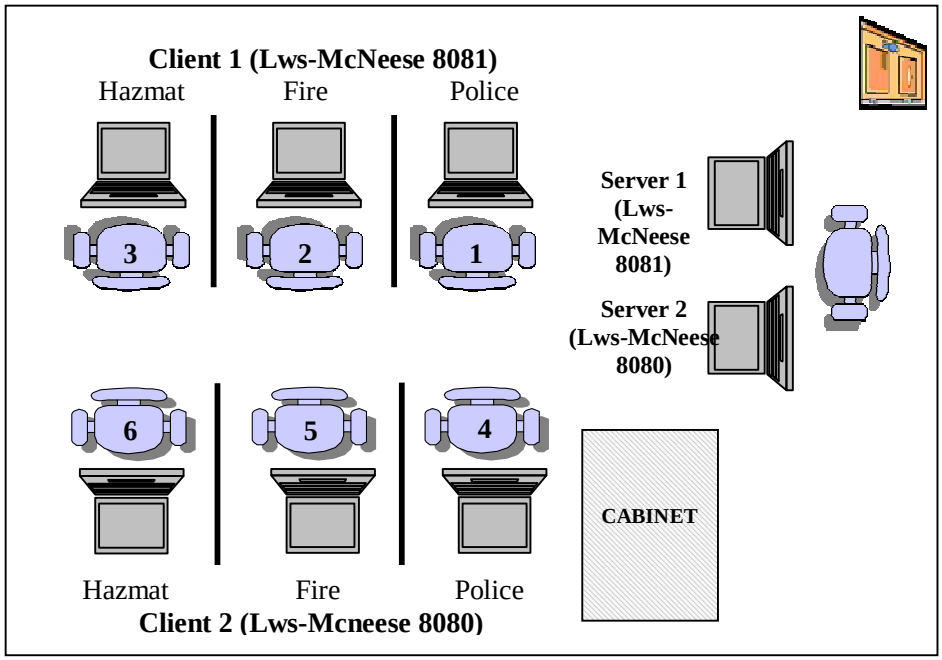


Figure 3. Diagram of the IST lab setup.

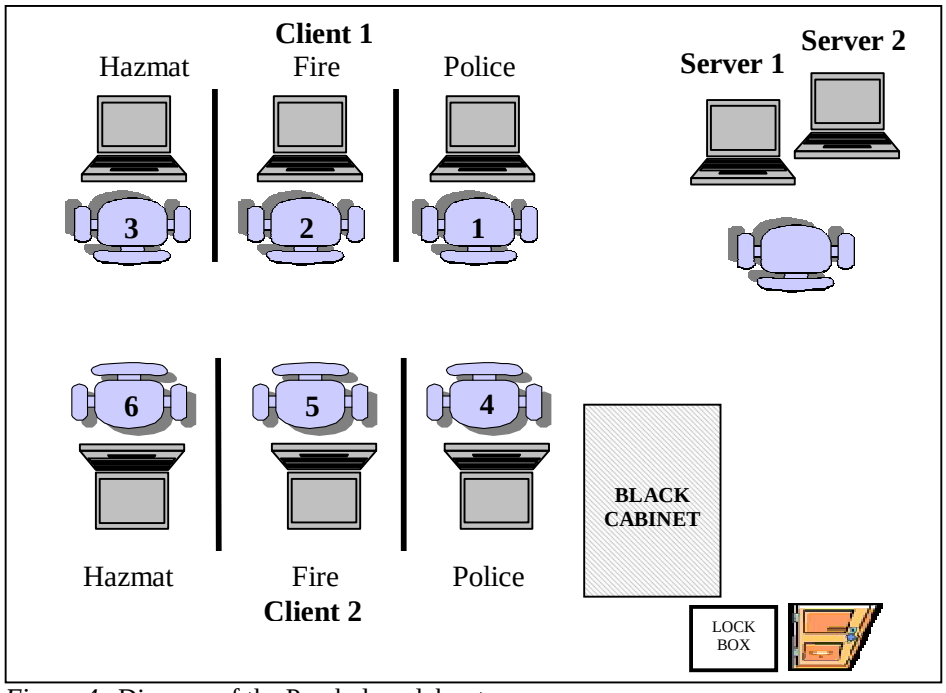


Figure 4. Diagram of the Psychology lab setup.

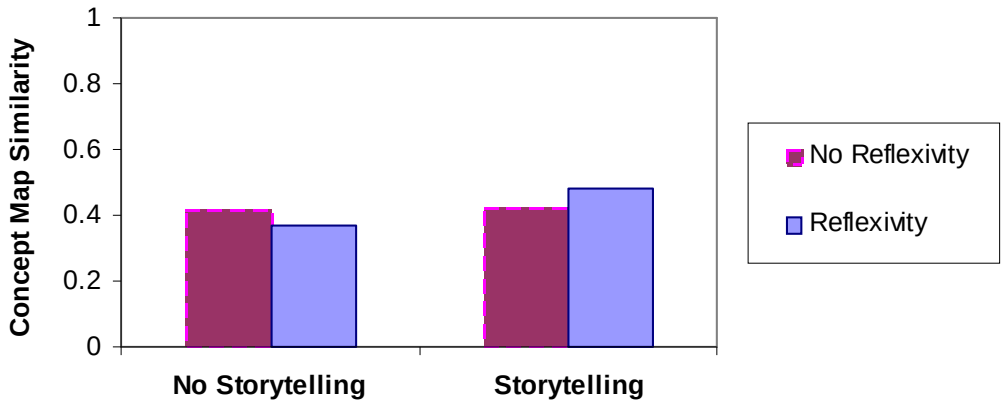


Figure 5. The interactive effect of storytelling and reflexivity on TMM similarity (concept mapping).

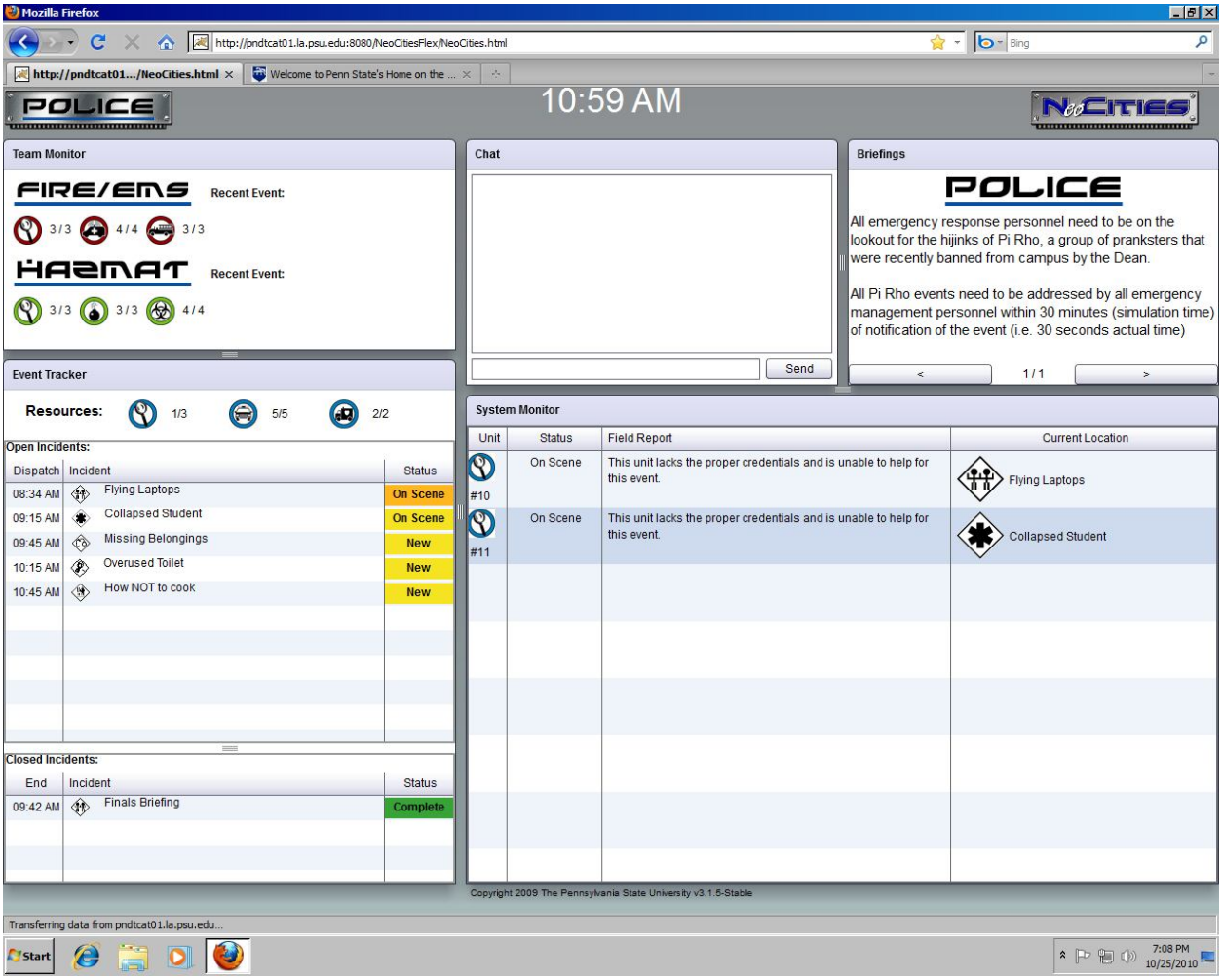
Appendix C

NeoCITIES Roles and Resources

POLICE		**FIRE**		**HAZMAT**	
Resource	# Avail.	Resource	# Avail.	Resource	# Avail.
Investigator	3	Investigator	3	Investigator	3
Squad Car	5	Fire Truck	4	Chemical Truck	3
SWAT Team	2	Ambulance	3	Bomb Squad	4

Appendix D

Screenshots of the NeoCITIES Simulation



Appendix E

Story Script

Please Note: The following is based on a true story. The names have been changed, but the facts are real.

Dan, a 22-year-old graduate student, was working late on an experiment in the Chemistry Lab while a severe snow storm raged outside. Just as he was carrying a beaker full of hydrochloric acid across the room, there was a power outage. In the dark, Dan tripped on his backpack and fell, spilling the contents of the beaker all over his left arm. Dan howled in pain, causing a student in the adjacent lab to come running over. The lights from the backup generator flickered on to reveal Dan lying on the floor unconscious with severe chemical burns on his left arm. The student immediately called 911, but since the beaker had shattered during the fall, the student could not tell the 911 operator what the specific substance was that had burned Dan's arm. The operator then notified all emergency response departments of the incident, which included Police, Fire, and Hazmat.

The operator also conveyed a written message to all of the units using a data terminal, a small wireless computer that can be accessed by each department. In this briefing, the operator described the incident and stated that both Hazmat and Fire were requested on scene. Hazmat was needed to identify the chemical before Fire could administer the appropriate treatment. The briefing also notified all units of the severe weather conditions. Several roads were impassable because of heavy snow build-up and debris. There were also many small accidents along the major roadways. According to the briefing, Police were to be on hand to clear the roads of traffic and to supervise snow removing crew. Therefore Police needed to be dispatched to the incident to clear the way for Hazmat and Fire to arrive on scene and treat Dan.

A few minutes went by and Hazmat noticed that Police had not responded to the incident. Their data terminal showed that no one was on scene. Due to the extent of Dan's injuries, it was critical to treat him quickly. Standard protocol dictates that chemical accidents need to be addressed within 10 minutes of notification. Because Police was busy responding to a series of fender benders they had not taken the time to read the briefing. They were therefore not aware that they needed to respond to the incident. Frustrated, Hazmat and Fire tried to save time and get to the Chemistry Building without the help of Police, but to no avail. Their units both ended up getting stuck in the snow.

With all other options exhausted, Hazmat and Fire contacted Police. When Police responded and arrived to clear the roads almost 45 minutes had passed since the original dispatch notification. After clearing the roads, it took an additional 15 minutes to tow the Hazmat and Fire units out of the snow. Once on scene, Hazmat was able to identify the chemical substance as hydrochloric acid, which then enabled Fire to correctly treat Dan's burns. However, since it was crucial for them to arrive within 10 minutes, their late arrival had led to tragic consequences. The tissue damage from the toxic chemicals had reached Dan's nerves causing his left arm to be completely paralyzed. Dan never fully recovered and had to live his life knowing that he would never again have any feeling or use in his left arm.

If Police had read the briefing and dispatched units to clear the roads, if Hazmat and Fire had waited to dispatch units in the proper order (Police, then Hazmat, then Fire), and if all of the units had arrived at the scene on time, this terrible tragedy would not have occurred. Hazmat and Fire would not have left first and gotten stuck in snow, Police would not have wasted time towing Hazmat and Fire out of the snow, and Dan would not have been paralyzed because the chemicals stayed too long on his arm.

Surface Structure Similarities:

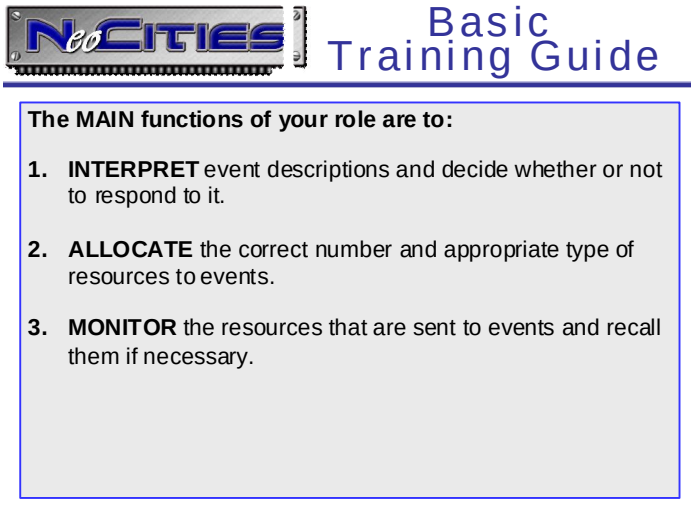
NeoCITIES	Story
1. Three distinct roles (Police, Fire, Hazmat).	Three emergency response units (Police, Fire, Hazmat).
2. Takes place on the Penn State campus.	Takes place in a Chemistry lab on a university campus.
3. Events involve university students.	Protagonist is a student.

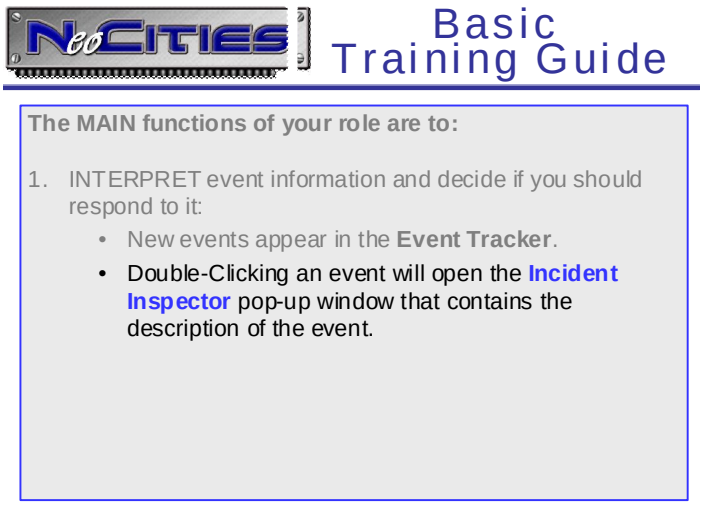
Deep Structure of Story:

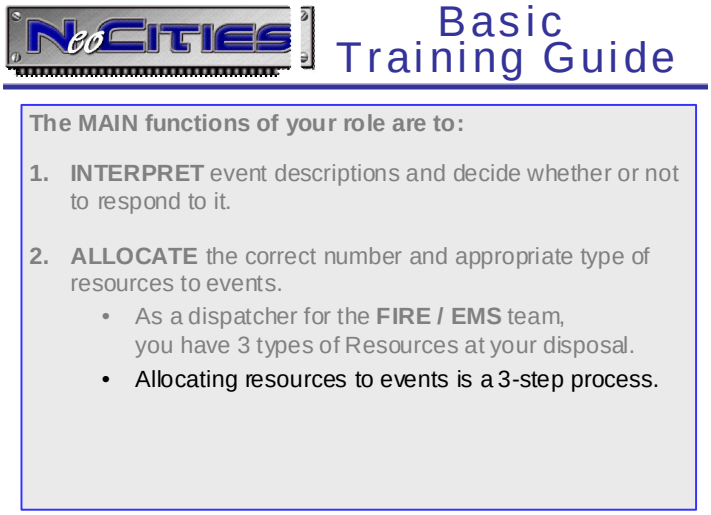
NeoCITIES	Story
1. Sequencing	Police had to clear the roads so that Hazmat could identify the chemical and then Fire could treat Dan.
2. Deadline	According to standard protocol, chemical accidents need to be addressed within 10 minutes.
3. Communication	Hazmat didn't tell Fire what the chemical involved in the accident was.
4. Interdependence	Fire couldn't treat the chemical burn no matter how hard they tried as long as they didn't have the correct chemical information from Hazmat.

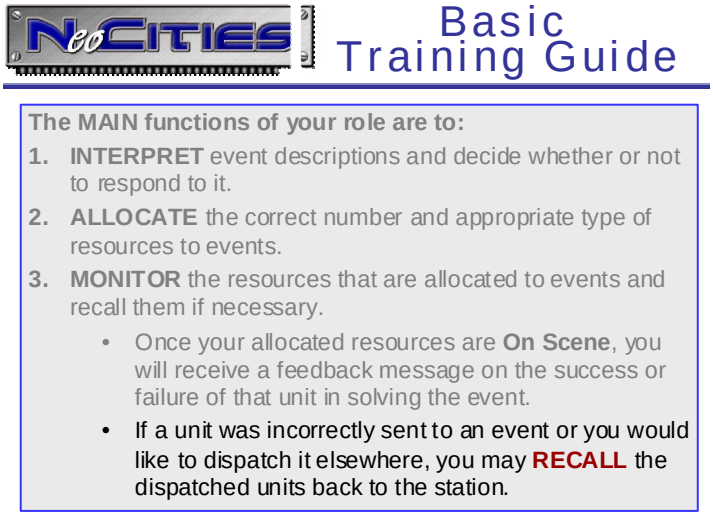
Appendix F

Sample Screenshots of the Basic Training Video









Appendix G

Sample Events from Performance Scenarios

Description	Team	Unit
Units needed for briefing on upcoming football weekend. Officials expect BOTH increased student rioting AND widespread injuries if we lose the game.	Police	Investigator
Send agents to Eisenhower to give training to campus security about bombs and bomb-making materials.	Hazmat	Investigator
Director requests disposal of a large number of barrels containing expired chemicals found in the basement of Beaver Stadium. Some containers may be volatile. Units are advised to proceed with caution.	Hazmat	Chemical Truck
A student called to report her friend vomiting and collapsing in the restroom at a local bar. Unit requested for treatment.	Fire	EMS
Employees report that an apartment has caught fire and spread to nearby store below. Units needed to suppress fire.	Fire	Fire Truck
Small group of students seen pouring strange blue and white chemicals on the Old Main lawn to try and create the Nittany Lion logo. Units need to arrive in the following order: FIRST to collect samples of the material and SECOND to interview possible witnesses.	Hazmat Police	Investigator Investigator
A tanker carrying aqueous ammonia has collided with a large truck. The driver thinks the tanker may explode within 60 seconds. Units are needed in the following order: FIRST to clear the area of on-lookers, SECOND to control the flames, and THIRD to cleanup the chemical material	Police Fire Hazmat	Squad Car Fire Truck Chemical Truck

Appendix H

Sample Intelligence Briefing

Intelligence Briefing

We have received an inside tip that there may be suspicious activity associated with banks during the course of the day.

The emergency crisis management manual recommends that all bank-related threats be addressed within the first 45 minutes (simulation time) of notification by all relevant parties (i.e., 45 seconds actual time).

Appendix I

Reflexivity Intervention

Reflexivity Condition- Experimenter Script

It is believed that when teams reflect upon and discuss their task performance and strategies, they can improve their scores. You will be given the opportunity to chat with your teammates about the NeoCITIES simulation and potential strategies for improving your performance. When discussing, **please remember to take into consideration all of your experiences in the experiment so far, including the training, the videos, and the performance scenario.**

Please note that you will have 6 minutes for this discussion, and that all communication must be conducted via chat. In addition, below there are three questions listed that will help guide your discussion. You may read all three questions, but please **do not discuss any given question until instructed to do so by the experimenter**. The discussion session time limit is fixed, so trying to rush through or skip questions won't get you out of here any faster. However, if you remember something you want to say after you've moved away from a question, please feel free to still bring it up.

We will now quickly explain the three questions below.

Discussion Questions

1. What were the main points you learned from the video you just watched that could apply to playing NeoCITIES?

<Wait for experimenter before proceeding to Question 2>

2. How well do you think you and your team just performed in Performance Scenario 1? What went right? What went wrong?

<Wait for experimenter before proceeding to Question 3>

3. You will be playing another performance scenario shortly. As a group please come up with strategies for how to improve your performance on this upcoming scenario.

So in summary, Question 1 focuses on the video you just saw and what its main take-away messages were, Question 2 focuses on how you did on the Performance scenario you *just* completed, so it's past focused, and Question 3 uses what you've discussed in Questions 1 and 2 to create strategies for your performance on the next Performance scenario, so it's future-focused.

Are there any questions?

You are now ready to begin your discussion. Please press Escape to get out of the slideshow mode, then click on the Internet Explorer window located on the bottom of your screen in the Task Bar that says "Google Talk" to open the chat room and begin your discussion.

Control Condition- Experimenter Script

We are interested in your opinions on the use of technology and its impacts on people. Please read the question below and discuss it with your teammates. Please note that you will have 6 minutes for this discussion, and all communication must be conducted via chat. The discussion session time limit is fixed, so trying to rush through won't get you out of here any faster.

Please discuss the pros and cons in relation to the following question. Feel free to discuss your own opinions as well as alternate viewpoints that could be taken:

Is the use of technology a benefit or detriment to society and interpersonal relationships?

You are now ready to begin your discussion. Please press Escape to get out of the slideshow mode, then click on the Internet Explorer window located on the bottom of your screen in the Task Bar that says "Google Talk" to open the chat room and begin your discussion.

Appendix J

TMM Measure- Concept Map Example

In the event "Stampede at the Forum" (in which a bomb threat is received by an instructor in Forum Building), please indicate the recommended order in which the units SHOULD HAVE arrived on the scene of the event.

NOTE: You may or may not need to use all of the textboxes listed below.

Please write the appropriate role (e.g., Fire, Hazmat, Police) in the boxes listed below:

Who should have responded 1st?	
Who should have responded 2nd?	
Who should have responded 3rd?	

Appendix K

Manipulation Checks

Storytelling

1. The video you just watched dealt with:
    a. A graduate student named Dan.
    b. I did not watch a video about Dan.

2. What was the video you just saw about?
    a. An emergency response team helping Dan, with Police, Fire, and Hazmat.
    b. A medical team in a hospital, with a Physician, Surgeon, and Nurse.
    c. A general explanation of how teammates can work together in NeoCITIES to tackle complex timing/ordering events (not a video about Dan).
    d. Students being trained as Centre LifeLink emergency response technicians.

Reflexivity

1. Which of the following did your group discussion mainly address?
    a. The pros and cons of technology development and use.
    b. Reviewing team performance and creating strategies.
    c. What it takes to get good grades in school.
    d. The conditions of professional success.

2. To what extent did your group discussion involve reviewing your team's performance in NeoCITIES and creating strategies for future performance?
    a. Scale ranging from 1= "Not at all" to 5= "To a great extent"

3. Reflexivity Scale (Items adapted from the Carter and West (1998) Task Reflexivity Scale)

Please think about your conversations with your teammates during discussion session you just had and answer the following questions using this scale:

1= Strongly disagree
2= Somewhat disagree
3= Neutral/Neither agree nor disagree
4= Somewhat agree
5= Strongly agree

1. The team reviewed its performance so far.
2. The methods used by the team to dispatch resources to events were discussed.
3. We discussed whether the team is working effectively together.
4. We re-prioritized our actions when circumstances changed.
5. We did not change our team strategies. (R)
6. We discussed how well we communicate information.

7. We reviewed our approach to successfully completing events in the simulation.
8. We did not alter the way decisions are made in this team. (R)

Affective Response

1. The group discussion session allowed me to bond with my teammates.
    a. Scale ranging from 1= "Strongly disagree" to 5= "Strongly agree"

2. PANAS

This scale consists of a number of words that describe different feelings and emotions. Read each item and then choose the appropriate answer next to that word. Indicate how you feel right now at the present.

1= Very slightly or not at all
2= A little
3= Moderately
4= Quite a bit
5= Extremely

1. Interested
2. Distressed
3. Excited
4. Upset
5. Strong
6. Guilty
7. Scared
8. Hostile
9. Enthusiastic
10. Proud
11. Irritable
12. Alert
13. Ashamed
14. Inspired
15. Nervous
16. Determined
17. Attentive
18. Jittery
19. Active
20. Afraid

Appendix L

Ancillary Analyses Items

Reflexivity Helpfulness- Administered Immediately Following Intervention

1. How helpful do you think the group discussion session you just had will be for improving your performance in the next scenario?
    a. Scale ranging from 1= "Not at all helpful" to 5= "Extremely helpful"

Reflexivity Helpfulness Scale- Administered After the Second Performance Scenario

Please indicate the extent to which you agree or disagree with the following items regarding the group discussion session you had in Google Chat.

1= Strongly disagree or definitely false
2= Slightly disagree or mostly false
3= Neutral/equally true or false
4= Slightly agree or mostly true
5= Strongly agree or definitely true

1. The discussion session helped me to play the NeoCITIES game better.
2. I thought about the discussion session while playing the NeoCITIES game.
3. I saw connections between what we discussed in the discussion session and NeoCITIES while playing the game.

Strategy Generation Items

Please indicate whether your team came up with each of the following conclusions/strategies during your group discussion session in Google Chat.

1=Yes
2=No

1. Communicate with each other more when playing the NeoCITIES game.
2. Tell each other what your briefings said.
3. Wait to send out units in a particular order.
4. Make sure units arrive at events in time.
5. Other (optional).

Strategy Utilization Items

Please indicate the extent to which your team actually used the following conclusions/strategies during your most recent performance session.

a. Scale ranging from 1= "Not at all" to 5= "To a great extent"

1. Communicate with each other more when playing the NeoCITIES game.
2. Tell each other what your briefings said.
3. Wait to send out units in a particular order.
4. Make sure units arrive at events in time.