ABSTRACT

Title of Dissertation: MEASURING TRANSACTIVE MEMORY SYSTEMS USING NETWORK ANALYSIS

Kylie Goodell King, Doctor of Philosophy, 2017

Dissertation directed by:
Associate Professor Laura Stapleton
Department of Human Development and Quantitative Methodology

Assistant Professor Tracy Sweet
Department of Human Development and Quantitative Methodology

Transactive memory systems (TMSs) describe the structures and processes that teams use to share information, work together, and accomplish shared goals. First introduced over three decades ago, TMSs have been measured in a variety of ways. This dissertation proposes the use of network analysis in measuring TMS. This is accomplished by describing the creation and administration of a TMS network instrument, evaluating the relation of the proposed network measures and performance, and considering the validity of the proposed network measures. Although the proposed network measures do not appear to be valid in their current form, this study provides motivation for future exploration of using instrumental networks as measures of TMS.

MEASURING TRANSACTIVING MEMORY SYSTEMS USING NETWORK ANALYSIS

by

Kylie Goodell King

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:

Associate Professor Laura Stapleton, Co-Chair
Assistant Professor Tracy Sweet, Co-Chair
Associate Research Professor Joseph Bailey
Associate Professor Vijaya Venkataramani
Assistant Professor Ji Seung Yang

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Transactive memory systems (TMSs) involve the structures and processes that teams use to share information and perform tasks. First introduced in the field of social psychology (Wegner, Giuliano, & Hertel, 1985), TMSs have been used to describe team behavior and predict team performance for over three decades. Although a number of measurement techniques and instruments exist for the measurement of TMS (as described in Ren and Argote [2011]), the most widely used and highly regarded instrument was proposed by Lewis (2003). It is important that the performance of Lewis's instrument and all TMS measurement instruments are evaluated to ensure that measurements obtained from the use of these instruments are valid.

Although studies about TMS have been conducted in fields ranging from organizational behavior to psychology to education, the construct remains relatively unknown to many scholars and practitioners. Therefore, in addition to validating measures of TMS in order to promote the appropriate use of this construct, it is important to relate TMS to team outcomes. This exercise will enable TMS to be more broadly understood and used to help teams maximize performance.

Lewis's (2003) inventory uses 15 self-report items related to three aspects theorized to underlie a TMS: specialization, credibility, and coordination (Liang, Moreland, & Argote, 1995). Although Lewis reported preliminary evidence of her instrument's construct validity, I believe that the proposed inventory, and other previously proposed measures of TMS, have several shortcomings. These shortcomings are primarily related to the way that responses are aggregated. I believe that the constructs proposed by Liang and colleagues and evaluated by Lewis, while meaningful in theory, are not adequately measured by previously proposed inventories (e.g. Austin, 2003; Lewis) but could be tapped in a meaningful way by using a network analysis to measure TMS. By focusing on the relations between individuals on a team, in both gathering data and reporting results, I believe that a more valid measure of TMS may be obtained.

This dissertation describes the application of social network analysis to TMS measurement. More specifically, this document describes (1) the development of an inventory that provides information about trust, enjoyment, and coordination networks and (2) the use of density and transitivity of strong ties in those networks as measures of TMS. High levels of density or transitivity of strong ties in each network would indicate high levels of TMS. Briefly reviewed below, trust, enjoyment, and coordination networks, network density, network transitivity, and strong ties will be described in detail in Chapters 2, 3, and 4 of this dissertation.

In network analysis, the ties between individuals are the variables of interest. This is an interesting methodological technique that may be distinguished between other approaches to team construct measurement as data are collected and interpreted at the individual level. Networks that are comprised of individuals holding many ties are said to demonstrate high density (Wasserman & Faust, 1994). Considering previous operationalization of TMS, a team demonstrating high network density may be characterized as being comprised of individuals who indicate high levels of credibility and coordination among one another and also hold a high degree of specialization.

In addition to density, centrality, reciprocity, and transitivity are other frequently studied attributes of networks. While there are many different measures of centrality (Freeman, 1978), betweenness centrality and indegree centrality are often evaluated in network studies. Node betweenness relates to the importance of one person to the network. For example, if that person were removed, previously connected individuals would no longer be connected. Indegree centrality is a measure of the number of incoming ties received by an individual. In a friendship network, an individual with high indegree centrality would be indicated as a friend by many other individuals in the network. Reciprocity is a measure of network ties that are bi-directional. For example, reciprocity is present where two individuals both indicate a “trust” tie with one another. Transitivity is a measure of how often two individuals tied to one another are also tied to a third individual (e.g. “the friend of a friend is also a friend”). Density, centrality, reciprocity, and transitivity are described in detail in Chapter 2.

Previous studies have shown that network analysis holds promise for measuring TMS (e.g. Lee, Bachrach, & Lewis; 2014; Lewis & Herndon, 2011) and that various attributes of networks may influence team performance (Tröster, Mehra, & van Knippenberg, 2014). This dissertation, therefore, argues for the continued use of network analysis in measuring TMS. More specifically, this dissertation describes the motivation and construction of a network inventory that may be used to measure the strength of various networks related to TMS, the importance of evaluating the density and transitivity of those networks, and how results of the proposed network inventory may be evaluated (based on Messick’s [1995] conceptualization of construct validity). This study is conducted in the context of undergraduate student teams.

This work carries two important implications. First, it proposes and validates a new instrument that may be used to measure TMS. Although the generalizability of the scores obtained from this instrument was not evaluated beyond teams of students, this work provides motivation for future research regarding the ability for this instrument to be used more broadly, for example, within professional organizations.

Second, researchers of small teams will be able to use results regarding attributes of various team networks to understand how different types of ties between individuals lead to different levels of TMS and team performance. Practically, these results could influence the manner in which teams are assigned and managed to enhance performance. By obtaining a detailed understanding of the various forms that ties between individuals may take, teamwork may be structured so as to maximize opportunities for success and early interventions may be made (prior to evaluations of performance).

In summary, the goal of this project was to propose and validate a new measure of TMS. The development of this new measure has been influenced by literature on TMS and social network analysis, results of a pilot study in which a form of the proposed measure was tested, and feedback from subjects about the proposed instrument's usability. The validation of the new measure was based on Messick's (1995) review of the necessary elements to demonstrate construct validity. This project contributes to both the theory and practice of teamwork within small groups.

Chapter 2: Literature Review

In order to illustrate the importance and increasing prominence of team studies, this chapter provides an overview of TMSs among other team constructs. Construct validity and its various aspects (as described by Messick [1995]) are then introduced in order to (1) lay the foundation for the validity study that is discussed in Chapter 4 and (2) provide context for the evaluation of previous approaches to TMS measurement. These previous measurement schemes are then described, including a discussion of their limitations.

This review of previous techniques motivates the continued evaluation of methods that may be used to measure TMS, including social network analysis (SNA). This chapter goes on to describe SNA and some of the attributes that are typically evaluated in social network research. Following this brief introduction to SNA, two previous studies that employed SNA in evaluating TMS are described and opportunities for future analysis are proposed. Finally, the research questions that motivate this study are outlined.

Transactive Memory Systems

In recent years, the study of team-level constructs has become increasingly popular in a number of different fields. In disciplines including organizational behavior, social psychology, education, and many others, researchers are interested in evaluating the presence, emergence, and measurement of multilevel, or individual- and team-level, constructs (e.g. Chen & Kanfer, 2006; DeShon, Kozlowski, Schmidt, Milner & Wiechmann, 2004; Guzzo & Dickson, 1996).

The basic structure of many multilevel studies involves individuals who are nested within teams. In these contexts, an individual is a single person who interacts with other individuals. These individuals, collectively, comprise a team. Although various definitions of teams have been proposed, this paper will utilize Kozlowski and Bell's (2003) definition: teams are composed of two or more individuals who collectively hold one or more common goals, share interdependent tasks while maintaining task boundaries, and interact socially as part of a higher-level context that constrains the team. Although some previous work has provided a distinction between groups and teams (e.g. Katzenbach & Smith, 1993), for the remainder of this document, the term "team" will be used, replacing "group" or "cluster" in reviewing previous studies to maintain consistency of language.

In multilevel modeling, this presence of individuals within teams is known as "nesting" (Raudenbush & Bryk, 2002). The nesting of individuals within teams is an important precursor to the existence of many team-level constructs and enables evaluation at both the individual and team levels (Raudenbush & Bryk).

Some of the team-level constructs that can be evaluated in multilevel contexts include well-known attributes or outcomes such as team communication (Adams & Galanes, 2009), team problem solving (Bormann & Bormann, 1988), team decision making (Napier & Gershenfeld, 1973), collaborative learning (Bransford, Brown, & Cocking, 1999; Johnson, Johnson & Smith, 2006; Springer, Stanne, & Donovan, 1999), performance and effectiveness (Chen & Kanfer, 2006; DeShon et al., 2004; Guzzo & Dickson, 1996; van Knippenberg, De Dreu, Homan, 2004; Weber & Murninghan, 2008), team creativity (Chen, Campbell-Bush, Farh & Wu, 2013; Richter, Hirst, van

Knippenberg, & Baer, 2012), team motivation (Park, Spitzmuller, DeShon, 2013), negative affectivity (van Knippenberg, Kooij-de Bode, van Ginkel, 2010), and collaboration (Kuljanin, 2011).

Other team-level constructs studied under a multilevel framework are less widely known. Two such constructs include shared team mental models (DeChurch & Mesmer-Magnus, 2010), defined as collective, team-level beliefs about the roles and abilities of a team, and transactive memory systems (TMSs) (Wegner et al., 1985), which involve distributed team knowledge and the processes for sharing that knowledge (Wegner, 1987). Shared team mental models, TMSs, and other less commonly recognized team-level constructs have begun to receive recent attention in both laboratory and field settings. The focus of this dissertation is the measurement of TMSs in a field setting.

TMSs were first introduced by Wegner and colleagues (1985) to describe the behavior of couples in close relationships. They have since been evaluated by a number of researchers in different contexts (e.g. Hollingshead, 1998a, 1998b, 2001; Lewis, 2003; Wegner, Erber, & Raymond, 1991). In many of these studies, three aspects of TMSs were evaluated: specialization (differentiated expertise), credibility (trust in the knowledge of other team members), and coordination (the ability to work together smoothly) (Lewis, 2003; Liang et al., 1995). Although similar to shared team mental models and other team constructs, a key feature of TMSs is that they emerge dynamically (Chan, 1998). That is, a team's TMS develops over time through repeated team member interactions.

Although the term "TMS" may be relatively unfamiliar, there are many practical examples of TMSs in use. One simple example involves a family decorating for a holiday. Suppose one member of the family, Anne, is responsible for remembering the holiday, a second member, Ben, is responsible for storing decorations, and a third member, Christine, is responsible for hanging decorations. This TMS proceeds as follows: 1) Anne remembers that the holiday is approaching; 2) Anne knows that Ben has stored decorations, so Anne asks Ben to retrieve these items; 3) Ben retrieves the decorations; 4) Ben knows that Christine can hang these items, so he asks her to do so; 5) Christine hangs the decorations. This process is repeated with each approaching holiday.

In this scenario, Anne, Ben, and Christine convey a high level of specialization (they each hold different skills and abilities), credibility (they believe one another to be credible in their unique skills), and coordination (they use knowledge of one another's skill to complete the task of hanging decorations). Thus, they demonstrate a well-functioning TMS. If they were to be lacking in any of these three areas, for example, if they all had overlapping knowledge and their "specialization" were reduced, their team-level TMS would also be lowered.

Although this specialized knowledge is certainly important to a well-functioning TMS, it is again important to note that TMSs include both a structural component of distributed knowledge (this distribution of knowledge falls on a spectrum from knowledge that is completely shared to knowledge that is completely unique to each team member) and the processes by which the knowledge of who knows what is shared amongst individuals within a team. Again, this is the process component of a TMS. Using the above example, it is likely the TMS of Anne, Ben, and Christine developed over several holiday seasons before reaching its current state.

The TMS of Anne, Ben, and Christine is a very simple example; however, the three aspects of TMSs outlined above; specialization, credibility, and coordination, (Liang et al., 1995), may be applied in a number of other settings. For example, TMSs exist among work teams, student groups, healthcare providers, and in many other contexts.

Lewis (2003) describes TMSs as “the cooperative division of labor for learning, remembering, and communicating relevant team knowledge” (p. 587). TMSs are different than transactive memory, as the former is measured at the team level while the latter is measured at the individual level (Kozlowski, Chao, Grand, Braun, & Kuljanin, 2013). According to Lewis, transactive memory is influenced by the knowledge that one person has about the knowledge of another person, or “knowledge of who knows what” (Richter et al., 2012). This individual-level concept, also referred to as KWKW, is important to Lewis’s definition of TMSs.

Also according to Kozlowski and colleagues (2013), TMSs may be distinguished from transactive memory as TMSs involve the active process of two or more people collecting, sharing, and utilizing information. Transactive memory, on the other hand, is an individual-level variable that describes one person’s KWKW and that person’s beliefs about whether or not information that is held by others may be accessed.

TMSs, therefore, emerge at the team-level through the aggregation, or more specifically, compilation, of individual-level transactive memory. Kozlowski and Klein (2000) distinguish between team-level constructs that emerge via composition versus compilation. Whereas composition is based on assumptions of isomorphism, or similarity across levels of measurement (e.g. the construct is the same at the individual-level and team-level and a simple summation of individual-level attributes appropriately represents a construct at the team-level), compilation describes any construct that is distinct as it emerges from the individual- to team-level. TMSs have been conceptualized as emerging through compilation. A team's TMS may be greater than the sum of its individual parts, enabling a team to access information that is not available to any one individual (Wegner et al., 1991).

Although some scholars and practitioners may be interested in evaluating a team's TMS to understand how that team functions early in a team project, I believe that TMSs are also useful as they relate to team performance. Teams with high levels of TMS are typically high performing (Lewis, 2003). Also referred to as team effectiveness, team performance is a team's capacity to achieve its goals and objectives (Kozlowski & Klein, 2000). Independent of the context (e.g. work, school, sports), a team's ability to achieve goals leads directly to enhanced team outcomes.

The relation between TMS and team performance serves as a measure of the external validity of this construct. Especially where innovative or creative tasks are required, teams must be able to demonstrate the ability to seek information, share information, and coordinate tasks in order to achieve high levels of performance (Parrotta, Pozzoli, Pytlikova, 2014; Williams & O'Reilly, 1998).

As previously mentioned, TMSs are related to, but distinguishable from, other team-level constructs, including shared team mental models (Cannon-Bowers & Salas, 2001). Shared team mental models are the knowledge structures within a team that enable the team to form expectations for tasks and coordinate their activities based on these expectations. This construct has obvious ties to TMSs in that it involves the coordination of knowledge to meet team goals. A distinction between shared team mental models and TMSs is that TMSs must involve not only team knowledge and expectations (a structural component), but also the mechanisms by which that knowledge is shared within a team (a process component) (Kozlowski et al., 2013; Lewis & Herndon, 2011). This process of sharing information within a team must be considered by TMS measurement schemes.

Another team-level construct that may contribute to a well-functioning TMS is functional background diversity, or FBD (Richter et al., 2012). According to Richter et al., FBD "reflects differences in knowledge, information, and perspective that are relevant to a team's tasks" (p. 1284). FBD contributes to knowledge specialization, one of the three tenets of TMSs identified by Liang and colleagues (1995). FBD, therefore, is an important precursor to the emergence of TMSs and measurement of TMS should account for the FBD of team members.

Although studied under a number of different contexts, many of the conceptual roots of TMS come from educational psychology. For example, in cooperative learning teams, students learn from and teach one another (Johnson et al., 2014). Typically composed of students with different skills or functional backgrounds, the use of cooperative learning teams is based on the principles of constructivism, the idea that individuals learn by connecting new ideas to existing knowledge (Bransford et al., 1999), and social interdependence theory, which states that teams emerge as dynamic entities that exhibit interdependence when sharing a common goal (Koffka, 1935; Lewin, 1935). The use of small groups in educational contexts has been noted as being essential to the process of cooperative learning as students are able to benefit from goal and resource interdependence (Johnson et al., 2014).

Cooperative learning has been found to increased academic achievement in a number of different types of tasks (from knowledge retention to creative problem solving) and in many different domains (Johnson et al., 2006; Springer, Stanne, & Donovan, 1999). Furthermore, beyond academic benefits, cooperative learning activities have also been shown to contribute to self-esteem and engagement (Kuh, Cruce, Shoup, Kinzie, & Gonyea, 2008; Springer et al., 1999). Considering these benefits, it is no surprise that constructs such as TMS, which require cooperative learning, have become popular research topics over the past few decades. In using TMS as a measure of cooperative learning, the expected relation of TMS and team performance should follow the relation between cooperative learning and individual performance. This highlights the ability for performance to serve as a measure of the external validity of a TMS measurement instrument.

All of measures of team structures and processes, including shared team mental models, FBD, and TMS, are subject to considerations of validity. Construct validity, whether results obtained from a measurement instrument adequately represent the construct of interest, is important to consider in evaluating previous measurement instruments and proposing new measurement techniques. In the next section, I discuss the aspects of validity that are especially important to consider in the evaluation of this new instrument.

Construct Validity

A single study cannot alone prove construct validity. Evaluations of construct validity are ongoing and should incorporate new information as additional validity studies are carried out. In first proposing an instrument, pilot studies may be conducted to test the feasibility of a proposed instrument while allowing for adjustments before an instrument is deployed. This study relies on pilot work to evaluate the utility of a social network inventory to tap a construct of interest, TMS.

Construct validity is extremely important in all social science research. It is comprised of both the meaningfulness and interpretability of scores (Messick, 1995). Measurements obtained from a given instrument are said to hold construct validity to the degree that those measurements paint an accurate picture of the phenomenon of interest. Involving content, substantive, consequential, structural, and external aspects of validity along with generalizability, construct validity requires both theoretical and empirical support and provides justification for score interpretation (Messick, 1995). Evaluations of validity must include each of the aspects listed above or a strong argument about why certain aspects of validity were excluded (Messick).

Content aspects of validity involve content relevance and representativeness of an assessment. Substantive aspects of validity provide theoretical rationales for observed responses. External aspects of validity encompass both convergent validity (a correspondence between measures of the same construct) and discriminant validity (distinctness between measures of different constructs). Consequential aspects of validity involve an evaluation of intended and unintended positive and negative consequences of using measurements obtained from a given instrument in both the short- and long-term. Structural aspects of validity consider the relation between measurement instruments and the structural relations present in a construct, typically demonstrated using factor models or item-response theory (IRT) models. Finally, generalizability involves the degree to which measurements obtained using an instrument in one use case are able to generalize across different groups and in different settings. This project considers content, substantive, external, and consequential aspects of validity of the proposed measurement instrument. It explicitly tests measures obtained from the proposed instrument for external aspects of validity and generalizability. Structural aspects of validity are not evaluated in the present analysis as a new structural model of TMS is not proposed. More detail about these considerations and evaluations are discussed in detail in Chapter 4.

Critique of Previous TMS Measurement Schemes

TMSs have been measured in both laboratory and field contexts. Early laboratory studies (Wegner et al., 1991; Hollingshead 1998a, 1998b, 2001; Liang et al., 1995; Moreland, 1999; Moreland & Myaskovsky, 2000; Lewis, 2003) used a variety of measurement techniques. Many of the early measures of TMSs used an outcome, word recall, to infer the existence of this construct (Wegner et al., 1991; Hollingshead 1998a, 1998b, 2001). Other studies utilized rater evaluations to measure TMSs in a team setting (Liang et al., 1995; Moreland 1999; Moreland & Myaskovsky, 2000). In a third approach, Lewis (2003) employed factors first proposed by Liang et al. (1995) to develop self-report indicators of TMSs. Below is a brief history of the measurement of TMS. Although the measurement of TMS has evolved over the past three decades, there remain opportunities to improve how TMS is measured by incorporating new measurement techniques that consider the individual-level relations between team members.

Performance Measures of TMS

The first studies to propose and evaluate the construct of TMS include those conducted by Wegner and colleagues (1991) and Hollingshead (1998a, 1998b, 2001). All of these studies involved pairs of participants and the outcome measure of interest in these studies was word recall. While quite similar, these studies varied in some of the experimental conditions imposed upon pairs of participants.

Wegner and colleagues (1991) varied experimental conditions to include either randomly assigned partners or pairs who were involved in a romantic relationship. Expertise was also manipulated to be either naturally occurring or randomly assigned. Hollingshead (1998a) also varied partner assignment to be either random or based on romantic relationships and evaluated performance based on communication constraints (communication was either permitted or not permitted). Hollingshead (1998b), again, varied the familiarity of pairs in addition to communication occurring in person or in an online environment. Additionally, this study incorporated the use of raters who watched recorded videotapes and evaluated a set of predetermined behaviors. Finally, Hollingshead (2001) created conditions that varied expectations of partner knowledge (similar to or different form one's own) and incentives for memorizing different types of information.

Several interesting findings emerged from these early studies. Wegner et al. (1991) found that in existing relationships, performance was hindered under the assignment of arbitrary expertise. More precisely, existing romantic couples assigned categories for memorization recalled significantly fewer items than existing romantic couples not assigned categories. Additionally, when categories were not assigned, existing romantic couples recalled significantly more information than randomly assigned couples.

Hollingshead (1998a) found that that existing couples recalled more information than randomly assigned couples when pairs could not communicate during learning and randomly assigned couples recalled more words than existing partners when communication during learning was permitted. Stated in a different way, there was a significant interaction of relationship (existing versus randomly assigned pairs) and communication during learning.

Findings of Hollingshead (1998b) indicate that, in face-to-face settings, existing couples performed better than randomly assigned couples. Additionally, existing couples in face-to-face settings outperformed existing couples in the computer conferencing condition. In terms of communicative behaviors, the raters noted several significant differences between existing couples and randomly assigned partners. In randomly assigned pairs, assertions of expertise and references to common knowledge occurred significantly more regularly than in existing couples and transactive information searches and references to shared knowledge occurred significantly less frequently.

Results of Hollingshead's third study (2001) demonstrated that individuals were more likely to consider their partner's expertise in the memorization of knowledge when incentives for differentiation of knowledge were present. Additionally, results showed that individuals "specialized" (recalled more words in their areas of expertise) most when their expectations of partner expertise aligned with their incentives. For example, individuals paired with partners with different expertise performed significantly better when incentivized to recall unique information.

Considered collectively, these early studies illustrate some of benefits that a pair can reap from a well-functioning TMS (the ability to recall information) and some of the problems that can arise when this system is disrupted. Furthermore, these studies demonstrate that existing couples may utilize non-verbal communication in face-to-face settings, which, in turn, may contribute to a well-functioning TMS. While these early studies of TMS contributed to the conceptualization of this construct and its relation to team performance, they were improved in later research that considered the presence of TMS in small teams as well as the dynamic nature of TMS emergence.

Evaluating TMS in Teams

Although originally explored in studies of pairs, other early researchers of TMSs expanded their conceptual framework to evaluate the presence of this construct in small teams (e.g. Liang et al., 1995; Moreland, 1999; Moreland & Myaskovsky, 2001). In the previously discussed pairs studies (Wegner et al., 1991; Hollingshead 1998a,b, 2001), although the relation between pair workflow processes and performance outcomes is clearly defined, a conceptual framework for TMSs is missing. The studies that evaluated TMSs in small teams improved upon the conceptualization of TMSs by providing a continuum along which strength of a TMS may exist.

Each of these studies involved the evaluation of temporary work teams of three students tasked with building an AM/FM radio. Outcome variables in these studies included recall of assembly steps, performance (the number of correctly placed radio parts and the time required to assemble the radio and the number of steps), and measures of TMS.

In all three of these studies (Liang et al, 1995; Moreland, 1999; Moreland & Myaskovsky) TMS was indicated by three cognitive aspects[1]; memory differentiation (the specialization of team members in different aspects of the assembly task), task coordination (the ability of team members to work well together), and task credibility (how much team members trusted the knowledge of other members regarding radio assembly). Two external raters reviewed videotapes of team interactions and evaluated the three aspects. Raters were provided with a list of specific behaviors exemplifying each factor and were then asked to provide one summative rating for each aspect (the list of behaviors was not published). All evaluations were made at the team level using a 7-point scale with higher ratings indicating greater levels of each factor (no information was provided about the labels of the scale points).

These three early studies of TMS in teams differed slightly in the team training conditions that were imposed. Liang and colleagues (1995) had two training conditions. Teams were either trained together or individuals who were later placed on a team were trained alone. Moreland (1999) used four training conditions. Individual training, individual training with team-building exercises (added to enable the evaluation of TMSs versus general teamwork skills), team training, and team training with reassignment of teams prior to radio assembly (added to enable evaluation of differences between TMSs, developed through the process of repetitive interactions, versus teamwork skills obtained through general team training). Moreland & Myaskovsky (2001) used three training conditions, including individuals trained by themselves, teams trained together, and

---

[1] Although some of the studies reviewed in this dissertation make use of the term “factor” to describe memory differentiation, task coordination, and task credibility, this dissertation will replace “factor” with “aspect” to maintain consistency of terms.

individuals trained alone and receiving information about the skills of their team members (to enable evaluation of the impact of KWKW).

Notable results of these studies include that there were significant differences in TMS aspects across training conditions. Liang and colleagues (1995) found that teams trained together demonstrated significantly higher levels of memory differentiation, task coordination, and task credibility than individuals trained alone. After analyzing the three aspects individually, the three team-level scores were combined to yield an overall measure of a TMS, justified by the authors based on the significant and positive correlations among the three aspects. Furthermore, Liang and colleagues noted that collective team training led to significantly better performance in terms of number of steps remembered and error reduction. There was not a significant difference in assembly time.

Moreland (1999) found that teams who trained together and later completed the assembly task together outperformed teams where individuals were trained individually and teams who trained together but were then shuffled into different assembly teams. According to Moreland, this supports the notion that generic training programs, which assume that experimental learning in teams transfers from one team to another, are not likely to succeed. Additionally, it reinforces that idea that TMS is distinct from general team skills.

Findings from Moreland and Myaskovsky (2001) showed that TMS index scores were significantly different across the three conditions (individual training, team training, individual training with KWKW). Post hoc (Tukey) tests indicate that significant differences were present between teams trained as individuals and teams trained in the other conditions while differences between the team training and performance feedback conditions were not significant. These results support the claim that the performance benefits associated with group training are not due to communication among team members but rather TMS.

Results of these studies contribute to the theoretical conceptualization of TMS as a construct that is learned through repeated interactions among pairs or team members (“process composition”, as defined by Chan, 1998) and, according to Liang and colleagues (1995) provides preliminary evidence that TMSs operate in small teams in addition to pairs. Furthermore, these studies advance the conceptual understanding of TMS by providing mechanisms for measuring TMS as distinct from other, more static constructs that may lead to enhanced team performance (e.g. general teamwork sills). Additionally, the measurement of TMS structure in terms of complexity, accuracy, and agreement contributes to TMS conceptualization.

A Self-Report Measure of TMS

A third approach to the measurement of TMSs was proposed by Lewis (2003). Rather than involving recall of information or ratings by evaluators, this study used self-reported ratings of each of the three previously proposed factors underlying TMSs (Liang et al., 1995). Lewis created a 15-item scale to measure TMS based on a specialization subscale, a credibility subscale, and a coordination subscale. Five items were created for each subscale to elicit individuals’ self-reports. These items, which may be found in Appendix A of this document, were scored using a 5-point Likert scale. Lewis, therefore, was the first author to explicitly conceptualize TMS as a second-order factor comprised of three first-order factors, with each first-order factor indicated by five items. Lewis was also the first author to provide specific item detail (e.g. the 15 item inventory), removing the "black box" of TMS measurement and enabling future researchers to compare findings by utilizing the same scale.

Lewis's (2003) study uses confirmatory factor analysis to evaluate the structure of TMS, exploring the utility of three subfactors as indicators of TMS and the plausibility of a 15-item inventory to indicate each subfactor. Figure 1 below illustrates Lewis's (2003) higher-order factor model of TMS. In this model, TMS is conceptualized as a higher-order factor, comprised of three subfactors, specialization, credibility, and coordination.

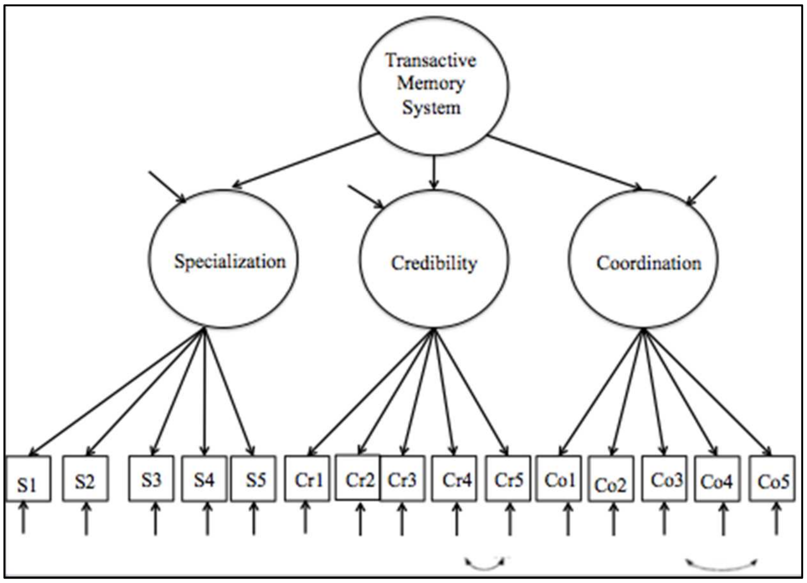


Figure 1. Lewis's (2003) higher-order factor model of TMS

This model was used to evaluate data from three scenarios: a laboratory study in which students were placed into teams and asked to assemble an electronic device, a field study in which MBA students provided management consulting services to local firms, and a field study involving managers and team members from 11 high-tech companies. In each of these three studies, items measured at the individual level were aggregated for the team and average team score was used as a team-level measure of TMS. Factor loadings from the laboratory study and field study of MBA students are provided in Appendix B. These loadings indicate the strength of the relation between TMS and the three subfactors as well as the relation between the three subfactors and each of their respective items.

It should be noted that in aggregating individual responses to the team level, potentially useful information about differences between team members might be lost. For example, in measuring credibility, Lewis's method of score aggregation assumes that there are no differences in team-level credibility between a team with four reasonably credible members and a team with two highly credible members and two members who know nothing. In both of these scenarios, the teams could have an average credibility score of 3 (3, 3, 3, and 3 versus 5, 5, 1, and 1). Under the method of score aggregation proposed by Lewis (2003), it is implied that this team-level rating is sufficient. It is not clear that this is the case. Evaluating TMS using individual responses may yield very different results. For example, the relation between team-level scores and performance outcomes is likely to be stronger in situations where there is less within-team variance. This idea will be described in detail later in this chapter.

Lewis conducted reliability and validity testing of aggregated TMS scores using data from each of these three studies (involving teams tasked with assembling a radio, MBA consulting teams, and high-tech teams). In the radio assembly study, convergent validity, a measure of Messick's external aspect of validity (1995), was evaluated by comparing the aggregated team scores to observational ratings of team behaviors indicative of TMS. Correlations between raters' scores of the teams and subscale team averages were all moderately strong and statistically significant with r=0.34 for specialization, r=0.41 credibility, and r=0.51 for coordination. Correlations between the aggregate performance measure and TMS were not provided.

Discriminant validity, another external aspect of validity, the ability to discriminate between different constructs, was established by conducting two comparisons. Illustrations of the models that were compared are provided in Appendix C. In the first portion of each model comparison, an additional construct, team motivation or team cohesiveness, indicated by two items each (provided in Appendix A), was added. New items were specified to load on their respective factors and were allowed to covary with the TMS second-order factor. In the second portion of each model comparison, motivation and cohesiveness were specified to load on the TMS factor instead of their own factors. Results of chi-square difference tests of both model comparisons indicate that the first models (the models with motivation or cohesiveness items loading on their own factors) demonstrated better fit with the data. This led Lewis to conclude that motivation and cohesiveness were distinct from the three factors comprising TMS and provides some evidence of discriminant validity.

Convergent validity was examined in two additional ways. First, between-team differences were evaluated for teams trained together versus teams trained individually. Lewis expected to find that collective training would lead to higher levels of each of the three factors and, therefore, higher levels of TMS. For all three factors, teams trained together did, in fact, have higher scores than those trained individually (d = 0.377, d = 0.422, d = 0.637 for specialization, credibility, and coordination, respectively). Convergent validity was again evaluated by comparing TMS scale scores to performance, as measured by assembly error and time required to complete the assembly task. Results of this testing illustrate that the scale had a significant, strong, positive relation with performance.

In the second study of MBA student consulting teams, internal consistency was indicated by high correlations of items with their respective subscales ($r > 0.40$). Additionally, results of a confirmatory factor analysis indicate that the proposed model demonstrated “good fit” (Lewis, 2003) with the individual and team-level data, respectively (SRMR=0.06, CFI=0.93 and SRMR=0.09, CFI=0.93). Note that, while described by Lewis as having reasonably good fit, simulation studies have demonstrated that CFI values should be greater than or equal to 0.95 and SRMR values should be less than or equal to 0.08 (Hu & Bentler, 1999). This discrepancy will be explored further in Chapter 3.

Convergent and discriminant validity were evaluated by comparing the TMS scale scores with two additional constructs; one construct was expected to be related to TMS (agreement on team member expertise), the other construct was not expected to be related to TMS (team autonomy). Both of these constructs were indicated by indices of self-report items. Results support the original expectations with a moderate correlation being found between TMS and agreement of team expertise (r=0.55, p<0.05) and a weak correlation between TMS and autonomy (r=0.28, p<0.05) (p. 598). External validity was again evaluated by comparing TMS to team performance, as measured by team advisors and team clients at the end of the semester. Findings suggest that this relation was moderately strong and significant for both advisors (r=0.53) and clients (r=0.48).

In the third study of managers and team members at a high-tech company, the TMS scale correlated positively with team self-assessments of performance (r=0.79,

$p<0.05$) and manager ratings of performance ($r=0.73$, $p<0.05$), supporting convergent validity, a piece of Messick's (1995) external aspect of validity. Discriminant validity, another piece of Messick's external aspect of validity, was evaluated by comparing the relation between autonomy and TMS. This relation ($r=0.36$) was not significant.

A number of subsequent studies have used Lewis's (2003) inventory as a measure of TMS. According to Ren and Argote's review of TMS literature (2010), of 31 field studies of TMS, 12 used Lewis's 15-item scale. These studies vary by field and purpose, but all are consistent in describing an application of TMS measurement. Unfortunately, none of these studies evaluated the fit of Lewis's model in their discussion of the inventory's use and results. In fact, only one of the studies reviewed by Ren and Argote provided any discussion of the Lewis's model fit by describing factor loadings (Gino, Argote, Miron-Spektor & Tordorova; 2010). This acceptance of Lewis's model without evaluation of its fit is somewhat worrisome and warrants further investigation.

Lewis (2003) concluded that TMS is a construct tapped by specialization, credibility, and coordination. Although these contributions certainly enhance the reader's understanding of TMS and how this construct may be measured, future researchers should consider whether aggregating individual ratings to the team level sufficiently measures specialization, credibility, and coordination. As previously mentioned, it seems as though important information may be lost in this aggregation, as differences between team members are included in measurement results as error rather than as useful information that would help define the TMS of the team.

The use of social network analysis, which focuses on ties between individual members of a team, measures the strength and pattern of intra-team relations. The analysis of ties among individuals on a team may be used as a measure of TMS. This dissertation outlines a study that evaluates the ability of network data to indicate TMS, validated, in part, by the relation of these network data to team performance.

TMS and Social Network Analysis

In recent years, researchers have extended TMS research to incorporate additional analytic techniques. One such technique is social network analysis (SNA), where individuals are represented as nodes and their relationships to other individuals are represented as ties. A social network is a map of all of the relevant ties between individuals and these ties between individuals are the variables of interest (Wasserman & Faust, 1994). Considered broadly, five general theories underlie most network analyses: self-interest theory (e.g. Coleman, 1988), social exchange theory (Homans, 1958), mutual or collective interest theory (Samuelson, 1954), homophily (e.g. Brass, 2011), and various cognitive theories. Specific to this project, one of the most prominent cognitive theories to the network literature of group interaction is that of TMS (Katz et al., 2005).

Small teams are not typically analyzed under SNA as ties between individuals are nearly always present based on the close proximity within which teams work and the lack of alternative individuals with which ties may be formed. I believe, and previous research suggests, however, that the strength of a team's TMS may be indicated by social network attributes, including network reciprocity, transitivity, density, and centrality. This may be true in both small and large teams if the strength of ties between individuals is considered. Each of these social network attributes (reciprocity, transitivity, density, and centrality) is discussed in detail below.

Density

A frequently studied network attribute is density. Conceptually, density is a measure of the total present ties in a network relative to the total possible ties. Mathematically, density is the sum of all entries in a dataset of network ties divided by the possible number of entries in that dataset: $\Delta = \frac{\sum_{i=1}^{g}\sum_{j=1}^{g} x_{ij}}{g(g-1)}$, where i is an individual in a network indicating ties to others, j is an individual in a network receiving ties from others, x_{ij} represents the presence or absence of a tie from individual i to individual j, and g is the number of individuals in the network of interest (Wasserman & Faust, 1994). The relation between network density, TMS, and performance has been studied in the past and will be reviewed later in this chapter.

Reciprocity and Transitivity

Reciprocity, another network measure, represents the proportion of reciprocal ties in a given network. Reciprocal ties are defined by Wasserman and Faust (1994) as $x_{ij} = x_{ji} = 1$, where x_{ij} indicates a tie that to actor j that is nominated by actor i and x_{ji} indicates a tie to actor i that is nominated by actor j. Reciprocity is frequently evaluated in network studies (e.g. Lee, Bachrach, & Lewis, 2014). Network reciprocity is high when individuals in a network who indicate ties with other individuals in a network also receive ties from those individuals (illustrated below in Figure 2). The density of ties in this network would be equal to 1. There are two possible ties and both ties are present.

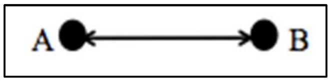


Figure 2. Reciprocity between individuals A and B

In network analysis, transitivity is a relation that is represented by ties between three individuals. When one individual is tied to a second individual and that second individual is tied to a third individual, the first individual is also tied to the third individual ("the friend of a friend is also a friend"). Transitivity is a network measure that is indicated by a transitivity index, $\frac{t}{T}$, where t represents the number of transitive triads present in a network and T represents the number of potentially transitive triads in a network (Wasserman & Faust, 1994). While ties between individuals in a transitive triad are often reciprocal, by definition, this does not need to be the case.

An illustration of a transitive triad of individuals A, B, and C is provided below in Figure 3. In this figure, from the perspective of individual A, the friend (individual C) of a friend (individual B) is also a friend. The density of this network would be equal to 0.67. There are six possible ties between individuals (A to B, B to A, A to C, C to A, B to C, and C to B) and four of those six ties are present (A to B, A to C, B to C, C to B).

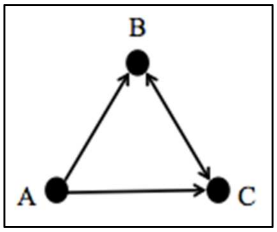


Figure 3. Transitive triad with individuals A, B, and C

There are varying opinions in the literature about the impact of reciprocity and transitivity on team performance. According to some theories, reciprocity and transitivity encourage the exchange of repetitive information and are, therefore, a hindrance to performance (e.g. Katz, 1982). According to others, reciprocity and transitivity help ensure that information is thoroughly and completely shared within a team (e.g.

Granovetter, 1985). A prior study evaluating the impact of reciprocity and transitivity on TMS networks will be reviewed in detail later in this chapter.

Centrality

Centrality is another attribute that is regularly used to describe individuals and networks. Many forms of centrality have been proposed, including degree centrality (the total number of direct ties held by an individual), indegree centrality (the number of incoming ties held by an individual), outdegree centrality (the number of outgoing ties indicated by an individual), and betweenness centrality (a measure of how an individual in a network links otherwise unconnected others) (Prell, 2012). Previous researchers have studied the relation between betweeenness centrality and team performance (e.g. Mell et al., 2014). This will be described in detail below.

Review of Previous Literature Involving TMS and SNA

This project is the first to propose the use of a specific network instrument to measure TMS; however, there are several examples of previous studies that evaluated TMS along with social network attributes. Two of these studies, involving the relation of TMS and transitivity, density, and reciprocity (Lee et al., 2014) and the relation of TMS and centrality (Mell et al., 2014) are discussed below. These studies provide examples of the relation between TMS and SNA and suggest how various network attributes may relate to performance.

TMS and transitivity, density, and reciprocity. Lee and colleagues (2014) developed and tested a path model to evaluate the impact of network reciprocity and transitivity on the development of a TMS. In their study, Lee and colleagues evaluate a number of hypotheses. First, they propose that there is a negative relation between reciprocal ties and network density on TMS, measured using Lewis's (2003) inventory. Second, they hypothesize that there is a positive relation between reciprocal network ties and network density on the number of transitive triads in a team. Finally, Lee and colleagues speculate that there is a positive relation between the number of transitive triads in a team's network and the team's TMS, as measured using Lewis's inventory. Essentially, Lee and colleagues evaluate the impact of network ties (density and reciprocity) on team performance, mediated through TMS and the number of transitive triads. This is depicted below in Figure 4. The overall purpose of this study was to evaluate the relation between network density, transitivity, and reciprocity and team performance.

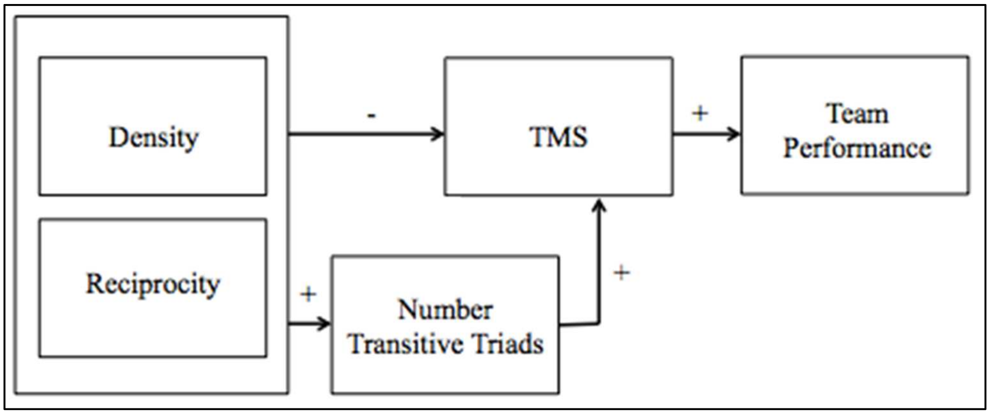


Figure 4. Social network attributes, transitive triads, TMS, and performance

In this study, data were collected from 132 four-member teams of undergraduate students who participated in a business simulation game. The outcome variable, team performance, was indicated by two measures of performance in the simulation game. Information about network ties was obtained by asking each team member to indicate how often each of their teammates provided them with relevant information (on a scale from 0 to 3 with 0 representing never and 3 representing very often). This score was aggregated to the team level. Network reciprocity was calculated as the ratio of the number of present reciprocal ties relative to the total number of possible reciprocal ties. Ties receiving scores of 1, 2, and 3 were considered to be present and included in the calculation of reciprocity. Ties receiving scores of 0 were considered to be absent. Transitivity was evaluated using counts of the number of transitive triads present within a network (team). Network density, also aggregated at the team level, was evaluated as the total number of present ties (i.e. ties that received scores of 1, 2, and 3) relative to the total number of possible ties.

In density calculations, directed ties are used. Directed ties indicate a one-way nominations of ties. For example, a directed tie is present if individual A indicates a tie with individual B but individual B does not indicate a tie with individual A. Therefore, reciprocity is not required in calculations of density. TMS was also measured at the team level using Lewis's (2003) 15-item TMS scale.

This study relies heavily on the TMS inventory proposed by Lewis (2003). While it is worthwhile to review and interpret results, it should be noted that different findings may have emerged had a different TMS measurement instrument been used, as suggested by this dissertation. Nevertheless, results indicate that the TMS scale score, calculated using the Lewis (2003) inventory, was related to team performance (r=0.40 and r=0.33). Furthermore, there was a significant negative relation between TMS and network density when controlling for transitive triads (β=-0.16). This means that the strength of a TMS may be due to the presence of transitive triads than network density, meaning that efficient ties (indicated by transitive triads) may be more important to TMS than the total number of ties.

Despite their dependence on Lewis's (2003) TMS inventory, the results of this study are noteworthy. Lee and colleagues offer a novel approach to evaluating the impact of reciprocal and transitive ties on TMS. I believe that future studies should follow this approach of relating social network attributes to TMS and team performance. Furthermore, I believe that future studies should move beyond simply relating network attributes, such as density, reciprocity, or transitivity, to TMS and consider these attributes as measures of TMS in and of themselves, as argued for in this dissertation.

TMS and Centrality

Mell and colleagues (2014) compared models where KWKW was centralized within one team member versus models where this metaknowledge was distributed evenly among a team. This work was motivated by previous TMS studies evaluating only extremes of knowledge sharing; no team members having any metaknowledge versus all members having complete metaknowledge. The authors argue that both of these situations seemed unrealistic and that partially distributed metaknowledge was more likely to be the rule rather than the exception.

This paper uses a form of centrality, betweenness, in relating TMS to performance (Mell et al., 2014). If an actor holding high betweenness centrality is removed from a network, previously linked entities will become disconnected from other portions of the network. In Figure 5, individuals B and E hold high betweenness centrality. Individuals A, B, and C and individuals D, E, and F are highly connected. Individuals B and E bridge the two subnetworks of triads. If either are removed, the triads will become disconnected.

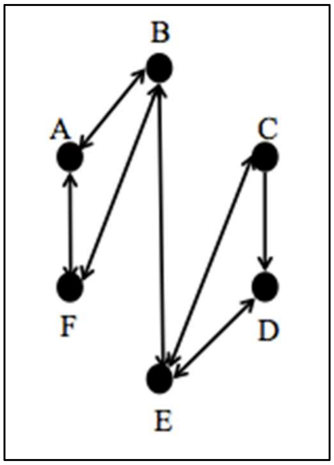


Figure 5. Example of betweenness centrality

Mell and colleagues (2014) define a centralized TMS structure as a situation where at least one team member holds an above-average amount of metaknowledge. They propose that in teams having a more centralized TMS structure, there will be greater transactive retrieval, more information elaboration (two indicators of TMS) and better team performance. In teams with a less centralized TMS structure, transactive retrieval, information elaboration, and performance will be lower. Similar to Lee and colleagues, this hypothesis is formed based on the principle of efficient exchanges of information. It is assumed that if one person has an above-average level of knowledge about the knowledge and skills of others in the network, that they will be able to direct efficient information exchanges and reduce redundancy of information sharing.

In this study, teams of three people participated in a task of evaluating the profitability of product innovations where the information required to find an optimal solution was distributed among the three members. 112 teams participated in this exercise, yielding approximately 28 teams in each of the four conditions. Performance, measured at the team level, was indicated by the similarity of team decisions to an objectively correct solution (a ranking of the five product alternatives).

Results indicated that a centralized TMS network structure (indicated by one team member holding an above-average level of betweenness centrality) resulted in more information elaboration and higher performance than a decentralized structure (Mell et al., 2014). These results enhance TMS theory by integrating the concept of metaknowledge distribution, or network centrality, with TMS. Most important to the proposed study; however, the results of Mell and colleagues indicate that metaknowledge centralization and other attributes of social networks could be useful in predicting how quickly and how effectively teams are able to demonstrate a well-functioning TMS. More information about how these social network attributes may be used in the proposed study is provided below.

Impact of Past Studies on Current Analysis

Density. Lee and colleagues (2014) suggest a negative relation between density and TMS in the case that repetitive information is shared between individuals, resulting in an inefficient use of resources (including time spent communicating redundant information). If transitivity is present, this means that, for the three people involved in the transitive triad, redundant information is unlikely to be shared, as all individuals in the triad are aware of the knowledge that is present and being shared among the others.

In the case of small teams, I believe that these findings will not hold. More specifically, I believe that greater strong-tie density will indicate greater levels of TMS and will have a positive relation with team performance. Strong-tie density, a newly proposed network attribute, will represent the proportion of strong ties (ties receiving scores of 4 or 5 or ties receiving scores of 5) divided by all possible strong ties. This attribute may be calculated where ordered data are collected. I believe that in certain networks, TMS will be higher when there are frequent, strong, reciprocal ties between individuals.

Transitivity. I believe that strong-tie network transitivity and strong-tie network density will have a similar relation to team performance and TMS. That is to say, performance will be higher and levels of TMS will be greater in networks with high strong-tie density and high strong-tie transitivity. This is due to the fact that these network attributes are calculated in a similar manner.

Reciprocity. I do not believe that reciprocity should be used as a measure of TMS where small teams are concerned. This stands in contrast with some of the findings previously described (Lee et al., 2014); however, I believe that by focusing on strong ties rather than all ties (e.g. using just scores of 3 instead of using scores of 1, 2, and 3) different findings regarding the ability for reciprocity to serve a a measure of TMS will emerge. This is especially important to consider in small teams as these networks have the tendency for ties to be present between all individuals due to the close proximity within which individuals work with one another.

Centrality. I do not believe that the findings of Mell and colleagues (2014) regarding the relation between centrality, TMS, and performance will consistently hold in small-group settings. When there are a small number of individuals interacting in a network, there is less of a need for some individuals to coordinate the flow of information or have abnormally strong relationships with all members of their team. Therefore, I believe that betweenness centrality will not be a consistently appropriate measure of TMS in small teams.

A better measure of TMS in small teams may be average indegree centrality, or the average number of ties received by individuals on a team. This measure is closely related to network density and may be defined as $\frac{\sum_{i=1}^{g} x_i}{g(g-1)}$ where g represents the number of individuals in the network of interest, i represents an individual in the network, and x_i represents a present incoming tie for that individual (Wasserman & Faust, 1994). Average indegree centrality is easy to scale where different network sizes are present. It is calculated based on the ties held by all team members rather than a few centralized individuals. It should be noted, however, that calculated values of average indegree centrality are likely to be quite similar to calculated values of density. This will be discussed in more detail later in this chapter.

Considering reciprocity, transitivity, density, and centrality, previous research suggests that TMS is shown by networks that demonstrate strong, frequent, reciprocal ties between individuals and that the coordination of ties (enhanced by high levels of transitivity, reciprocity, and betweenness centrality) positively influences a team's level of TMS. I believe that these previous findings will generally hold true where small teams are concerned. More specifically, I believe that high levels of strong-tie density and high levels of strong-tie transitivity in networks related to TMS will correlate with team performance and thus indicate high levels of TMS.

In using strong-tie density and strong-tie transitivity as measures of TMS a number of assumptions are made regarding the way ties are formed and networks are structured. One important assumption is small network size mitigates the risk of redundancy of information sharing. Past research has demonstrated that dense networks may lack in number of external ties, may share repetitive information, and may not lead to high levels of creativity (e.g. Lee, Bachrach, & Lewis, 2014). This research assumes that a small team size protects against the sharing of repetitive information.

A second assumption that is made in using strong-tie density and transitivity as measures of TMS is that external ties, or ties to individuals not on one's team, does not impact the proposed TMS network measures. Individuals are limited in the number of strong ties they develop. It could be argued that in forming ties with beyond one's team, within-network density and transitivity may be affected. This work assumes that external ties do not impact a team's TMS.

Demonstration of TMS Tie Formation

This demonstration serves two purposes. First, it illustrates the need to consider different patterns of tie formation in different types of networks. Second, it shows how tie formations in different networks emerge over time to reach a relatively steady state or equilibrium of ties.

In SNA, a number of different network ties may be evaluated for individuals on a team. For example, past studies have evaluated communication, collaboration, trust, and information retrieval networks for the same teams of individuals (Contractor & Su, 2011; Katz et al., 2005). In the proposed study, a number of different networks are evaluated for each team of participants. These "multiplex" networks will provide information related to each team's TMS.

The influence of parameters previously theorized to contribute to ties in two networks related to TMS will be illustrated below through step-by-step demonstrations of how individual, team, and contextual attributes influence the formation of ties in expressive (e.g. enjoyment) and instrumental (e.g. coordination) networks. Enjoyment ties may be good indicators of TMS as these ties reflect some of the process elements of TMS (e.g. beliefs about credibility). Coordination ties reflect the structural elements of TMS that enable teams to coordinate and perform tasks.

Instrumental networks are based on ties that are goal-oriented or task-based. Examples of instrumental networks include information seeking or coordination networks. Expressive ties, on the other hand, are emotionally based. Expressive networks may demonstrate ties including friendship, love, or trust.

According to Palazzolo and colleagues (2006), the emergence of well-functioning TMS is complex and nonlinear and it is “extremely difficult to mentally construe interdependent implications over time” (p. 233). By providing a step-by-step illustration of the drivers of network ties, this demonstration articulates the attributes that contribute to TMS, the relations among individuals on teams, and the co-development of these attributes and relations over time.

This demonstration of tie formation is motivated by a quantitative direct approach (Kozlowski et al., 2013) that models system dynamics of TMS emergence. This is meant to illustrate how various networks evolve from having no ties to many ties. For a more rigorous demonstration of the quantitative direct approach, please see Kuljanin (2011). In the present demonstration, I illustrate how a number of different input factors (including frequency of interaction, preferences for interaction, and behaviors upon successful vs. unsuccessful interactions) interact to result in network ties.

There are a number of individual-level, team-level, and contextual mechanisms that account for the emergence of ties in networks related to TMS. Individual-level mechanisms include KWKW, preference for collaboration, the success of previous interactions, the personalities of various team members, team member preference for collaboration, and principles of reciprocity and transitivity (Kozlowski et al., 2013). Team-level mechanisms that may influence TMS but are static throughout the team's time working together include functional background diversity, shared goals, shared team mental models, feasibility of task, and creativity required of task (Kozlowski et al.). Contextual mechanisms are the availability of new information, opportunities for repeated interactions, and constrained teams, or teams that are assigned to work with one another and do not work on project tasks with individuals outside of their team (Kozlowski et al.).

The influence of each mechanism may be different for different networks, especially for expressive versus instrumental networks. The emergence of ties in an expressive network (e.g. a trust network) is illustrated below in Figure 6. This figure shows how a network may evolve from exhibiting no strong trust ties between individuals to exhibiting many strong ties between individuals, thus, demonstrating high strong-tie density.

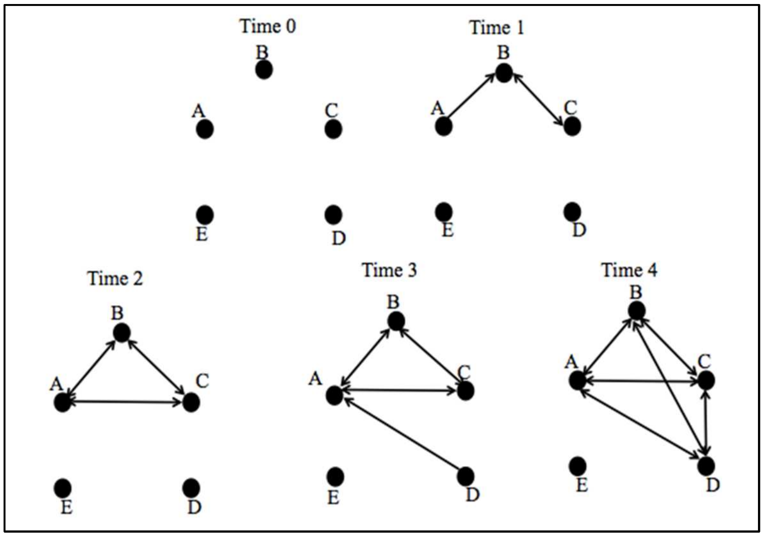


Figure 6. Trust tie formation

As is illustrated in Figure 6, at time 0 there are no trust ties. At time 1, individual A indicates that they trust individual B and individuals B and C indicate that they trust one another. These first ties are likely due to personality of individual team members and may reflect initial preferences for collaboration (early tie formation is discussed in detail in Chapter 3). At time 2, the unidirectional tie between individuals A and B becomes reciprocal. This is due to the tendency for expressive ties (e.g. enjoyment, trust, friendship) to be bi-directional, based on the principle of reciprocity. Also at time 2, a tie forms between individuals A and C. This is due to the principle of transitivity, which indicates that if one individual indicates a tie to a second individual and that second individual indicates a tie to a third individual, it is likely that the first and third individuals will also develop a tie. Transitivity is regularly present in expressive networks. At time 3, individual D indicates a trust tie with individual A. Again, due to principles of reciprocity and transitivity, at time 4, there are reciprocal ties between all individuals in the network save individual E who remains an isolate.

Figure 7 provides an illustration of coordination ties between individuals in a network. While this illustration has some similarities to the trust network described above, there are important differences based on the instrumental nature of these ties.

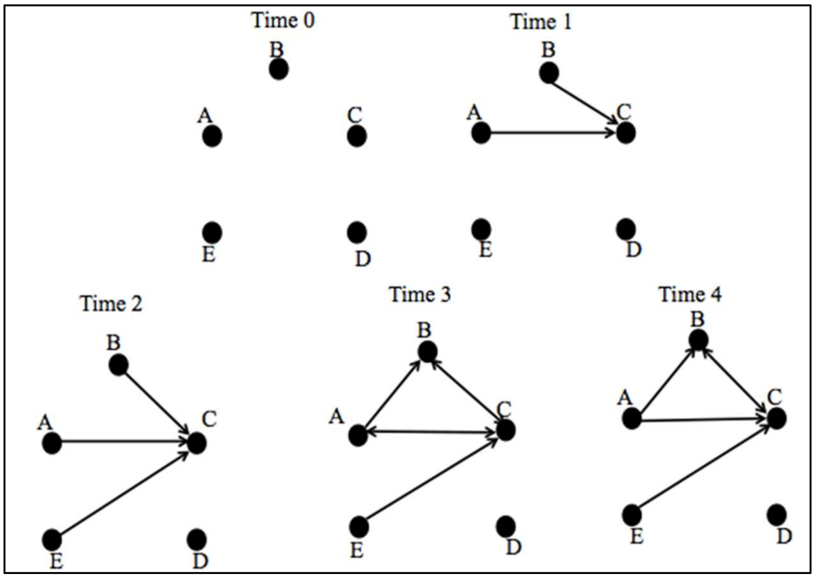


Figure 7. Coordination tie formation

As illustrated above, at time 0, there are no ties in this network. At time 1, individuals A and B indicate that they choose to coordinate with one another. Again, these early ties are most likely due to factors such as team member personality or availability (see Chapter 3 for a more detailed discussion of these early ties). At time 2, the previous ties still exist, however, now individual E also indicates that they choose to coordinate with individual C. At time 3, individual E now also chooses to coordinate with individual C and individual C indicates that choose to coordinate with individual A. At time 4, individual C no longer chooses to coordinate with individual A.

Several points emerge that differentiate this instrumental network from the expressive network previously described. For example, ties can be directed rather than bi-directional. This is due to the fact that there is less social pressure to indicate instrumental ties to those who indicate instrumental ties with you. This stands in contrast to the principle of reciprocity that was previously described and evidenced in the trust network illustration. Additionally, the principle of transitivity does not hold. In this case, just because two individuals choose to coordinate with a third individual does not mean that those two individuals are more likely to coordinate with one another. This is, again, due to lower social pressure in instrumental networks. Finally, previously existing ties may be removed over time. Although this is still somewhat unlikely, it happens more regularly in instrumental networks than in expressive networks.

Summary, Recommendations, and Research Questions

Over the past three decades, TMS has received growing attention from researchers in many different fields. From its roots in social psychology, TMSs have been evaluated in a number of contexts and through many measurement approaches. Although this is a positive trend that has encouraged a deeper understanding of the structures and processes teams use to share information, there are opportunities to more accurately measure a team's TMS.

First and foremost, the impact of analyzing TMS at the team-level by aggregating individual responses about team functions should be considered. In doing this, information may be lost and statistical power may be reduced. Using SNA to measure TMS has the potential to improve upon previous techniques. First, the way that data are collected and analyzed under SNA focuses directly on the relationships and behaviors of individuals on teams. This approach strengthens the content validity of obtained measurements by assessing behaviors at the individual (within-team) level. If data related to various networks (e.g. trust, enjoyment, coordination) are collected, these within-team relationships and behaviors may provide direct insight into the structures and processes that comprise a TMS.

Additionally, as noted by previous researchers (e.g. Lee et al. 2014; Lewis & Herndon, 2011), many of the attributes studied under SNA are closely related to TMS. For example, network reciprocity and network density have been shown to influence TMS, which, in turn, would be expected to influence team performance (Lee et al.).

Although Lee and colleagues found a negative relation between density and performance (mediated by transitive triads), I believe that these results may not hold true in other contexts. In small teams, density is bound to be very high. I believe that more interesting and insightful than measuring density as previously defined (the number of total present ties divided by the number of total possible ties) would be measuring *strong-tie density* (the number of total present strong ties divided by the number of total possible strong ties). More specifically, I argue for the measurement of TMS using the density of strong TMS network ties. This is a new approach to TMS measurement.

While my conceptualization of strong-tie density is unique, the impact of strong versus weak ties is a concept that has been evaluated by several SNA studies over the past four decades (e.g. Burt, 2001; Granovetter, 1973; Uzzi, 1996). Strong ties are typically found between people who interact often and exhibit trust in one another whereas weak ties are characterized by infrequent contact and greater fragility (Prell, 2012). Tie strength, therefore, has several dimensions including frequency of contact and robustness to external influence.

Most theories about strong versus weak ties have been conducted in the context of large networks. According to Granovetter (1973), weak ties may be more useful than strong ties in that they prevent the sharing of redundant information. This is in contrast with findings of Uzzi (1996), which indicate that strong ties lead to better problem solving and higher-quality decision-making. TMS theory, based on the belief that a well-functioning TMS is built on repeated interactions, especially among individuals on small teams, is more in line with Uzzi's belief that strong ties may lead to greater levels of TMS and higher levels of team performance. Therefore, an evaluation of the density of strong ties in TMS networks (e.g. trust, enjoyment, coordination) could prove to be useful as a measure of TMS, even in small networks. Under this conceptualization, moderate or weak ties would not be considered.

Using the density of expressive and instrumental network ties as measures of TMS contributes to TMS measurement theory in several ways. First, this is the first study to introduce and provide a network instrument to measure TMS. This encourages the future use of SNA in measuring TMS and the focus of TMS measurement to be on ties between individuals rather than aggregated perceptions of the team's functioning.

Second, the distinction between density and strong-tie density could be useful in the measurement of other team constructs, especially where small teams are concerned. To date, a majority of small group research has not used social network analysis to evaluate constructs of interest. This is because there is typically very little variance in the total number of ties present in networks of small teams. For example, in teams of 4 or 5, nearly all members indicate information-sharing ties as the size of their group requires at least some level of relation among individuals. By considering strong ties rather than all ties, differences between networks may be noticed and interesting findings may be obtained. For example, perhaps previously noted relations between reciprocity, transitivity, and performance will no longer hold.

In this study, I collect and analyze data to evaluate how the strong-tie density and transitivity of expressive networks (e.g. enjoyment, trust) and instrumental networks (e.g. coordination) may demonstrate the strength of a team's TMS and may influence team performance. In doing this, I aim to answer six research questions and explore a number of related hypotheses:

1. Is there evidence that the proposed network measures correlate with team performance, overall and after controlling for team-level attributes?

   Hypothesis 1A: The network measures correlate positively with performance

   Hypothesis 1B: Of the proposed network measures, measures of density will have a stronger relation with performance.

2. Do the proposed network measures demonstrate external aspects of validity, including convergent and discriminant validity?

   Hypothesis 2A: Measures will correlate more strongly with teamwork than professionalism.

3. Do the proposed network measures demonstrate generalizability?

   Hypothesis 3A: Correlations between the proposed network measures and team performance will hold across different types of teams (multidisciplinary and single-discipline).

4. Do the proposed network measures demonstrate sufficient test-retest reliability?

5. Do the proposed network measures demonstrate incremental validity over the measure proposed by Lewis (2003)?
6. Does the structural model proposed by Lewis (2003) demonstrate goodness of fit with a new dataset?

This work uses a definition of TMS that is informed by Liang et al. (1995) and Lewis (2003) who considers TMS to be “the cooperative division of labor for learning, remembering, and communicating relevant team knowledge.” In this project, TMS is defined as the processes and sentiments that lead teams to work together to accomplish shared goals. It should be noted that this definition assumes that team members have shared definitions of success. Furthermore, this definition relies more strongly on the affective nature of teamwork, including such attributes as trust and enjoyment among team members. Limitations of using this definition of TMS are described in Chapter 6.

Chapter 3: Pilot Work

It is important to obtain measures of how a team is functioning before performance assessments are made. This allows for early identification of high-performing teams and for interventions to be made in instances where teams are not on track for success. For this reason, a number of measures of TMS have been proposed and employed. The higher-order factor model and method of score aggregation presented in Chapter 2 (Lewis, 2003) has been the most widely adopted conceptualization of TMS and its corresponding TMS inventory (Lewis, 2003) has been highly recommended for use in future studies (Ren & Argote, 2011).

In order to evaluate the performance of this inventory and provide direction for future work in this area, I conducted a pilot study. The goals of this pilot study were to determine whether responses to Lewis's (2003) inventory provided a valid measure of TMS and evaluate if there were additional measures of TMS that related to team performance. This was accomplished by fitting the higher-order factor model presented in Chapter 2 to a new dataset and comparing findings regarding factor loadings and model fit between the new data and information reported by Lewis. The relation between TMS and team performance, as well as the relation of other attributes of the TMS data (e.g. within-team variance) and team performance, were examined to evaluate the external aspect of validity of the proposed instrument.

Additionally, an inventory was created to capture how the three previously-studied aspects of TMS, specialization, credibility, and coordination (Liang et al., 1995; Lewis, 2003), would manifest in practice. This inventory was used to collect data from teams of students. Using these data, calculations of newly proposed network operationalizations of TMS, e.g. strong-tie density, were made for a number of networks related to TMS. Strong-tie density of these networks was then compared to team performance, TMS composite scores from Lewis's inventory, and within-team TMS variance from Lewis's inventory to preliminarily evaluate the validity of the use of the strong-tie density of various TMS networks to measure TMS.

In addition to evaluating the relation between strong-tie density and performance, the relation between other network attributes and performance was also evaluated. These attributes include the reciprocity of enjoyment, trust, and efficiency networks, the average indegree centrality of enjoyment, trust, and efficiency networks, and the transitivity of enjoyment, trust, and efficiency networks. Correlations between these attributes and performance were considered to evaluate previously proposed relations between TMS networks and performance (e.g. Lee et al., 2014; Mell et al., 2014,).

Qualitative data were also collected as part of this pilot study. These data were collected for two purposes. First, information provided by participants was compared with TMS theory to evaluate if previously theorized relations between individuals on teams held in the context of the proposed study. Second, participants provided feedback on the proposed measurement instrument that enabled for both instrument revision and evaluations of content and substantive aspects of validity.

This quantitative and qualitative pilot work provided motivation for this dissertation; the continued study of how TMS may be better conceptualized and measured and how these new measures of TMS relate to team performance.

Method

Measures

Lewis's (2003) 15-item TMS inventory was used. Based on recommendations from Lewis's study, two credibility items were reworded so as to not require reverse coding. Otherwise, the items remained exactly the same as those asked in the original Lewis inventory (provided in Appendix A).

Additionally, participants responded to a five-item inventory that obtained information about four networks and one attribute related to TMS (manifestations of specialization, credibility, and coordination including information seeking, trust, efficiency, enjoyment networks and a complementary skills attribute). These items were chosen based on their relation to theorized aspects of TMS as well as their similarity to other networks typically analyzed in SNA. For example, the item "I trust this person" provides information related to credibility and trust networks are frequently evaluated by network researchers.

Each team member answered these five questions for each of their teammates. In SNA, network data are often gathered by obtaining information about ties of individuals within a fixed boundary. In this case, the boundary was the assigned team. All questions were asked to generate information about ties at the individual level. Names were programmed into the survey instrument to ensure that complete data were captured. All items followed a five-point Likert scale with response options ranging from (1) Strongly Disagree to (5) Strongly Agree. Items are provided below and in Appendix D. Newly proposed attributes of strong-tie density, reciprocity, transitivity, and indegree centrality (scores of 4 or 5) for each of these networks were considered as possible measures of TMS strength.

- I regularly seek information from this person (*Information Seeking*)
- I trust this person (*Trust*)
- This person and I work together in an efficient manner (*Efficiency*)
- I enjoy working with this person (*Enjoyment*)
- I have complementary skills to this person (*Complementary Skills*)

It should be noted that, while gathering information about ties between individuals, many of these questions elicit different information than is obtained in typical social network studies. Network studies usually evaluate either expressive (e.g. friendship) or instrumental (e.g. advice) networks (Fang et al., 2015). Of the above questions, the Trust and Enjoyment items may be considered to relate to expressive networks and Information Seeking and Efficiency items may be considered to elicit information about instrumental networks. The Complementary Skills item provides attribute rather than network data which influences the way in which these data were analyzed. Rather than providing information about Specialization networks, these data are more useful in describing features of individuals in the other networks.

While responses to the expressive network items should follow patterns expected of social network studies (e.g. reciprocity, transitivity), many of the expectations typically held for patterns of network behavior might not apply to the other items. For example, Information Seeking behavior may not demonstrate the patterns of reciprocity that are often expected in social network studies. It would not be uncommon to find individuals on a team who are often on the receiving end of Information Seeking ties but who do not reciprocate by seeking information from their peers. This could lead to different findings about network structure than observed in previous studies. For example, perhaps the presence of one individual holding high betweenness centrality does not positively influence team performance, as found by other researchers (Tröster, Mehra, & van Knippenberg, 2014).

At the end of the semester, course instructors and teaching assistants evaluated team performance. This evaluation asked for a summative rating of the team's performance using a five-point scale (provided below and in Appendix D). Two ratings from the previously described evaluators were obtained for each team and those ratings were summed to provide an overall performance evaluation.

- 5 - Outstanding; The team is polished, thorough, and on par with what could be expected from a team of professionals
- 4 - Good; The team exceeds expectations for a team of honors students
- 3 - Fair; The team meets expectations for a team of honors students
- 2 - Poor; The team falls short of expectations for a team of honors students but still provides satisfactory work in some areas
- 1 - Unacceptable; The team does not accomplish satisfactory work at any level

Procedures

Data were gathered from two multidisciplinary undergraduate courses (cross-listed by business and engineering departments) where teams spent a semester (15 weeks) working on team-based projects. Both inventories were delivered concurrently approximately halfway through the semester (during week 8 of the 15-week class) through an online survey in the Spring of 2016. All participants provided informed consent prior to participating.

One of the courses was at the sophomore level; the other course was junior level. In the sophomore-level course, nine teams of four to five students were formed to maximize within-team functional background diversity. Functional background diversity, or FBD, represents differences in knowledge, information, and perspective relevant to a particular task (Richter et al., 2012). There were 40 total student participants in this course. In the junior level course, 12 teams of four to five students were formed based on both maximizing functional background diversity and maximizing student project preference (students indicated which projects they preferred prior to being assigned). There were 43 total students participants in this course. Student teams in both courses were tasked with completing projects that required creativity and functional background diversity of team members. Data were only included in cases where at least three members of the team responded to the survey. This was true for 19 of the possible 21 teams. Individuals who did not provide responses were excluded from the analysis.

Analysis

Data gathered using Lewis's (2003) inventory were fit to the higher-order factor model described in Chapter 2. Mplus version 7.4 (Muthén & Muthén, 2012) was used. To account for the multilevel nature of the data of students within teams, maximum likelihood estimation with robust standard error corrections was employed with the use of TYPE=COMPLEX. Data gathered using the newly proposed social network inventory were analyzed using the R package statnet (Handcock et al., 2006). This package was used to analyze and display a number of networks for each team. All other calculations, including analyses of external validity (e.g. correlations between scale scores and performance) and calculations of density and within- and between-team variances, were performed in R (R Core Team, 2015).

As previously stated, valued data about four networks related to TMS and one TMS attribute (complimentary skills) were collected. These data provided information about the strength of ties between individuals. In small networks, ties (especially instrumental ties) are typically present between all individuals. For example, based on the limited number of alternatives in a small team, it is likely that everyone seeks information from everyone at some point. Therefore, rather than considering the overall network density, the density of strong network ties was used. Ties in this network were dichotomized as follows: ties receiving scores of 4 or 5 were coded with values of "1" and all other ties (receiving scores below 4) were coded with values of "0". Therefore, the networks analyzed in this study were based on binary data. The attribute providing this data, strong-tie density, was created for use in this study.

Results

Before analyzing the parameter estimates for Lewis's (2003) TMS Inventory, it should be noted that the model fit indices (provided in Table 1) indicate poor fit between the collected data and the higher-order factor model presented in Chapter 2. All fit indices obtained after fitting the higher-order factor model presented in Chapter 2 to the collected data fall beyond the thresholds of acceptable fit, as suggested by commonly used fit guidelines (e.g. Hu & Bentler, 1999). It should be noted, however, that the sample size used in the present analysis (N=19) is quite small considering the number of estimated parameters (48), which could lead to instability of fit information (Hu & Bentler). Factor loadings, error variances, and factor disturbances are provided in Appendix E.

Although the model fit results presented in Lewis (2003) demonstrate slightly better fit (see Table 1), it should be noted that, based on recommended standards for fit indices, the higher-order factor model presented in Chapter 2 should be viewed with hesitation for these data. The RMSEA, reported for the newly collected data, is an important index that evaluates the extent to which a model fits "reasonably well" compared to fitting perfectly while accounting for model parsimony (Brown, 2006). Although no parsimonious fit indices were reported by Lewis, the RMSEA may be calculated using the provided information as $RMSEA = \sqrt{\frac{\frac{\chi^2}{df}-1}{N-1}}$ = 0.09. This index, greater than the suggested maximum of 0.06 as indicating "good fit" (Hu & Bentler), provides some concern regarding the structural validity of Lewis's findings. Additionally, the SRMR reported by Lewis falls above the recommended figure of 0.08 and the CFI reported by Lewis falls below the recommended threshold of 0.95 (Hu & Bentler). This, again, means that the higher-order factor model presented in Chapter 2 does not demonstrate "good fit" (Hu & Bentler) with Lewis's data.

Table 1

Model fit indices from Lewis (2003) and pilot study

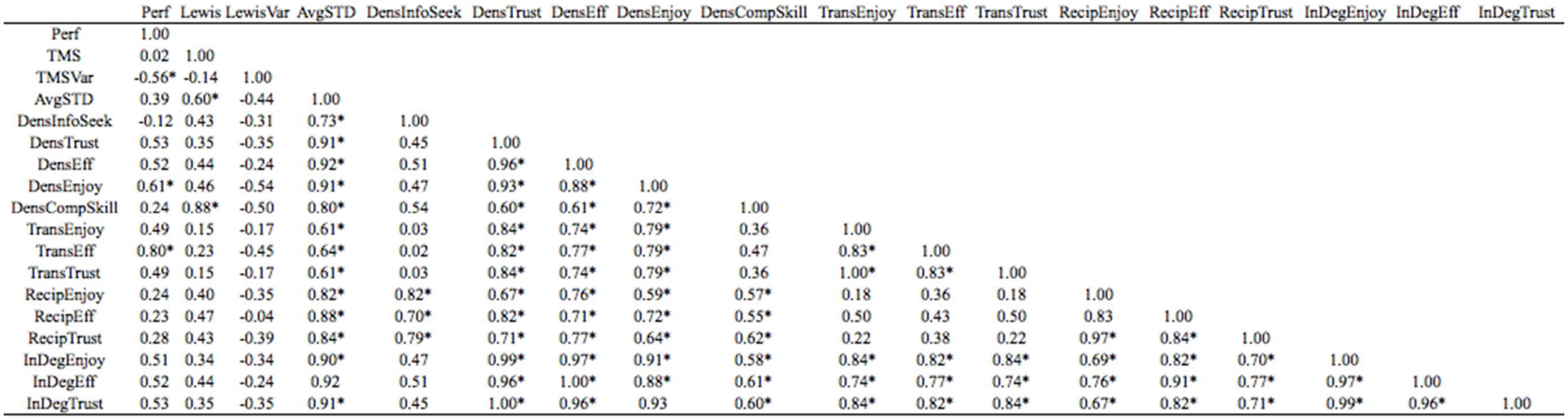
	Lewis, 2003	Present Study
N	260	43
RMSEA	0.09	0.14
CFI	0.93	0.83
SRMR	0.09	0.10

Beyond model fit, other findings reported by Lewis (2003) were not replicated in this pilot study. These findings relate to the convergent validity of the proposed measure. First, the minimum within-team correlation of 0.7, reported by Lewis as necessary to justify aggregating results to the team level, was not observed (average within team correlation ranged from 0.51 to 0.62 for the three factors). There was typically more variance found within teams than between teams (the percent of within-team variance ranged from 49 percent to 81 percent).

Additionally, the moderately strong correlations that were found by Lewis (2003) to exist between the TMS composite score (the aggregated score for each team as measured using Lewis's [2003] inventory) and team performance (from 0.34 to 0.53) were not observed in the present study; all correlations are provided below in Table 2. The correlation between TMS composite score (Lewis) and performance (Perf) was 0.02, which is considered to be quite weak. This demonstrates that Lewis's measurement approach does not adequately predict team performance in the current setting.

Table 2

Correlation matrix of TMS measurement outcomes and performance (N=19)



	Perf	Lewis	LewisVar	AvgSTD	DensInfoSeek	DensTrust	DensEff	DensEnjoy	DensCompSkill	TransEnjoy	TransEff	TransTrust	RecipEnjoy	RecipEff	RecipTrust	InDegEnjoy	InDegEff	InDegTrust
Perf	1.00																	
TMS	0.02	1.00																
TMSVar	-0.56*	-0.14	1.00															
AvgSTD	0.39	0.60*	-0.44	1.00														
DensInfoSeek	-0.12	0.43	-0.31	0.73*	1.00													
DensTrust	0.53	0.35	-0.35	0.91*	0.45	1.00												
DensEff	0.52	0.44	-0.24	0.92*	0.51	0.96*	1.00											
DensEnjoy	0.61*	0.46	-0.54	0.91*	0.47	0.93*	0.88*	1.00										
DensCompSkill	0.24	0.88*	-0.50	0.80*	0.54	0.60*	0.61*	0.72*	1.00									
TransEnjoy	0.49	0.15	-0.17	0.61*	0.03	0.84*	0.74*	0.79*	0.36	1.00								
TransEff	0.80*	0.23	-0.45	0.64*	0.02	0.82*	0.77*	0.79*	0.47	0.83*	1.00							
TransTrust	0.49	0.15	-0.17	0.61*	0.03	0.84*	0.74*	0.79*	0.36	1.00*	0.83*	1.00						
RecipEnjoy	0.24	0.40	-0.35	0.82*	0.82*	0.67*	0.76*	0.59*	0.57*	0.18	0.36	0.18	1.00					
RecipEff	0.23	0.47	-0.04	0.88*	0.70*	0.82*	0.71*	0.72*	0.55*	0.50	0.43	0.50	0.83	1.00				
RecipTrust	0.28	0.43	-0.39	0.84*	0.79*	0.71*	0.77*	0.64*	0.62*	0.22	0.38	0.22	0.97*	0.84*	1.00			
InDegEnjoy	0.51	0.34	-0.34	0.90*	0.47	0.99*	0.97*	0.91*	0.58*	0.84*	0.82*	0.84*	0.69*	0.82*	0.70*	1.00		
InDegEff	0.52	0.44	-0.24	0.92	0.51	0.96*	1.00*	0.88*	0.61*	0.74*	0.77*	0.74*	0.76*	0.91*	0.77*	0.97*	1.00	
InDegTrust	0.53	0.35	-0.35	0.91*	0.45	1.00*	0.96*	0.93	0.60*	0.84*	0.82*	0.84*	0.67*	0.82*	0.71*	0.99*	0.96*	1.00

*Significant at $\alpha = 0.10$

This could mean that (1) the higher-order factor model presented in Chapter 2 does not adequately represent TMS, (2) the aggregation across individual scores described in Chapter 2 leads to a misrepresentation of TMS, (3) the survey items do not adequately tap the constructs of interest, or (4) TMS and team performance are not always related. Based on the theoretical underpinnings of TMS, I believe that the reason for the lack of correlation between TMS and performance is likely due to problems with the previously used method of aggregating individual scores to the team level.

This relation between TMS composite score and team performance is very important. In addition to the evidence of external aspects of validity that are provided by a strong correlation between TMS and performance, reporting TMS as a stand-alone descriptor of a team is oftentimes not very meaningful. As previously stated, TMS is not a widely recognized team construct, therefore, a substantial amount of practical value comes from relating TMS to team outcomes, including team performance. A strong correlation between TMS composite score and performance was not demonstrated in the present analysis.

Beyond illustrating the relation between TMS composite score and team performance, the correlation matrix provided above in Table 2 illustrates the relation between all outcome measures of interest in this pilot study. It should be noted that the small number of teams considered in this analysis ($N = 19$) yields results with limited power. Despite considerations of relatively low power, several interesting findings emerge. The correlation between within-team TMS composite score variance (LewisVar) and performance (Perf) was much stronger than that of TMS composite score and performance (r=-0.56 versus r=0.02), suggesting that there is a negative relation between within-team TMS variance and performance, or that teams with greater agreement about their levels of TMS outperformed those with less agreement.

Correlations between strong-tie density, reciprocity, transitivity, and indegree centrality in individual networks (e.g. trust network, enjoyment network) and team performance were also calculated. Additionally, the average strong-tie density of the five networks (the mean strong-tie density of all five networks) was also correlated with within-team variance and team performance. The correlation between students reporting strong ties (values of 4 or 5) in the evaluated networks and team performance were typically of moderate to high strength, as is illustrated in Table 2, although it should be noted that statistical significance was not reached in many of these cases.

The results presented in Table 2 motivate the continued study of how network data may be used to measure TMS and how strong-tie density in networks may relate to performance. Perhaps the most important result obtained from this pilot analysis is that aggregated TMS results from the previously proposed inventory (Lewis, 2003) did not demonstrate any relation to performance (r= 0.02). While the relations between the proposed networks and TMS were also not statistically significant at the level of α =0.05, this could be due to the small number of teams evaluated in the present study. By incorporating more teams in the analysis, it is likely that similar strengths of relations carrying statistical significance would be obtained. Additionally, it is likely that by improving upon the wording of the proposed TMS network questions, the relation between networks related to TMS and performance would be strengthened. The current inventory provided many responses of 5. By updating the questions to include strongly worded stems this could result in more representative data being obtained.

While data were consolidated for all networks to provide an overall measure of the networks' collective relation to TMS composite scores and performance, data were also analyzed for each network individually. Analysis of individual networks was purposeful as one of the goals of this study was to understand how ties in different types of networks related to performance. While previous TMS studies aggregated results of specialization, credibility, and coordination items to get an overall TMS score, I believe that a preliminary study of network data related to TMS should consider responses to items individually in addition to aggregating these data.

A graph depicting the relation between enjoyment networks (responses to the item "I enjoy working with this person") and performance of 10 teams of students in the junior-level course is provided in Figure 8. This figure includes only 10 teams as it does not include information about teams in the sophomore level course. As can be seen in this graphic, and as is demonstrated in Table 2, there appears to be a relation between performance and density of the TMS enjoyment network ($r = 0.61$, $p = 0.06$). For example, the worst performing team (with a score of 4) has the least dense enjoyment network, where density is equal to the number of present ties divided by the number of possible ties (Δ=0.58). The best performing teams (with scores of 10) have densities equal to 0.67 and 0.83. Similar patterns of high densities for high performing teams were noticed for each of the social network items.

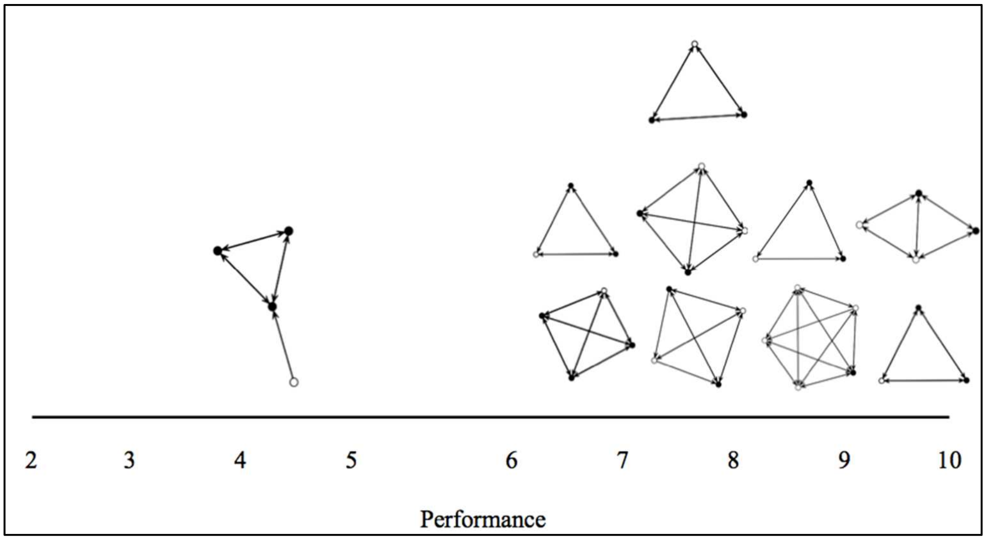


Figure 8. Enjoyment network ties and performance

As is evidenced by the relations described and shown in Table 2, there were strong correlations present between trust and efficiency networks and the Lewis (2003) inventory composite score (r_{trust}= 0.53; $r_{efficiency}$= 0.52). These correlations are much higher than those observed between Lewis's inventory and team performance. These findings, therefore, promote the continued evaluation of whether networks related to trust, enjoyment, and efficiency (or coordination) may be used in measuring TMS.

Results of an analysis of the relation between other network attributes (e.g. reciprocity and average indegree centrality) and performance support some previous findings about these attributes. For example, average indegree centrality may be related to team performance in enjoyment, trust, and efficiency networks ($r_{enjoyment}$ = 0.51; r_{trust} = 0.53; $r_{efficiency}$ = 0.52). As reviewed in Chapter 2, Mell and colleagues (2014) evaluated the relation of betweenness centrality and performance. Indegree centrality was used in the present analysis. I believe that, in small teams, indegree centrality will be a better measure of TMS as coordinated ties are not critical based on the small number of team members. Furthermore, this attribute is easier to scale for different sized teams (average indegree centrality is equal to the average number of ties received by each member of the team).

The relation between average indegree centrality and performance is quite similar to that of strong-tie density and performance. This is somewhat expected as the calculations of indegree centrality and density are based on similar network attributes (the number of ties). I believe, however, that network density is a better overall measure of a network as this is based on all network ties (versus incoming only), does not require scaling, and is more regularly used in network research. Therefore, I believe that strong-tie density is a better measure of TMS than average strong-tie indegree centrality and that strong-tie density should be used in future research.

The relation between reciprocity and performance, noted by previous researchers (Lee et al., 2014) was not found in this pilot study as correlations between reciprocity and performance for all networks were all rather weak and not significant ($r_{enjoyment}$ = 0.24; r_{trust} = 0.28; $r_{efficiency}$ = 0.23). While this relation is not expected in instrumental networks (e.g. efficiency), network theory suggests that reciprocity will be present in expressive networks (e.g. trust and enjoyment). These results suggest that reciprocity of strong network ties may not be a good measure of TMS.

A moderate to strong relation between transitivity and performance, however, was noticed in the present study ($r_{enjoyment}$ = 0.49; r_{trust} = 0.49; $r_{efficiency}$ = 0.80). While consistent with other findings, it should be noted that there was not great variability in the transitivity of strong-tie networks among the teams. For example, only one team's strong-tie transitivity of the efficiency network was less than 1. Future work, described in Chapters 4 and 5 of this document, evaluates whether the moderately strong to strong correlations noticed in this pilot study hold when data on more teams are obtained and when there is greater variability in network transitivity.

It should also be noted that the only network attributes that had a significant relation with Lewis's (2003) TMS composite score were the strong-tie density of the complimentary skills network and the average strong-tie density of all analyzed networks. This indicates that the newly proposed networks provide data that is distinct from that reported by Lewis (2003), which has been demonstrated to lack external aspects of validity (based on its relation with team performance in the present analysis).

Although only a small number of teams were analyzed (*N*=19), I believe that the relation between strong-tie density of networks related to TMS and performance motivates a more in-depth exploration of how these networks (networks related to enjoyment, trust, and coordination) capture important information about team processes that may serve as early indicators of team performance. To improve upon the original network items, a qualitative study was conducted to evaluate proposed changes to the TMS network inventory and to better understand the individual, team, and contextual characteristics that underlie the formation of various network ties related to TMS.

Qualitative Data Analysis

As part of the pilot work for this project, qualitative data were gathered for two purposes. First, cognitive interviewing enabled a modified network inventory to be evaluated. Second, in-depth interviews provided rich and contextual information about the way individuals function within teams. This information was compared with the illustration of network tie formation provided in Chapter 2.

Interviews were conducted with a subsample of students (*n*=6) who had recently completed the junior-level course previously described. All participants completed team projects requiring creativity and innovative solutions. As noted earlier in this chapter, teams were composed to maximize student project preference and within-team functional background diversity.

Procedures and Analysis

Data were gathered through six 30-minute interviews. Interviewees were participants in the empirical study previously described, selected from a pool of those who expressed an interest in continuing to be involved with the research project. Participants were chosen through purposeful sampling (Marshall & Rossman, 2014; Merriam, 1998; Patton, 2002), which identified information-rich cases of individuals who exhibited a range of experiences working on their team (very negative to very positive).

Recommended guidelines for qualitative research were followed, including using clear, open-ended questions, carefully selecting the question order, maintaining neutrality, quickly transcribing data, comparing findings to other research on the topic, and analyzing negative cases, or cases that were not line with emerging data patterns (Caudle, 2004; Schuh & Upcraft, 2001). Additionally, the validity of qualitative findings was enhanced by involving participants in the verification of data and keeping explicit records of all aspects of data gathering and analysis (Marshall & Rossman).

The interview protocol, provided in Appendix F, also followed recommended guidelines of informing participants of the purpose of the study, explaining that their participation would be voluntary and that results would be confidential (Simone, Campbell & Newhard, 2012). Immediately following the interview, a number of important findings were noted, including the most important themes and ideas and how these ideas differed from what was expected (Morgan & Krueger, 1998).

Results

Cognitive interviews were conducted with six students who had taken the previously described TMS network inventory to obtain feedback on proposed changes. Feedback from these interviews include that the proposed TMS network questions are "straightforward and easy to understand" (Participant A, personal communication, July 13, 2016). The train of thought followed by the participants in reviewing the questions was in line with what was expected. For example, for the specialization item, participants thought about differences in educational backgrounds and experiences and the degree to which a teammate held specific knowledge about a portion of the project. Participants noted that, relative to the social network inventory used in the previously described pilot study, in using the updated inventory, they would be more likely give different scores to different team members for all questions, more accurately representing the true nature of ties between themselves and their teammates. This is an improvement over the previously described TMS network inventory, on which many participants provided the same, typically high, scores for each team member.

According to the six participants, motivation to collaborate has a direct influence on the likelihood of an interaction occurring between teammates. The reported likelihood of a subject choosing to collaborate had a wide range, with individuals indicating that they choose to collaborate from 30 percent to 99 percent of the time. Furthermore, the reported motivation to collaborate with specific individuals changed as the project progressed. All six individuals reported starting with the same motivation for collaboration across each of their teammates (regardless of factors such as FBD or gender). Changes in motivation were due to whether or not past collaborations were positive or negative. Positive interactions were noted to lead to higher perceived probabilities of choosing to collaborate in the future and negative interactions were reported to lead to reduced likelihoods of future collaboration.

Participants generally indicated that belief in teammate credibility changed minimally throughout their interactions. For example, one participant noted that there were “no changes in credibility, just reduced comfort in working with some team members” (Participant B, personal communication, July 20, 2016). This is in contrast to existing TMS theory that describes belief in teammate credibility as a dynamic attribute of teams.

All participants indicated that the ability for individuals on a team to coordinate generally increased after each interaction, regardless of whether or not an interaction was successful. These increases were perceived to occur during early interactions and leveled off at a certain point. This dynamic nature of coordination is in line with TMS theory.

All participants noted that their knowledge of whom on their team knew what (KWKW) influenced their likelihood of choosing to collaborate. As participants described similar levels of KWKW across teammates, KWKW may be viewed as a team-level variable. Participants reported that the likelihood of collaboration increased early in the project and reached a steady state approximately halfway through the semester.

The presence of shared goals and shared team mental models has been found to lead to greater levels of TMS (Kozlowski et al., 2013). According to the six interviewed participants, the degree to which goals were shared was relatively consistent within teams. Individuals on different teams, however, did note a difference in the “sharedness” of their team’s goals. Implications of this possible asymmetry of team goals are reflected by previous literature on teams (e.g. Pearsall & Venkataramani, 2015).

Much of the gathered information supports existing TMS theory, although some responses are in contrast with previously theorized patterns of individual and team behavior. Personality and availability were cited as the reasons why individuals chose to collaborate with specific teammates, especially at the start of projects. The personality of teammates does not relate to any of the aspects theorized to underlie TMS but is more associated with expressive ties, such as those of trust or enjoyment. Availability relates more to coordination than credibility and specialization aspects of TMS, yet TMS theory posits that all three aspects contribute equally to TMS development (Lewis, 2003). These are issues that will be improved upon by my proposed measure of TMS.

Some of the divergence between these results and TMS theory should be considered in developing future instruments for TMS measurement. For example, credibility was not reported to be important to an individual's choice to interact with their teammates and credibility was noted to be relatively constant throughout the duration of the team project. This could motivate new measures of TMS to be used that are less reliant on credibility.

Summary, Recommendations, and Conclusions

I conducted a pilot study in which Lewis's (2003) 15-item inventory was used to measure the TMSs of undergraduate students working in small teams. Additionally, a five-item inventory was provided to the same students to obtain information related to four social networks and one network attribute. Results of this pilot study include that the previously proposed higher-order factor model of TMS (Lewis) did not demonstrate good fit with the collected data. Additionally, this pilot work found moderately strong to strong relations between strong-tie density of these social networks and performance, providing preliminary support that strong-tie density may serve as a valid measure of TMS.

The poor data-model fit, high within-team variance, and low correlation between Lewis's measure of TMS and team performance provide evidence against the validity of the method of aggregating individual scores to represent a team construct. These preliminary results provide ample motivation for continued study of how TMS may be better conceptualized and measured and how these new measures of TMS relate to team performance. More specifically, the relation between enjoyment, trust, and efficiency (or coordination) networks and team performance provides a promising avenue for continued research. The qualitative information described above provides further support for this work.

Chapter 4: Method

This dissertation evaluates the application of network analysis to TMS measurement and involves two elements. First, the development of a TMS network inventory and, second, a validity study of the proposed TMS measurement instrument. This chapter describes the proposed measurement instrument's development and deployment, providing information about the study's participants, the various outcome measures to be evaluated, procedures for data collection, and analyses undertaken, including a validity study. Limitations of this study and potential implications of possible findings are also discussed.

Previous researchers have stated the perceived value in incorporating a network perspective into the study of TMS (e.g. Austin, 2003; Contractor & Su, 2011; Katz et al., 2005; Palazzolo et al., 2006). By focusing on the "web of relationships" (Katz et al., 2005) that characterize teamwork, network analysis is an appropriate tool to measure a team's TMS. The study of expressive ties (enjoyment and trust) and instrumental ties (coordination) distinguishes this approach from previously proposed TMS measurement schemes (e.g. Austin, 2003; Hollingshead, 1998a,b, 2001; Lewis, 2003; Wegner et al., 1991).

Measurement Instrument Development and Validity Study

This dissertation evaluates the use of a network inventory to measure TMS followed by a validity study to evaluate the proposed instrument's construct validity. This study follows Messick's (1995) view of construct validity as a unitary concept encompassing content, substantive, structural, external, and consequential aspects of validity as well as generalizability. Described in detail later in this chapter, aspects of external validity and generalizability were directly evaluated. Other aspects of validity were considered in the design of the instrument and the evaluation of measures obtained from the instrument. These considerations are discussed in Chapter 6.

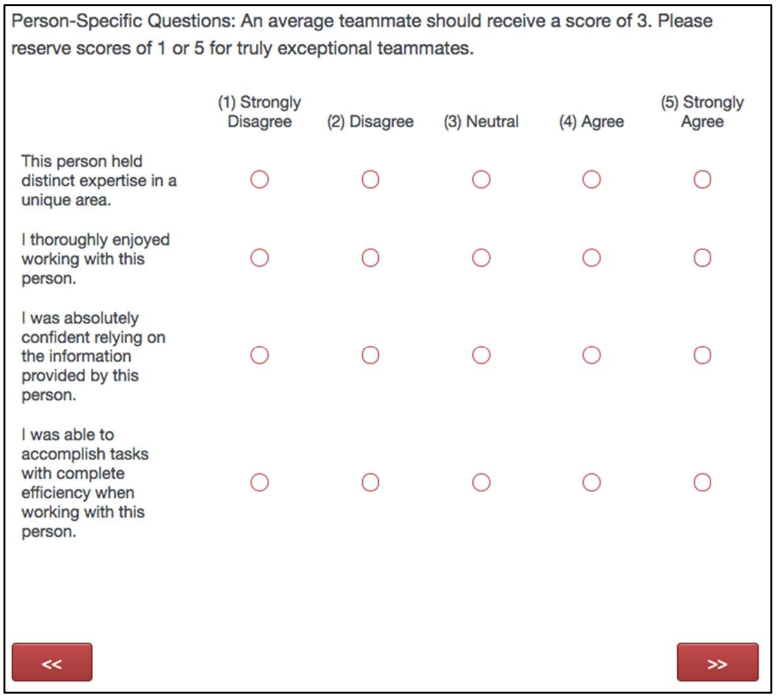
Participants

Participants in this study were enrolled in three undergraduate honors courses that required semester-long multidisciplinary team projects and one senior-level mechanical engineering capstone design course. Of a total 281 possible participants, 204 sophomore, junior, and senior students participated in this study. One hundred forty-eight participants identified as male and 56 identified as female. One hundred twenty-two identified as white, 51 identified as Asian, 15 identified as being of two or more races, 14 identified as black, and 2 identified as Hispanic (there were no other participant racial backgrounds). The range of student grade point averages was between 2.46 and 4.0 with a mean and median of 3.50. The participating students worked on 39 teams of three to seven students each.

While the maximum number of participants in this study was a constraint, the achieved sample size enabled reasonable statistical power to be obtained. Using G*Power (Faul, Erdfelder, Lang, & Buchner; 2007) to conduct a power analysis, the 95% confidence interval for the population correlation between the proposed network measures and team performance was anticipated to be approximately 0.60 ± 0.24. This assumes an effect size similar to that noticed in the pilot study (a correlation of approximately 0.60) with a level of power of 0.80, providing evidence of a moderate to strong effect. It should be noted that this confidence interval is sufficiently narrow to distinguish evidence of discriminant validity versus convergent validity, as described later in this chapter; however, the power for any partial correlation coefficients or interaction terms will be lower than the levels described above.

Measures

TMS Network Inventory. TMS was measured through the use of a TMS network inventory. A screen-shot of this inventory is provided below in Figure 9. Inventory items are also included in Appendix G. This inventory was informed by TMS theory (e.g. Wegner, 1987), social network theory (e.g. Wasserman & Faust, 1994), the results of previous TMS studies (Lewis, 2003), the pilot study described in Chapter 3, the cognitive interviews described in Chapter 3, and recommendations for the construction of survey questionnaires (Tourangeau et al, 2000).



Person-Specific Questions: An average teammate should receive a score of 3. Please reserve scores of 1 or 5 for truly exceptional teammates.

	(1) Strongly Disagree	(2) Disagree	(3) Neutral	(4) Agree	(5) Strongly Agree
This person held distinct expertise in a unique area.	○	○	○	○	○
I thoroughly enjoyed working with this person.	○	○	○	○	○
I was absolutely confident relying on the information provided by this person.	○	○	○	○	○
I was able to accomplish tasks with complete efficiency when working with this person.	○	○	○	○	○

Figure 9. Screen shot of TMS network inventory

As can be seen above, all questions follow a 5-point Likert scale with responses ranging from (1) Strongly Disagree to (5) Strongly Agree and are preceded by the following guidelines: “An average teammate should receive a score of 3. Please reserve scores of 1 or 5 for truly exceptional teammates.” Therefore, it is assumed that a participant responding with a score of 5 would strongly agree with the above statements.

Participants were asked to respond to the same four questions for each of their teammates. Teammate names were provided to each participant as they completed the survey (this information was entered into the survey tool Qualtrics 360 [Qualtrics, 2016] prior to administration of the survey instrument). This format of asking participants to provide all ratings for each team member one at a time was selected to reduce the cognitive requirements of rotating back and forth among all team members. Prior to administering the survey, I confirmed that there were no errors with the roster (e.g. incorrect lists of teams, students dropped from the course) by verifying lists of teams with instructors and teaching assistants.

Data were coded into binary responses, using a similar method to that described in Chapter 3. That is, only scores indicating agreement or strong agreement were coded with values of “1” as demonstrating strong ties. Scores of 3 and below were coded with values of “0”. This scheme of condensing ordered data into binary data was selected purposefully. The tendency of participants to provide overly generous ratings (evidenced in the pilot study described in Chapter 3 as well as in information obtained in qualitative interviews) would likely lead to an over-abundance of ratings of 1 if items simply asked whether or not strong ties existed. By coding values of 3 (neutral) as 0, I believe that the responses obtained reflected the true nature of the present ties.

Given that this is a new measure of TMS and that there is some uncertainty regarding the cutpoint of what distinguishes strong ties, in addition to using scores of 4 and 5 as indicative of strong ties, I also considered the use of scores of 5 alone in calculating strong-tie density and transitivity. Differences in the correlations between strong-tie density and performance using these different methods were evaluated, as described in Chapter 5 of this document.

Question 1, as shown in Figure 9, was used to generate specialization attribute data, question 2 was used to create enjoyment networks for each team, question 3 was used to create trust networks for each team, and question 4 was used to create coordination networks for each team. These questions are similar to but slightly different from the items asked in the pilot study for several reasons. First, they demonstrated the strongest relation to team performance based on my pilot data. Second, they provide information that may be easily analyzed under a social network framework. Third, they were perceived by participants in my qualitative interviews as being most effective at eliciting meaningful and thoughtful responses. Fourth, and most importantly, they relate to the theorized aspects underlying a TMS (specialization, credibility, and coordination [Liang et al., 1995]).

Specialization is critical to the theoretical underpinnings of TMS (Liang et al., 1995). Built on the expectation that individuals share diverse knowledge to collectively know more as a team than they would as a group of individuals, diversification of knowledge is critically important to the formation of a well-developed TMS (Huber & Lewis, 2010).

Coordination is similarly important to previously proposed definitions of TMS (e.g. Liang et al). Related to both the structure and process of organizing teamwork, a team that is able to coordinate their tasks is better-prepared to make use of diverse knowledge and reach an understanding of who within the team knows what.

Trust, as defined in this analysis, is closely related to the aspect of credibility that has been used in other TMS studies (e.g. Lewis, 2003). More specifically, the quality of being “absolutely confident” in the information provided by a teammate is very closely tied to credibility, or trust in the knowledge of other team members (Lewis). Previous literature on the topic of trust in TMS indicates that this aspect may serve as an antecedent of TMS, a unique dimension of TMS, or a moderator of TMS and performance (Ren & Argote, 2010); however, as defined in this evaluation, it is very closely related to the credibility aspect of TMS as previously defined.

Finally, the enjoyment network data that are gathered in this analysis provide information about a team’s TMS above and beyond what has been collected by past researchers (e.g. Lewis, 2003; Liang et al., 1995). Indicating that one “thoroughly enjoys” working with a teammate is likely related to teammates’ abilities to coordinate tasks and perceive one another as credible; therefore, there is a loose theoretical tie between this item and previous TMS scholarship. I believe that enjoyment may serve as an antecedent to TMS or that it may capture information that is closely related to both trust and coordination items. Not explicitly evaluated by past analyses of TMS, plans for evaluating the relation between enjoyment network attributes, performance, and other collected variables are described in detail later in this chapter.

The use of single questions to obtain information about the networks analyzed in this study may cause some concern regarding item reliability; however, previous studies have found that these single-item network questions perform reasonably well (Marsden, 1990). Additionally, test-retest reliability of these measures was preliminarily examined in this study. Chapters 5 and 6 provide further information on this analysis.

An assumption of using questions 2, 3 and 4 to collect network data is that individuals on a team are aware of how their teammates perceive one another in terms of trust, coordination, and enjoyment. This assumption was verified during qualitative interviews in the pilot study. This assumption is also logically reasonable considering the small number of individuals on a team.

The guidelines provided to participants prior to taking the survey (included in Figure 9 above) were written to discourage overly generous ratings. From my pilot study (described in Chapter 3), and from recommendations from Dillman, Smyth, and Christian (2014), when participants perceive a certain response as more favorable (e.g. indicating that they work well with their teammates), they are more likely to provide this information even if it is not accurate. This tendency to provide generous ratings also influenced the wording of survey questions. By responding to items with language such as “solely responsible”, “absolutely confident”, and “total ease”, participants were less likely to provide negatively skewed responses.

Strong-tie density is a potential measure of TMS explored in this study. As previously described, strong-tie density is the sum of all strong ties (scores of 4 and 5 or 5) in a dataset of network ties divided by the possible number of strong ties in that dataset

($\Delta = \frac{\sum_{i=1}^{g}\sum_{j=1}^{g} x_{ij}}{g(g-1)}$, where *i* is an individual in a network indicating ties to others, j is an individual in a network receiving ties from others, x_{ij}represents the presence or absence of a tie from individual *i* to individual j, and g is the number of individuals in the network of interest). Strong-tie density was selected for use above other network measures (e.g. strong-tie reciprocity) based on the prominence of density as an outcome of interest in social network literature as well as the results of the pilot study described in Chapter 3.

Strong-tie transitivity was also selected for use as a measure of TMS based on the results of the pilot study. Transitivity is a relation that is represented by ties between three individuals. When one individual is tied to a second individual and that second individual is tied to a third individual, the first individual is also tied to the third individual ("the friend of a friend is also a friend"). Transitivity is a network measure that is indicated by a transitivity index, $\frac{t}{T}$, where *t* represents the number of transitive triads present in a network and *T* represents the number of potentially transitive triads in a network (Wasserman & Faust, 1994).

Although density and transitivity are related, it is not expected that strong-tie transitivity will perform as well as strong-tie density. This is based on the nature of the small teams involved in this study who regularly interact with all team members and do not require a formal coordination of efforts.

Lewis inventory. In order to evaluate the incremental validity of the proposed network measures, Lewis's (2003) inventory was also administered to all participants. This instrument, provided in Appendix A, was modified from its original form so that no questions were reverse formatted. As described in Chapter 3 of this document, this action was recommended by Lewis. Otherwise, the inventory was exactly the same as originally used by Lewis and asked participants to provide information about their team experiences at the team-level. While this information was important to obtain, it should be noted that the network data previously described were of a higher priority in this study. The TMS network inventory was provided first and the Lewis inventory was provided second so as to encourage participants to provide accurate TMS network data even if they became fatigued from the length of the survey.

Team composition. Information about team composition was also collected as part of this study. These data, including race, gender, grade point average (GPA), and team member familiarity, were used as control variables. For example, by obtaining data about GPA, the influence of teams with high-performing individual students versus the impact of highly effective teams (as measured by TMS scores) was distinguished. In order to encourage participation in the TMS network inventory and assuage participant fear of a lack of anonymity, this information was collected after participants provided responses to the TMS inventory.

An index of racial heterogeneity, h, was calculated as follows: $h = 1 - \sum_{i=1}^{R} p_i^2$, where R is the number of racial backgrounds present on a team and p_i is the proportion of team members coming from a given racial background, *i* (Gibbs & Martin, 1962). This provided a heterogeneity index for each team with a minimum of 0 (a racially homogenous team) and a maximum of $1 - \frac{1}{n}$ (a team composed of individuals all coming from different racial backgrounds).

For example, consider three teams, each with five members. On one of these teams, all members self-identify as Asian. h is then equal to 0 ($1 - \sum_{i=1}^{1} p_i^2 = 1 - 1^2 = 0$). The second team has two Asian, two white, and one black member. In this case, h is equal to 0.64 ($1 - \sum_{i=1}^{3} p_i^2 = 1 - (0.4^2 + 0.4^2 + 0.2^2) = 0.64$). The third team has one team member who self-identifies as Asian, one who self-identifies as white, one as black, one as Hispanic, and one as coming from two or more racial backgrounds. Here, h is equal to 0.80 ($1 - \sum_{i=1}^{5} p_i^2 = 1 - (0.2^2 + 0.2^2 + 0.2^2 + 0.2^2 + 0.2^2) = 0.80$).

Team-level information about gender was captured by calculating the mean gender of each team, with males coded with values of 0 and females coded with values of 1. Therefore, the gender variable used in this analysis provides information about the proportion of females on a team.

Team member familiarity was accounted for by asking participants to consider their level of familiarity with each of their team members prior to the start of the project ("Prior to beginning this project, how well did you know this person?"). The response scale included a range from 1 ("I did not know this person at all") to 5 ("I interacted regularly with and was extremely close to this person"). An index of familiarity was calculated for each team by taking the mean of all individual responses; therefore, this index had a minimum value of 1 and a maximum value of 5.

Performance measure. The performance measure used in this study mirrors that described in Chapter 3's pilot study (from "1 - Unacceptable; The team does not accomplish satisfactory work at any level" to "5 - Outstanding; The team is polished, thorough, and on par with what could be expected from a team of professionals"). The performance measure was written with purposefully strong language to solicit unbiased responses. Similar to the pilot study described in Chapter 3, and as described in the "Procedures" section of this chapter, ratings were provided by course instructors and teaching assistants. Although it was my intention to obtain multiple ratings for each team, this was not accomplished as in some courses only one rater had sufficient knowledge to provide an informed rating.

Beyond including strongly worded items, other recommendations adhered to in creating this instrument include using simple and familiar words, using specific and concrete language, asking one question at a time, using complete sentences, and limiting scales to 5 categories (Dillman et al. 2014; Tourangeau et al., 2000). Additionally, I followed recommendations of obtaining feedback on draft questions through cognitive interviews to confirm the instrument would function as intended (Dillman et al.). This process is described in Chapter 3.

External aspects of validity. In support of Research Question 2, measures of external aspects of validity were evaluated by calculating the bivariate correlations between strong-tie network densities and transitivities and measures of teamwork (to demonstrate evidence of convergent validity) and professionalism (to demonstrate evidence of discriminant validity). Ratings of these measures were provided by course instructors and teaching assistants. The surveys that were used to collect information related to external aspects of validity are provided in Appendix H. Although these measures of teamwork and professionalism have not been validated in previous studies, they have been used to measure teamwork in an undergraduate program for the past five years and have been described by King and Herrmann (2015).

Procedures

Data were collected toward the end of a fifteen-week semester (during weeks 13 and 14) in Fall 2016. Participants were recruited based on their enrollment in one of four courses. The TMS network inventory, Lewis inventory, and request for demographic information were delivered to participants through an online survey created using the Qualtrics 360 platform (Qualtrics, 2016). Before administering the survey, participants were provided with more information about the research project and informed consent was obtained via an online signature.

Administration of the survey took place during class time. This method was selected to facilitate high levels of participation. Although administering a survey in class may influence the accuracy of ratings received, students' use of laptop computers provided privacy for participants while responding. Furthermore, participant names and survey items requiring ratings did not appear on the same computer screen. This encouraged more accurate ratings.

By collecting data from students during class (data for the pilot study described in Chapter 3 were collected outside of class) and having a well-developed follow-up procedure (phone calls, emails, and in-person meetings, as needed) I was able to achieve a mean participation rate of 90.6% for the proposed network inventory. There is, however, a risk that these data collection methods introduced bias into the responses obtained. I believe that the benefit of obtaining high response rates greatly outweighs this risk. While I also considered asking participants to respond to additional questions about how they believed they were perceived by their teammates and using this to protect against the risk of missing data, I believe that creating a longer survey instrument would have done more to hinder participation than enhance the likelihood of obtaining complete data.

The TMS network inventory was re-administered through email to all participants approximately one month following the end of the course projects in order to allow for the evaluation of test-retest reliability. Lower levels of participation were expected as this re-administration was not carried out during class time and the project was of less importance to participants at this time. Unfortunately, a low response rate of 8% was achieved due to this less than ideal setting. Strategies for improving this poor response rate are discussed in Chapter 6.

Performance data were collected from “experts” including course instructors and teaching assistants. All raters provided evaluations for multiple teams (course instructors and teaching assistants evaluated from six to nine teams each). The impact of involving different raters is discussed in Chapter 6 of this document. These expert raters were recruited through in-person meetings and email requests. Performance data were collected from expert raters at approximately the same time as TMS data were collected from participants (the end of a fifteen-week semester) via an online survey.

One or two ratings were obtained for each team. The presence of multiple ratings from different raters of a single team was managed by simply selecting the first provided rating for use in the analysis. Justification for this decision is discussed in Chapter 5. If ratings were not provided after the first request, they were obtained through follow-up emails. While there is a slight risk of this direct involvement introducing bias to the information that is obtained from raters, the reward of greater participation far outweighs the risk presented by this approach.

Evaluations of teamwork and professionalism used in considerations of external aspects of validity were obtained from course instructors and teaching assistants as described above. Similar to performance ratings, one or two evaluations of teamwork and one or two evaluations of professionalism were obtained for each team. Also similar to performance ratings, the first rating provided was used in all analyses and in any instances where there were two ratings per team the second rating was not used.

The impact of using only one rating per team was evaluated by collecting two evaluations for teams in one course. An analysis of these two sets of ratings indicated that the two raters provided the same performance rating in 77.8% of all cases. Considering the range of possible ratings (that is, the maximum possible difference of 4 scale points), a 5.5% difference between performance ratings was observed. Ratings of professionalism and teamwork were similarly consistent between raters. Raters provided the same score 62.5% of the time and there was a 14.4% difference in ratings considering the maximum difference of 4 scale points. It can, therefore, be assumed that obtaining one rating per course provides sufficiently accurate information for this analysis.

Analysis

First, Pearson and Spearman correlations between variables were calculated, as dictated by variable type. Next, a series of ANOVA, multinomial logistic regression, and ordered probit regression models were run. All analyses were conducted to evaluate the relation of trust, enjoyment, and coordination network attributes (e.g. strong-tie density) with team performance, controlling for team composition variables including gender, race, GPA, and familiarity. The distributional properties of all variables were evaluated. Results of these analyses are provided in Chapter 5 of this document.

The outcome metrics of interest in this study are the beta coefficients for the strong-tie densities and transitivities of enjoyment, trust, and coordination networks as predictors of team performance in both multinomial logistic and ordered probit regression models. Strong-tie density and transitivity of each network were first considered individually by evaluating the bivariate correlation between each network's strong tie density and strong tie transitivity and team performance. Figure 10 below illustrates these analyses, providing conceptual models of the relation between the proposed network measures and team performance.

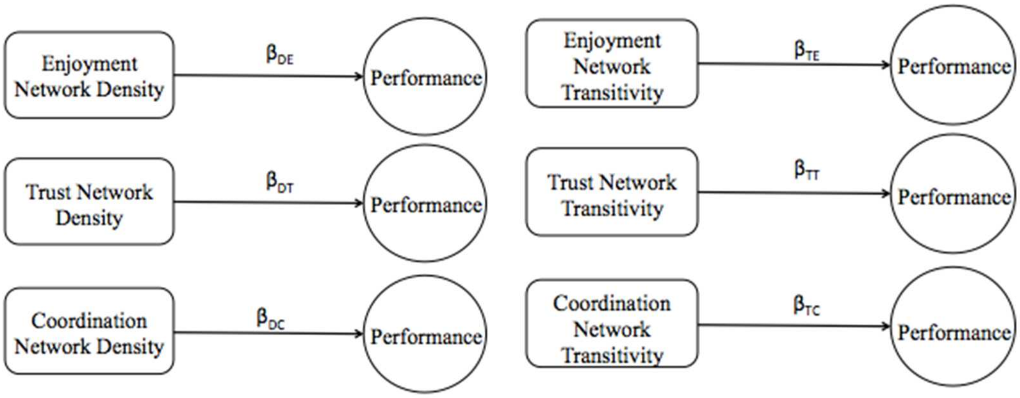


Figure 10. Conceptual models of individual networks and performance

In addition to evaluating the relation between individual networks and performance, the relations between a simple sum aggregation of these measures and performance were also evaluated. Using a composite of the density measures and a composite of the transitivity measures for each network (trust, enjoyment, and coordination) provides an overall measure of TMS. This aggregation of density and transitivity networks is illustrated below in Figure 11.

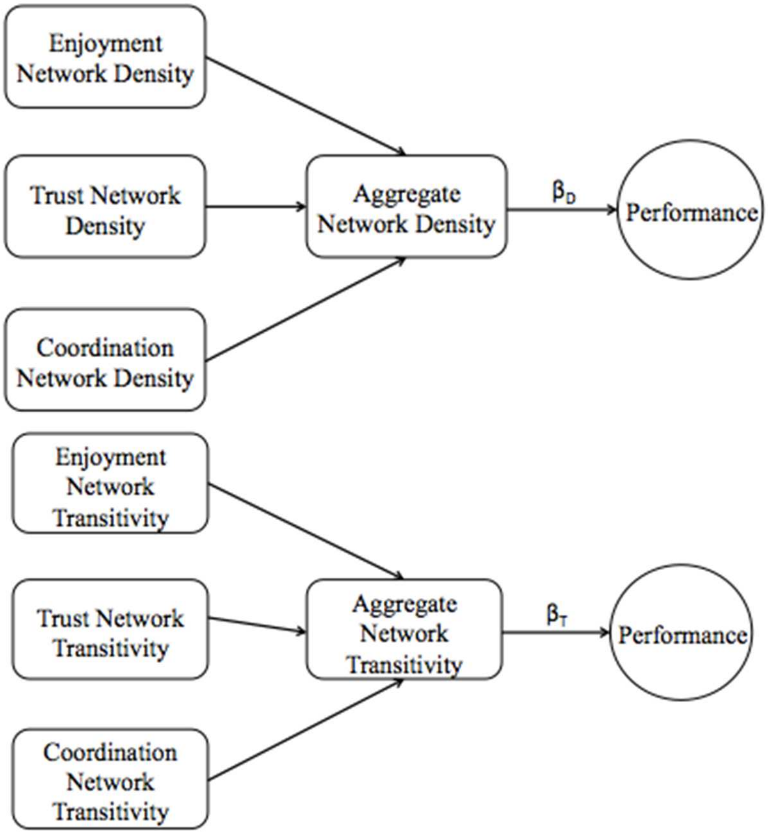


Figure 11. Conceptual models of aggregate density and transitivity networks and performance

Assessing Validity

In evaluating whether a new measure of any performance or behavioral construct is fit for use, it is important to confirm the measure's construct validity. According to Messick's (1995) theory of construct validity as unifying a number of sub-elements, validity should be evaluated from multiple perspectives to ensure that the results of an instrument are both theoretically meaningful and practically useful. In support of

Research Questions 2 and 3, below I provide a discussion of how a number of aspects of construct validity were evaluated statistically (external aspects of validity, generalizability, incremental validity) or considered theoretically (content, substantive, and consequential aspects of validity) in this study.

External aspects of validity. To investigate Research Question 2, whether the proposed network measures demonstrated external aspects of validity, I evaluated the convergent validity of my instrument by correlating the strong-tie densities and transitivities of enjoyment, coordination, and trust networks to expert ratings of (1) team performance and (2) a measure of multidisciplinary teamwork (survey provided in Appendix H). Scores for each element were aggregated to obtain an overall measure of teamwork. Discriminant validity was evaluated by correlating the proposed TMS network measures to a team-level measure of business ethics and professionalism, also provided in Appendix H. Again, scores for each element were aggregated to obtain an overall measure of professionalism.

Although there are no explicit rules regarding cutoff values that demonstrate convergent and discriminant validity (Crocker & Algina, 1986), at the very least, measures of convergent validity should be greater than measures of discriminant validity. That is to say, I believe that the correlation between both the proposed TMS network items and team performance and the proposed TMS network items and multidisciplinary teamwork should be greater than the correlation between the proposed TMS network items and the measure of professionalism previously described.

More specifically, I believe that correlations at or above 0.60 provide sufficient evidence of convergent validity. This correlation is also known as a validity coefficient and would provide a coefficient of determination of 0.36, meaning that 36 percent of the variance in performance or multidisciplinary teamwork would be attributed to TMS (Crocker & Algina, 1986). Similarly, I believe that correlations at or below 0.10 provide evidence of discriminant validity, meaning that 1 percent of the variance in professionalism would be associated with TMS.

Generalizability. To evaluate Research Question 3, I considered the generalizability of my instrument across teams of multidisciplinary students and teams of students coming from a single discipline. Although it is important to evaluate the relations between network densities and performance in non-academic settings, this was not carried out in the present analysis. It is my hope that future studies will be able to evaluate the correlations between network densities and performance under different contexts to provide further evidence of generalizability; however, I would expect some of the results from student teams to translate to teams of professionals due to similarities between teams working in these settings.

For example, I believe that the applied and experiential nature of the student projects evaluated in this analysis hold many consistencies with professional projects that would lead to similar patterns of relations between network measures and performance. These student projects took place over many weeks, were generally unstructured, required team members to take on different roles, and were influenced by participants' abilities to manage complex tasks. All of these features are also present in the workplace. Aspects of the student projects that may lead to different results include motivation (e.g.

grades versus salary or bonuses), commonality of goals (e.g. individual versus company success), and familiarity of team members (for example, professional workers may have more of a history of working together on projects than students). In order to evaluate some of these potential discrepancies, the motivation and commonality of student goals was evaluated in the pilot study described in Chapter 3 and team familiarity was used as a control variable.

In further considerations of the generalizability of this study, a reduced sample size was present as I calculated a correlation coefficient on two subsamples. With a lower limit of 15 subsamples (e.g. the number of mechanical engineering teams available for evaluation), the 95% confidence interval for the population correlation coefficient was expected to fall between 0.15 and 0.86. Again, this interval assumes an effect size of 0.60 and an achieved power of 0.80. It should be noted that this is quite a broad interval and, therefore, for this analysis the results are unlikely to demonstrate adequate power.

Test-retest reliability. To evaluate Research Question 4, whether the proposed network measures demonstrated sufficient test-retest reliability, I re-administered the network survey instrument approximately 2 months after the original inventory was completed. Although response rates were disappointingly low (8%), responses from individuals participating in both of these survey administrations were used to calculate Cohen's Kappa (Cohen, 1960). This evaluation is described in Chapter 5.

Incremental validity. Incremental validity, used to demonstrate whether a new measure increases predictive ability above existing measures, was calculated by comparing the variance explained by Lewis's (2003) model alone to the variance explained by Lewis's model with the addition of the proposed network measures. This analysis was conducted in support of Research Question 5. If models with the addition of the network measures explain more variance than the model with Lewis's model only, this would demonstrate the incremental validity of the proposed network measures.

Performance of Lewis's Inventory

Finally, in support of Research Question 6, responses to Lewis's (2003) 15-item inventory were fit to Lewis's higher-order factor model. A number of fit indices were explored including indices of relative, parsimonious, and absolute fit. This evaluation provides information about the performance of a previous inventory that may influence the future direction of work in the area of TMS measurement.

Summary of Methods and Potential Implications of Findings

This study involves the development, use, and analysis of a new TMS measurement instrument and a validity study of the newly proposed instrument. More specifically, this study evaluates the ability of strong-tie density and strong-tie transitivity of enjoyment, trust, and coordination networks to serve as valid measures of TMS (evaluating Research Question 1).

A preliminary consideration of the implications of potential findings provides prospective uses for the results of this study. As previously mentioned, it is my hope that these results will motivate the use and subsequent evaluation of the proposed instrument in different settings (e.g. including corporate offices, hospitals, sports teams). To the best of my knowledge, no previous studies have used social network measures alone to formally evaluate TMS in small teams. I believe that this provides a promising avenue for future research through the method’s focus on the ties between individuals.

Chapter 5: Results

This chapter describes the results of the analyses that were undertaken to evaluate the validity of a newly proposed instrument to measure TMS. More specifically, the validity of a four-item network inventory is reviewed, including a description of all dependent and independent variables employed in this analysis, a discussion of the network attributes that were used (strong-tie density and strong-tie transitivity), and an investigation of whether the proposed network measures relate to team performance, individually or in the aggregate (Research Question 1). External aspects of validity and generalizability are statistically evaluated (Research Questions 2 and 3), other aspects of validity are discussed, and the test-retest reliability of the proposed network measures is examined (Research Question 4). This analysis also examines the benefit of using results from the newly proposed measurement instrument over those from Lewis's (2003) previously proposed and widely used inventory (Research Question 5) and includes an independent evaluation of the performance of this previously proposed measure of TMS (Research Question 6).

Treatment of Missing Data

The overall participation rate for network inventory items was 90.6%. Two of the six participating courses, however, had markedly lower participation rates. After evaluating the impact of the reduced power that would result from removing these two courses from the study (58 individuals on 11 teams), it was determined that this was not overly detrimental to the study's achieved power and teams from these courses were removed from the analysis. The anticipated confidence interval for an effect size of 0.6 was widened by only 0.04 units with the removal of teams from these two courses.

Additionally, based on recommendations from Huisman (2009), I removed all teams from the analysis where participation was less than 40%. The total number of teams surveyed was 54. Upon removing teams from the two courses with low participation rates, my sample size was reduced to 43 teams. Removing other teams with lower than 40% participation rates resulted in 39 teams, or 72.2% of the total number of teams surveyed. These decisions, to remove teams from courses with low participation and where total participation was less than 40%, will be discussed in Chapter 6.

Of the teams that were included in the analysis, 28 had complete information provided by all individuals on the team. Two teams had missing information for two of the individuals on the team and nine had missing information for one of the individuals on the team. There were no instances of item non-response, where an individual neglected to respond to some but not all of the survey items. All missing data were due to unit non-response, where an individual did not provide responses to any of the inventory items. As recommended by Huisman (2009) and discussed in Chapter 6, these individuals were removed from the analysis.

In considering the potential impact of removing unit non-response data as described above, the teams with unit non-response that were included in the analysis were evaluated to determine how excluding unit-non response data impacted team-level averages for the network attributes and for percent female. Other variables included in the analysis (GPA, race) could not be analyzed, as this information was not available. Results of this analysis, provided in Appendix I, illustrate that team averages for specialization, enjoyment, trust, and coordination were generally slightly higher where unit non-response data were included. This means that teams generally had higher average scores for specialization, credibility, and coordination before scores provided about the individuals who did not respond to the survey were removed from the analysis. Considering each attribute individually, differences between including and excluding unit non-response were not statistically significant at the level of $\alpha = 0.05$. Still, the presence and treatment of missing data is a limitation that should be considered in future analyses.

Determination of Strong-Tie Density

As discussed in Chapter 4, the relation between team performance and strong-tie density was calculated by first considering ratings of 4 (Agree) and 5 (Strongly Agree) as being indicative of strong ties and then considering ratings of 5 alone as being indicative of strong ties. Table 3 provides Spearman correlations between performance (with levels of 2/3, 4, and 5) and strong-tie density enjoyment, trust, and coordination networks coded using both of these schemes. Although confidence intervals for both coding schemes include zero, it appears as though strong ties coded by including ties of 4 and 5 are likely better measures of TMS than strong ties coded by including ties of 5 alone. This is based on the positive point estimates of correlations calculated using ties of 4 and 5 versus the negative point estimates of correlations calculated using ties of 5 alone. Therefore, strong ties will be considered to be those indicated by scores of 4 and 5 for the remainder of this analysis.

Table 3

95% Confidence intervals and point estimates of bivariate Spearman correlation between strong-tie density and performance based on different coding schemes (N=39)

	Strong Ties 4 and 5 Point Estimate	Strong Ties 5 Confidence Interval	Strong Ties 5 Point Estimate	Strong Ties 5 Confidence Interval
Enjoyment	0.09	[-0.38, 0.24]	-0.08	[-0.40, 0.23]
Trust	0.17	[-0.46, 0.15]	-0.17	[-0.43, 0.19]
Coordination	0.21	[-0.44, 0.18]	-0.14	[-0.47, 0.14]

A visualization of the performance of two teams, one high performing and the other low performing, is provided below in Figure 12. The low performing team (performance rating of 2: "The team falls short of expectations for a team of honors students but still provides satisfactory work in some areas"), pictured on the left, has a strong-tie coordination density of 0.42. Of 12 possible ties, 5 are present. The high performing team (performance rating of 5: "The team's work is polished, thorough, and on par with what could be expected from a team of professionals"), pictured on the right, has a strong-tie coordination density of 0.83. Of 12 possible ties, 10 are present.

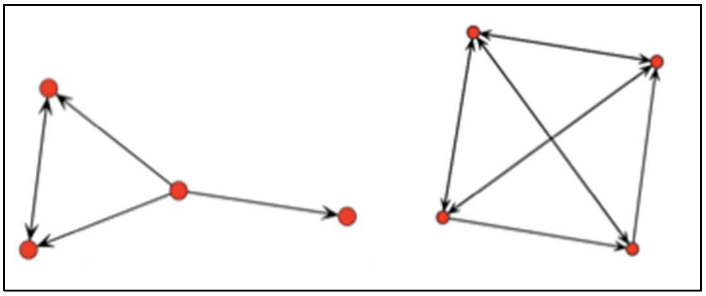


Figure 12. *Coordination strong-tie density of low and high performing teams*

Description of Variables

The independent variables used in this analysis include a number of control variables: percent female on each team, a heterogeneity index of race for each team (as defined in Chapter 4), mean team grade point average (GPA), the range of GPAs within a tea,, and a measure of team familiarity (as defined in Chapter 4). Mean team-level specialization and measures of strong-tie density and strong-tie transitivity of enjoyment, trust, and coordination networks are included as independent variables of interest. The dependent variable in all analyses is team-level performance. A description of these variables is provided in Table 4 and density plots and histograms of these variables are provided in Figure 13. Also evaluated in this analysis and described below are measures of convergent and discriminant validity, team-level evaluations of teamwork and performance.

Table 4

Descriptive statistics across 39 teams

	Mean	Median	Standard Deviation	Minimum	Maximum
Percent Female	0.29	0	0.23	0	1
Race Heterogeneity	0.43	0.45	0.18	0	0.72
Mean GPA	3.51	3.54	0.21	3	3.85
Range GPA	0.64	0.66	0.27	0.15	1.25
Familiarity	3.36	3.33	0.99	1	4.9
Enjoyment Density	0.77	0.77	0.15	0.43	1
Trust Density	0.76	0.81	0.14	0.43	1
Coordination Density	0.71	0.73	0.14	0.42	1
Enjoyment Transitivity	0.93	1	0.13	0.46	1
Trust Transitivity	0.95	1	0.08	0.71	1
Coordination Transitivity	0.90	0.92	0.12	0.60	1
Specialization	3.83	3.88	0.42	2.17	4.5
Performance	3.92	4	0.81	2	5
Teamwork	13.92	14	1.98	10	16
Professionalism	14.49	15	1.34	11	16
Lewis	55.4	56.4	5.5	42.5	66.4

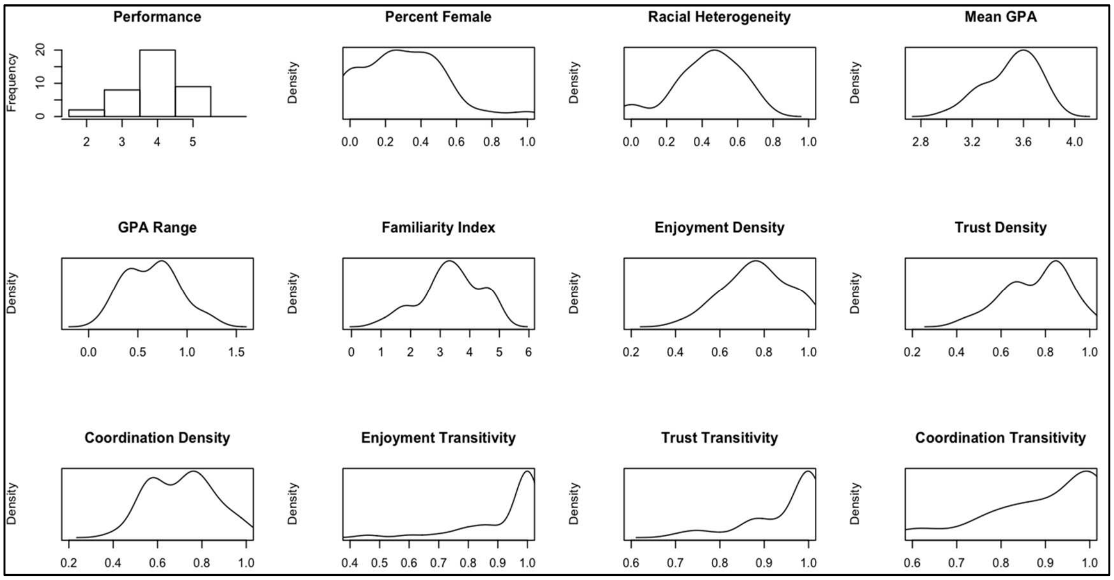


Figure 13. *Density plots and histograms of team-level variables and network measures (N=39)*

Correlations illustrating the bivariate relations between all independent and dependent variables used in this analysis are provided below in Table 5. Spearman correlations between performance and all team-level and network measures are provided as performance is an ordered categorical variable measured with levels of 2/3, 4, and 5. As is evidenced by this correlation matrix, there is a significant correlation between Lewis's measure of TMS and all of the proposed network measures, with the exception of trust transitivity, at the level of $\alpha = 0.10$. This significant relation between predictors is discussed later in this chapter. It should also be noted that none of the control variables had a significant relation with performance. Therefore, these control variables were not included in subsequent regression models. The significant relation between coordination transitivity and performance and the Lewis aggregate score and performance should also be noted as the relations between these variables are also discussed in detail later in this chapter.

Table 5

Bivariate correlations of team-level variables and network measures (N=39)

	Performance (Spearman)	Enjoyment Density	Trust Density	Coordination Density	Enjoyment Transitivity	Trust Transitivity	Coordination Transitivity	Aggregate Density	Aggregate Transitivity	Lewis Score	Percent Female	Racial Heterogeneity	Mean GPA	GPA Range	Familiarity
Performance (Spearman)	-	0.09	0.17	0.21	-0.05	0.07	0.22*	0.14	0.09	0.38*	0.23	0.21	0.17	0.08	0.06
Enjoyment Density		-	0.63*	0.71*	0.61*	0.36*	0.58*	0.88*	0.73*	0.64*	-0.08	0.26	0.09	0.20	-0.30*
Trust Density			-	0.74*	0.15*	0.63*	0.51*	0.88*	0.54*	0.52*	0.01	0.35	0.00	0.08	-0.17
Coordination Density				-	0.26*	0.33*	0.72*	0.92*	0.60*	0.39*	-0.06	0.23	-0.08	-0.01	-0.31*
Enjoyment Transitivity					-	0.27*	0.36*	0.38*	0.79*	0.33*	-0.19	0.10	0.32	-0.07	-0.11
Trust Transitivity						-	0.25	0.49*	0.60*	0.20	-0.01	0.28	0.03	-0.01	0.04
Coordination Transitivity							-	0.67*	0.77*	0.43*	-0.09	0.22	-0.05	0.03	-0.17
Aggregate Density								-	0.70*	0.58*	-0.05	0.31	0.01	0.10	-0.29*
Aggregate Transitivity									-	0.45*	-0.15	0.26	0.16	-0.03	-0.13
Lewis Score										-	0.03	0.14	0.18	0.20	-0.02
Percent Female											-	0.27*	0.50*	-0.24	0.14
Racial Heterogeneity												-	0.15	0.28*	0.00
Mean GPA													-	-0.32*	0.11
GPA Range														-	-0.04
Familiarity															-

* Significant at $\alpha = 0.10$

The information provided above in Tables 4 and 5 and Figure 13 enables for the evaluation of assumptions some assumptions of ANOVA and regression models. Univariate normality is generally satisfied in the case of all control variables. Proposed measures of density demonstrate slight negative skew and proposed measures of transitivity are highly negatively skewed. These deviations from normality should be noted and future analyses should consider transforming measures of transitivity to better approximate normality. Multicollinearity, which must be considered in regression models, may problematic as nearly all proposed network measures are correlated with the Lewis aggregate TMS measure. In order to test whether or not multicollinearity is present, evaluations of variance inflation factors (VIFs) should be conducted in order to evaluate the relation between variables included in the best peforming model(s).

In addition to reviewing descriptive statistics and bivariate correlations between independent variables, other diagnostic evaluations were carried out to ensure that assumptions of these models were met. These assumptions include an absence of outliers, an evaluation of homogeneity of variance (satisfied at *p* =0.10),

Evaluation of Network Measures of TMS

Because team measures of performance were rated on an ordinal scale and only one rating was used, prior to running regression analyses to evaluate the performance of the proposed network measures, omnibus ANOVA analyses were carried out to determine whether there were differences in each of the proposed measures at different levels of team performance. For this analysis, the team that was scored a “2” on performance was grouped with the teams that were scored a “3.” Results of omnibus

ANOVA tests include that, at the level of $\alpha = 0.10$, only one of the proposed network measures, coordination transitivity, had significant differences at different levels of performance (F [2, 37] = 3.6, $p = 0.07$). Subsequent analyses demonstrate that a linear trend was present between coordination transitivity at scores of 2/3, 4, and 5 with mean coordination transitivities of 0.68, 0.71, and 0.76 at each level, respectively. Post-hoc pairwise analyses found significant differences between coordination transitivities of scores of 2/3 versus 5 at the level of $\alpha = 0.10$. Significant differences were not found between scores of 5 versus 4 or 4 versus 2/3.

Figure 14 below illustrates the relation between strong-tie coordination networks and performance for a sample of 10 teams selected to illustrate a range of team performance scores. A sample of teams was used to allow greater visibility of the included network graphs[2]. As can be seen in Figure 14, teams with greater levels of strong-tie transitivity were generally evaluated to be better performing than teams with lower levels of strong-tie transitivity. For example, the team with the lowest performance, a score of 2, had coordination network transitivity of 0.60 while the teams with the highest performance, a score of 5, had coordination network transitivity measures of 0.88, 0.92, and 1. It should be noted that this trend is also true for strong-tie coordination density, where higher performing teams also had higher density.

---

[2] Although a scatterplot illustrating the performance of all 39 teams was considered as an additional visual representation of performance scores, the prevalence of scores of 4 and 5 make this graphic less useful.

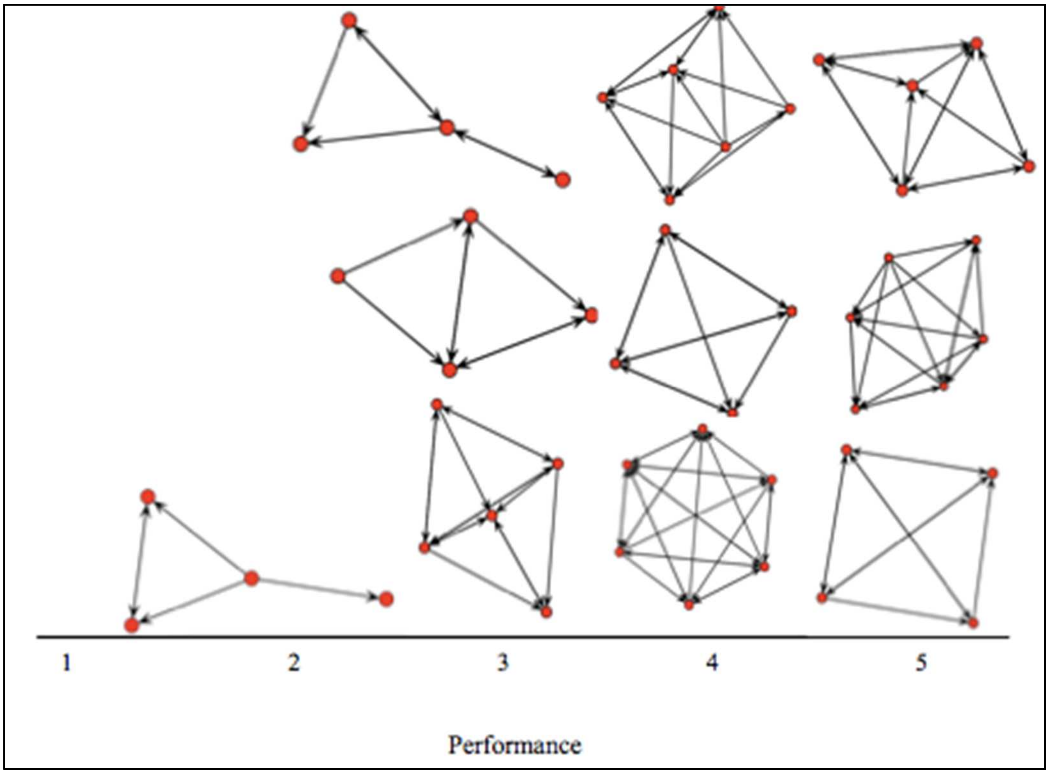


Figure 14. *Examples of strong-tie transitivity coordination network graphs and team performance ratings*

The Lewis aggregate TMS score was also significantly different at different levels of performance (*F* [2, 37] = 10.59, *p* = 0.02). Therefore, coordination transitivity and Lewis aggregate TMS score were the only covariates included in future regression analyses.

Based on these results and in evaluation of Research Question 1, a series multinomial logistic and ordered probit regression models were run. All models used team performance as the dependent variable. Model 1 used coordination transitivity as an independent variable, Model 2 used the Lewis aggregate TMS score as an independent variable, and Model 3 used both coordination transitivity and Lewis aggregate TMS score as independent variables. All independent variables used in these models were mean-centered.

Results of these analyses are provided below in Tables 6 and 7. Both multinomial logistic and ordered probit regression models were run. Multinomial logistic models assume that the outcome variable, performance, is not ordered. Ordered probit regression models consider the ordered nature of team performance. In considering this ordered nature of the dependent variable, ordered probit models provide an additional degree of freedom, providing more power to these models.

First, considering the multinomial logistic regression model, which uses teams with performance scores of 2 or 3 as the reference group, coordination transitivity does not provide a statistically significant distinction between different levels of performance. The Lewis aggregate TMS score, however, is significant at the level of $\alpha = 0.10$ in Model 2 and Model 3. Based on AIC values, Model 2, which includes only the Lewis TMS score, is best performing. Interpreting the results of this model, if a team were to increase their Lewis TMS score by one unit, the multinomial log-odds for receiving a performance score of 4 versus 3 would increase by 0.19 units and the multinomial log-odds for receiving a performance score of 5 versus 3 would increase by 0.23 units.

Table 6:

Beta coefficients and AIC values for multinomial logistic regression models (N=39)

	β_{TCoord_4}	β_{TCoord_5}	β_{Lewis_4}	β_{Lewis_5}	AIC
Model 1, Transitivity Coordination	3.66	7.30	-	-	85.18
Model 2, Lewis Score	-	-	0.19*	0.23*	80.79
Model 3, Transitivity Coordination and Lewis Score	0.91	4.52	0.19*	0.19*	83.90

* Significant at $\alpha = 0.10$

A consideration of the ordered probit logistic regression model results similarly finds Model 2 to be best performing, based on this model having the lowest AIC value. Also based on these AIC values, as expected, these models perform slightly better than the previously described multinomial logistic regression models based on their consideration of the ordered nature of the performance variable and additional degree of freedom. An interpretation of the results of this model is that if a team were to increase their Lewis TMS score by one unit, their ordered log-odds of receiving a performance score of 2/3 versus 4 or 5 would increase by 0.14 units.

Table 7:

Beta coefficients, cutpoints, and AIC values for ordered probit regression models using mean-centered variables

	β_{TCoord}	β_{Lewis}	AIC
Model 1, Transitivity Coordination	4.60*	-	83.20
Model 2, Lewis Score	-	0.14*	80.17
Model 3, Transitivity Coordination and Lewis Score	2.23	0.12*	81.58

* Significant at $\alpha = 0.10$

As is evidenced in the information provided above, although transitivity coordination was significant in Model 1, it had contributed only minimally to performance variance after also incorporating the Lewis measure.

Evaluation of Validity

External Aspects of Validity

Measuring the relation between the proposed TMS network measures and performance is one way to evaluate the external aspect of validity. Other evidence to support external aspects of validity may be provided by the correlation between the proposed TMS network measures and measures of teamwork and professionalism (to measure convergent and discriminant validity, respectively). In order to evaluate Research Question 2 and explore Hypothesis 2A, that the proposed measures will correlate more strongly with teamwork than professionalism, the bivariate correlations between each proposed TMS network measure and professionalism and teamwork were calculated.

As is evidenced by the results provided below in Table 8, some evidence of external aspects of validity is present as measures of convergent validity are demonstrated by positive correlations between TMS network measures and teamwork (with the exception of enjoyment transitivity and trust transitivity) while measures of discriminant validity are demonstrated by negative correlations between TMS network measures and professionalism. Although confidence intervals for the correlation between professionalism and the network measures extend above the criterion levels described in Chapter 4 (i.e., 0.10) in some cases; this provides weak evidence of discriminant validity. No correlations between the network measures and teamwork did reached the desired criterion level of 0.60 to demonstrate convergent validity. It should also be noted that none of the bivariate correlations are statistically different from zero.

Table 8

95% Confidence interval of bivariate correlation between TMS network measures and measures of external aspects of validity

	Teamwork	Professionalism
Enjoyment Density	[-0.17,0.45]	[-0.40, 0.21]
Trust Density	[-0.23,0.40]	[-0.42,0.20]
Coordination Density	[-0.23,0.40]	[-0.42,0.21]
Enjoyment Transitivity	[-0.37,0.26]	[-0.55,0.02]
Trust Transitivity	[-0.33,0.30]	[-0.51,0.08]
Coordination Transitivity	[-0.15,0.46]	[-0.42,0.20]
Density Aggregate	[-0.20,0.43]	[-0.42,0.20]
Transitivity Aggregate	[-0.27,0.36]	[-0.56,0.03]

Generalizability

In support of Research Question 3 and hypothesis 3A, that correlations between the proposed network measures and team performance will hold across different types of teams, generalizability of the best performing TMS network, coordination transitivity, was evaluated. This was accomplished by evaluating an ordered probit model with team performance as an outcome, coordination transitivity as a predictor, and incorporating team type (single discipline and multidisciplinary) as an interaction term in this model. Results of this analysis illustrates evidence of generalizability for this measure across the two types of teams, single-discipline teams (the reference group in these analyses) and multidisciplinary teams ($\beta_{interaction}$= 4.16, t =0.59, $p > 0.10$).

Test-Retest Reliability

In support of Research Question 4, test-retest reliability was evaluated by re-distributing the survey to participants approximately two months after it was initially completed. Participation in this redistributed survey was quite low (12 of 204 participants). While disappointing, this was expected based on the time that had passed since project completion and a lack of motivation for participation. These 12 participants, however, came from different courses and generally reflected the demographics of gender, race, and GPA that were present in the initial population of survey participants. For example, of all participants in the original survey, 58% identified as white, 25% identified as Asian, 7% identified as two or more races, 7% identified as black, and 1% identified as Hispanic. In the follow-up sample, 60% identified as white, 25% identified as Asian, and 17% identified as two or more races. Although those who responded to the second survey may have other differences from the original population, including perhaps greater levels of investment in the project, it is reasonable to assume that they are generally representative of the full survey population.

Cohen's Kappa, κ, (Cohen, 1960) was calculated based on the original and follow-up responses to the TMS inventory items. Scores for each participant from the original survey administration and the follow-up survey administration are included in Appendix J. As can be calculated from this information, $\kappa = 0.69$. Based on guidelines for this index, 0.69 indicates "substantial agreement" (Landis & Koch, 1977). This provides some preliminary evidence of test-retest reliability, however the very low response rate to the re-administration of the survey should be noted. In order to ameliorate this problem in future studies, it would be helpful to provide participants with incentives for participation if information is to be gathered after the completion of a project.

Incremental Validity

In evaluating Research Question 5, the ordered probit regression analyses described earlier in this chapter provide information about the incremental validity about the best performing proposed network measure. By comparing the AIC values for models with the Lewis measure of TMS to models with both Lewis's TMS measure and the coordination transitivity measure, it can be determined that this network measure of TMS does not demonstrate incremental validity (see Table 7). That is to say, no improvement is offered above and beyond previously proposed measures of TMS.

Data-Model Fit of Previously Proposed Measure

In support of Research Question 6, the fit of Lewis's (2003) TMS inventory was evaluated using the newly collected data. These data were collected from 204 student participants on 39 teams, as described in Chapter 4 of this document. Analysis was conducted using the lavaan package in the R platform (Rosseel, 2012). Similar to the suboptimal fit noticed in the pilot study described in Chapter 3, good fit was not found between the data and Lewis's higher-order factor model. While the SRMR index illustrated reasonable fit (SRMR = 0.08), absolute and parsimonious fit indices did not (CFI = 0.92; RMSEA = 0.07). Taken together, these results follow those discussed in Chapter 3; Lewis's approach to TMS measurement, while widely used, has room for improvement.

A review of modification indices suggest that data-model fit could be improved by adding additional factor loadings, more specifically, loadings between endogenous items (TMS, Specialization, Credibility) and the first coordination item ("Our team worked together in a well-coordinated fashion”). Although one approach to improving data-model fit could be using modification indices to examine other structural models based on the Lewis inventory items, that is beyond the scope of this work, which is to evaluate a new method for TMS measurement. Future work can and should evaluate the structure of Lewis’s inventory and consider updating this structure based on modification indices, if supported theoretically.

Chapter 6: Summary and Conclusions

Based on the results described in Chapter 5, the proposed network measures cannot be determined to be valid measures of TMS. Despite the disappointing nature of these results, it is nonetheless useful to analyze why they may have occurred and to discuss future directions for this work. Therefore, this chapter begins by discussing why the transitivity of coordination networks outperformed the other proposed network measures. This is followed by a discussion of a number of considerations related to construct validity. Next, limitations of the present study are discussed and reasons behind the superior performance of the Lewis (2003) inventory are considered. Finally, other applications of using networks attributes as measures of TMS are reviewed and directions for future research are recommended.

Performance of Proposed Network Measures

Of the proposed network measures, the strong-tie transitivity of coordination networks had a greater impact on team performance than the other proposed network measures. There are several explanations for why this coordination network may have outperformed trust and enjoyment networks. Most importantly, the coordination inventory item ("I was able to accomplish tasks with complete efficiency when working with this person") is very closely related to the underpinnings of TMS theory, namely the importance of well-established structures and processes that are used by teams as they go about their work (Hollingshead 1998a, 1998b, 2001). The ability for a team to coordinate tasks has a direct relation to how teams function in practice, in contrast to team trust or team member enjoyment. These results, therefore, provide some support that instrumental network ties may be better measures of TMS than expressive network ties. While it is important that individuals on a team have sufficient relationships to encourage KWKW, more important than these relationships is the team's ability to work together to share information and accomplish tasks.

It should also be noted that coordination network transitivity had a stronger relation with team performance than coordination network. As discussed in Chapter 4 and as outlined by Hypothesis 1B, I surmised that although the two measures were related (in networks with greater levels of transitivity, density is also higher, and vice-versa), density would be a better measure of TMS as the nature of the small teams involved in the study meant that all participants regularly interacted with each of their teammates and had minimal need for a formal coordination of efforts. The impact of transitivity is likely greater than density in larger networks where density is low and transitivity is more variable.

The general lack of significance of results could be due to a number of factors. Perhaps the number of teams included in the analysis did not enable sufficient power to detect correlations that were statistically different from zero. Or, perhaps there was a lack of reliability in the proposed network measures. Although an exploration of test-retest reliability demonstrated some evidence of the reliability of the proposed measures, it is not clear how much error is included in the proposed network measures. A potential lack of reliability of the proposed measures should be considered in conjunction with the lack of statistically significant results.

Considerations of Construct Validity

In addition to mathematically evaluating external aspects of validity, generalizability, and incremental validity of the proposed network measures, other aspects of construct validity (content aspects of validity, substantive aspects of validity, consequential aspects of validity) were considered throughout this analysis.

Content Aspects of Validity

Content aspects of validity were incorporated in a number of ways. First, inventory items were created such that they that tapped constructs previously posited and determined to be related to TMS. Additionally, content experts were involved in the process of constructing the TMS network instrument. These content experts included faculty members and graduate students from various departments at UMD who had familiarity with TMS and experience using social networks to analyze teams. These experts were consulted prior to collecting data and provided feedback about the proposed instrument's content relevance, representativeness, and technical quality (as recommended by Messick [1995]). Considerations of content aspects of validity centered around Lewis's (2003) definition of TMS as "the cooperative division of labor for learning, remembering, and communicating relevant team knowledge." This definition encompasses a number of elements including intra-team relations (trust, enjoyment) and team processes (coordination).

Substantive Aspects of Validity

Substantive aspects of validity were considered by gathering qualitative information from participants in my pilot study. As described in Chapter 3, information was obtained about how respondents perceived and reacted to items. This helped to ensure that the proposed network questions functioned as intended. For example, participants noted that the proposed TMS network items were “straightforward and easy to understand” (Participant A, personal communication, July 13, 2016). Furthermore, as participants talked through their reactions to the proposed items, it was clear that they understood what information was being asked and that they were able to draw on their team experiences to provide that information.

Consequential Aspects of Validity

Consequential aspects of validity were considered by evaluating intended and unintended positive and negative consequences of using strong-tie network densities and transitivities as measures of TMS in both the short- and long-term. To this end, brief post-hoc interviews were conducted with survey participants to evaluate whether there may have been any adverse impact of study participation, including increased strain between team members, which would detract from team performance. Approximately two months after the survey was distributed, participants were asked to provide feedback on their team experiences after completing the network inventory. More specifically, participants were asked if participation impacted their work with their team or their project results. A total of nine interviews were conducted with students in multiple courses.

The effect of participation was noted to have been moderately positive or negligible. One respondent noted that the instrument simply provided an opportunity to “take a step back and evaluate myself and each of my team members and their contributions" (Participant C, personal communication, February 6, 2017). Another stated that completing the inventory "was a nice reminder that made people aware of their level of contribution to the project" (Participant D, personal communication, February 6, 2017). A third found that "the survey had very little impact on my team work, overall impression of the course, or project results" (Participant E, personal communication, February 6, 2017). While the number of participants in these interviews was quite small relative to the number of total participants, there is no reason to believe that these results would not generalize across the larger body of survey participants. It can, therefore, be stated that survey participation had little to moderately positive effects on participants in the short- and medium-term. It was not possible to examine longer-term effects; however, there is no reason to believe that these effects would be any different from those described above.

Aspects of Validity Not Considered

Structural aspects of validity evaluate the relation between measurements obtained from a given instrument and the structural model of a construct of interest. As discussed in Chapter 5, the proposed study does not evaluate structural aspects of validity as the structural modeling of TMS is beyond the scope of the proposed work. Evaluations of structural validity should be carried out in the future, especially considering the lack of data-model fit with Lewis's higher-order factor model (2003).

Limitations

One limitation of this study is that it uses student data and does not evaluate teams of working professionals. As mentioned in Chapter 5, it should be noted that the results obtained in this analysis may not transfer to other settings. For example, in the current study, it is likely that individual participants are highly motivated by grades. This individual motivation will not translate to a professional setting. This limitation should be considered before applying findings to or using the proposed instrument in different settings than those explicitly analyzed in this project.

A second limitation of this study is that several assumptions are made regarding the ways individuals behave within teams and the way that information is shared within a network (e.g. that individuals have knowledge about all relationships present within a team). These assumptions were based off of recommendations from previous studies and data collected through qualitative interviews. It is necessary to make assumptions in any study and the multiple sources that influenced these assumptions should render them relatively accurate. Still, future research could explore the influence of other assumptions on TMS tie formation.

Another limitation is that data are collected at only one point in time. As previously mentioned, ties related to trust, enjoyment, and coordination emerge dynamically. The end state of TMS ties, measured in this study, may be different than the ties that exist at other points in the project. While I believe that these end-of-project ties are likely to be similar to ties present approximately halfway through the project (as noted by qualitative study participants, behaviors tend to reach a steady state after a few iterations of working together), future research should evaluate strong-tie density and transitivity of enjoyment, trust, and coordination networks at multiple points in time to obtain a better sense of how ties may change over time.

The fact that different raters evaluated different teams and that teams performance scores were nested within raters is another limitation of this work. This may have resulted in some raters providing overly generous ratings and other raters providing overly critical ratings. Despite efforts to explain the rating scale to all individuals providing performance evaluations, differences between raters will certainly emerge, even in cases where multiple ratings are obtained. This will adversely impact results by attenuating the calculated relation between TMS network measures and team performance. This problem would be ameliorated if future analyses were able to involve the same raters in the performance evaluation of each team or if future studies had sufficient number of teams to consider the nesting of performance ratings through a multilevel analysis.

Another limitation of this work is that teams were formed in different ways in each of the courses. In some courses, students selected their own team. In others, teams were assigned to maximize diversity of team members. Although a control for team-level familiarity was used, it is likely that the various team selection processes influenced the ways that teammates were scored. Future studies should aim to include teams that were formed using similar processes, if possible.

The collection of performance data is another limitation of this study. The focus of this work, the validation of proposed network measures of TMS, led to a thoughtful consideration of the network items that were used to gather data. The performance data that were collected, unfortunately, did not undergo the same process of thorough vetting and selection. It could be that the proposed network measures have a very high correlation with performance, but that the measure that was used does not sufficiently represent actual team performance. Perhaps rather than relying on ratings from instructors and teaching assistants, a better strategy for evaluating performance would be to have other raters evaluate some project deliverable, for example, a team report or presentation, and provide a performance score based on that artifact.

As previously discussed, the definition of TMS that is used in this study, the processes and sentiments that lead teams to work together to accomplish shared goals, impacted the data that were collected and analyzed. Although qualitative pilot work led to the conclusion that trust and enjoyment networks may have served as good indicators of TMS, it may be that the definition of TMS used in this work was overly reliant on affective relationships between team members, as demonstrated through expressive ties.

An important of this study is the relatively small number of teams that was used. This is a problem with a majority of teams research as it is not the number of participants that impact the power of the study, but the number of teams. As recommended by Huisman (2009), in teams where fewer than 40% of all members provided information, I chose to disregard some of the collected data rather than use multiple imputation techniques. This resulted in fewer assumptions, but also a lower sample size and a greater amount of missing data. Utilizing multiple imputation techniques or incorporating additional types of teams (e.g. professional work teams) could increase the sample size and power of future studies.

Item non-response did not pose any issues in this study. That is to say, if a participant responded to one item, they responded to all items. Unit non-response, a participant neglecting to participate all together, was more problematic. There are several reasons why unit non-response may have occurred. First, data were collected just before the Thanksgiving holiday. In two courses, there were very few students present. This resulted in a higher percentage of missing data in these courses than in the others. Second, some participants had difficulty logging into the survey system. Despite resolving this problem after class, many of these participants simply did not respond to the survey.

Discussion and Recommendations for Future Research

As described in Chapter 5, Lewis's TMS inventory (2003) demonstrated better performance than the proposed TMS network inventory. There are several reasons why this may have occurred. First, scores from the Lewis inventory are obtained from 15 items compared to scores coming from a single item in the proposed TMS network inventory. Although past research has demonstrated that single-item inventories perform reasonably well (Marsden, 1990), it could be that asking more than one question about enjoyment, trust, and coordination could have provided more valid measures of TMS. This should be considered as future work on this topic is conducted.

Furthermore, future studies should follow a similar analysis using valued network ties rather than dichotomizing valued data into strong ties and not strong ties. Although the choice to dichotomize valued data was purposefully selected to minimize the impact of participants providing overly generous ratings (as discussed in Chapter 4), the additional information provided by valued network data may provide more interesting and meaningful results. Furthermore, using valued data where it is available would be consistent with prior recommendations for social network research (e.g. Prell, 2012).

Another opportunity for TMS measurement with a focus at the individual level is considering a dispersion model of team performance (Chan, 1998). Under a dispersion model, within-team variance serves as an operationalization of the construct of interest. For example, within-team variance of items asked at the team-level (e.g. Lewis's 15-item inventory) would function as a representation of a team's TMS. According to Chan, dispersion is a team-level characteristic where low team agreement leads to low levels of the construct of interest. Considering Lewis's (2003) TMS inventory, if a team were to have high levels of agreement on specialization, credibility, and coordination items, that would be evidence of a well-functioning TMS, if some minimum baseline of ratings was achieved. For example, a team who strongly agrees that they have low levels of these aspects is unlikely to perform well.

Similar to results found in the pilot work discussed in Chapter 3, there were greater levels of TMS sum score variance within teams than between teams as within team variance accounted for 63% of the total variance noticed in the Lewis composite score. This further motivates future study of TMS using a dispersion model.

Summary and Conclusions

In this study, I collected and analyzed data to evaluate how the strong-tie density and transitivity of expressive networks (enjoyment and trust) and instrumental networks (coordination) may have demonstrated the strength of a team's TMS and may have influenced team performance. In doing this, I answered a number of research questions.

First, and most importantly, results of this study did not provide evidence that the proposed network measures correlated with performance, overall and controlling for team-level attributes (Research Question 1). Additionally, the proposed network measures demonstrated some evidence of convergent and discriminant validity based on their positive correlation with teamwork and negative correlation with professionalism (Research Question 2). Some evidence of generalizability was obtained as the best performing network measure performed similarly across single-discipline teams and multidisciplinary teams (Research Question 3); however, future evaluations of generalizability should evaluate the performance of the proposed TMS measures among professional teams in addition to student teams. Test-retest reliability was also demonstrated by comparing the responses to the original survey with responses to a re-administered survey among a subsample of the original participants (Research Question 4).

Finally, the proposed measures did not provide evidence of incremental validity over Lewis's previously proposed measure of TMS (Research Question 5), although it should be noted that the structural model proposed by Lewis did not demonstrate goodness of fit with a new dataset (Research Question 6). Though disappointing, these results have some utility in their ability to influence future research on this topic.

This study explored the utility of network items in measuring TMS. Although the proposed network measures do not appear to be valid in their current form, this study provides motivation for future exploration of using instrumental networks or other individual-level attributes (e.g. within-team variance as measured through dispersion models) as measures of TMS. In addition to considering alternative individual-level measures of TMS, future research should evaluate the performance of these proposed measures in different settings (e.g. corporate offices, hospitals, work teams) to provide further evidence of generalizability and to achieve greater statistical power.

To the best of my knowledge, no previous studies have used network measures alone to formally evaluate TMS in small teams. Although the results of this study did not demonstrate that the proposed network items provided valid measures of TMS, this work provides a promising avenue for future research through the method's use of various strong ties between individuals as indicators of performance and measures of TMS.

Appendix A

Transactive Memory System Scale Items (Lewis, 2003)

Specialization
1. Each team member has specialized knowledge of some aspect of our project.
2. I have knowledge about an aspect of the project that no other team member has.
3. Different team members are responsible for expertise in different areas.
4. The specialized knowledge of several different team members was needed to complete the project deliverables.
5. I know which team members have expertise in specific areas.

Credibility
1. I was comfortable accepting procedural suggestions from other team members.
2. I trusted that other members' knowledge about the project was credible.
3. I was confident relying on the information that other team members brought to the discussion.
4. When other members gave information, I wanted to double-check it for myself. (reversed)
5. I did not have much faith in other members' "expertise." (reversed)

Coordination
1. Our team worked together in a well-coordinated fashion.
2. Our team had very few misunderstandings about what to do.
3. Our team needed to backtrack and start over a lot. (reversed)
4. We accomplished the task smoothly and efficiently.
5. There was much confusion about how we would accomplish the task. (reversed)

Items evaluated to measure discriminant validity:

Motivation
1. I was highly motivated to perform this task well.
2. My team was highly motivated to perform this task well

Cohesiveness
1. This group worked as a cohesive team in order to complete the task.
2. There was a feeling of 'team spirit' within this workgroup.

Note. All items use a 5-point disagree–agree response format, in which 1=strongly disagree, 2=disagree, 3=neutral, 4=agree, and 5=strongly agree.

Appendix B

Parameter Estimates of Factor Loadings from Lewis (2003) and Pilot Study

Table 9

Factor loadings from Lewis (2003) laboratory and field studies

Factor Loading	Estimate from Lab Study	Estimate from Field Study
$\lambda_{S1,Spec}$	0.82	0.77
$\lambda_{S2,Spec}$	0.58	0.59
$\lambda_{S3,Spec}$	0.73	0.86
$\lambda_{S4,Spec}$	0.57	0.77
$\lambda_{S5,Spec}$	0.84	0.76
$\lambda_{Cr1,Cred}$	0.68	0.51
$\lambda_{Cr2,Cred}$	0.86	0.93
$\lambda_{Cr3,Cred}$	0.88	0.89
$\lambda_{Cr4,Cred}$	0.46	0.74
$\lambda_{Cr5,Cred}$	0.70	0.77
$\lambda_{Co1,Coord}$	0.61	0.85
$\lambda_{Co2,Coord}$	0.60	0.85
$\lambda_{Co3,Coord}$	0.76	0.65
$\lambda_{Co4,Coord}$	0.92	0.94
$\lambda_{Co,Coord5}$	0.66	0.78
$\lambda_{Spec,TMS}$	0.60	0.66
$\lambda_{Cred,TMS}$	0.64	0.87
$\lambda_{Coord,TMS}$	0.80	0.81

Appendix C

Lewis (2003) Models Evaluated for External Aspects of Validity

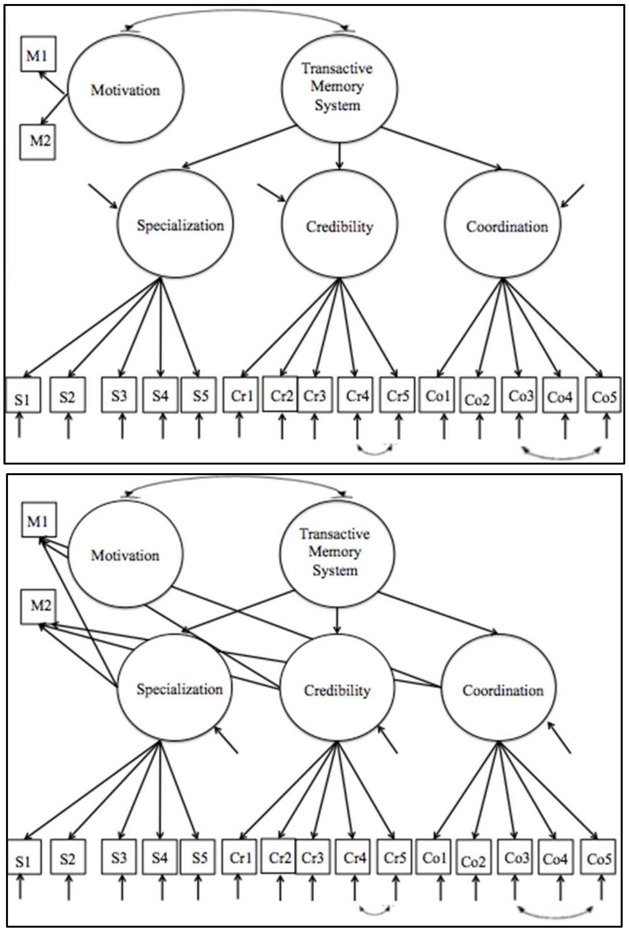


Figure 15. Motivation models

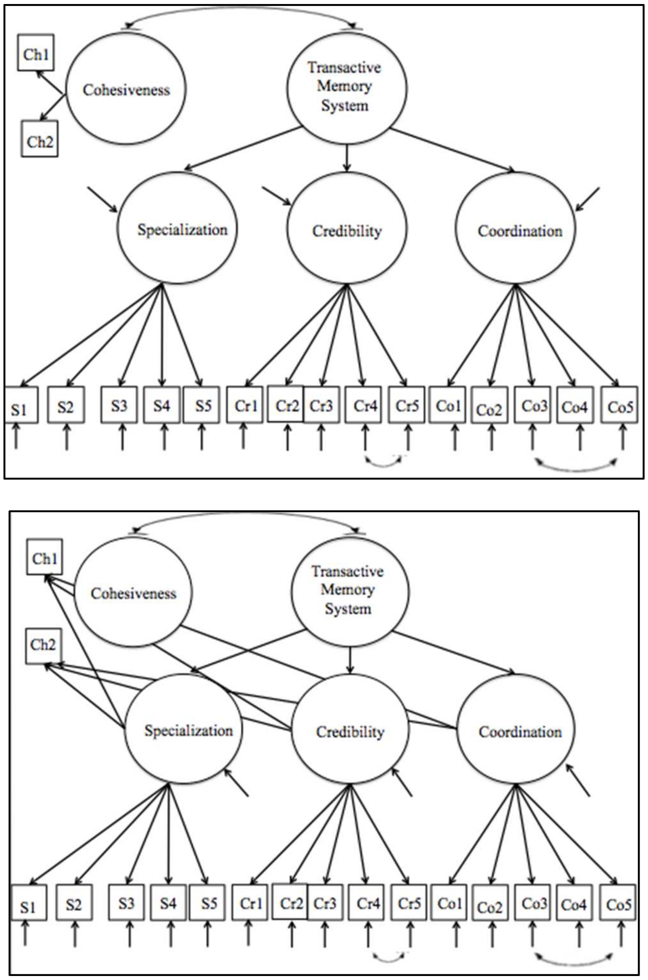


Figure 16. Cohesiveness models

Appendix D

Pilot Study Network Questions and Team Evaluation Scale

1. I regularly seek information from this person (*Information Seeking*)
2. I trust this person (*Trust*)
3. This person and I work together in an efficient manner (*Efficiency*)
4. I enjoy working with this person (*Enjoyment*)
5. I have complementary skills to this person (*Complimentary Skills*)

5. Outstanding; The team is polished, thorough, and on par with what could be expected from a team of professionals

4. Good; The team exceeds expectations for a team of honors students

3. Fair; The team meets expectations for a team of honors students

2. Poor; The team falls short of expectations for a team of honors students but still provides satisfactory work in some areas

1. Unacceptable; The team does not accomplish satisfactory work at any level

Appendix E

Pilot Study Results

Table 10

Standardized factor loadings from Lewis (2003) and pilot study

	Lewis, 2003	Pilot Study
Specialization	0.66	0.44
Credibility	0.87	1.08
Coordination	0.81	0.58
S1	0.77	0.63
S2	0.59	0.19
S3	0.86	0.89
S4	0.77	0.57
S5	0.76	0.53
CR1	0.51	0.64
CR2	0.93	0.80
CR3	0.89	0.99
CR4	0.74	0.81
CR5	0.77	0.83
CO1	0.85	0.69
CO2	0.85	0.75
CO3	0.65	0.63
CO4	0.94	0.83
CO5	0.78	0.80

Table 11

Error variance and factor disturbances from Lewis (2003) and pilot study

	Lewis, 2003	Present Study
Specialization	0.75	0.81
Credibility	0.50	0.17
Coordination	0.59	0.67
S1	0.64	0.60
S2	0.81	0.96
S3	0.51	0.22
S4	0.64	0.68
S5	0.65	0.72
CR1	0.86	0.59
CR2	0.35	0.35
CR3	0.45	0.03
CR4	0.68	0.38
CR5	0.64	0.31
CO1	0.53	0.53
CO2	0.52	0.44
CO3	0.76	0.61
CO4	0.33	0.31
CO5	0.63	0.36

Appendix F

Cognitive Interview Protocol

Overview of Project

I am a Ph.D. student in measurement and statistics in Maryland's College of Education. For my doctoral dissertation, I am evaluating the measurement of transactive memory systems, or the structures and processes teams use to work together and accomplish tasks.

I'd like to use information from your 490H experiences to help me with this. While I know who you are, your identity will remain completely anonymous in any results that I share. In the series of questions that I ask, I may ask you to think about specific team members. You can use their names if you'd like (again, I will not use this information in any results), or you may refer to them in the third person.

In some questions, I will ask you to assign a number or percentage. I know this can be difficult, but please try to approximate as best as you can.

I will be typing notes as we talk. Let me know if you have a problem with this or if it is distracting. After I finish my interviews, I will share notes with you to make sure I understood what you were saying. Do you have any questions for me?

Review Planned Survey Questions

First, I'd like you to review survey questions I plan to ask students in 490H next fall. Could you please read these questions and then talk through your understanding and how you would answer them for different people on your team? (Show 3 network questions to be asked.)

General Teamwork Questions

For each of these questions, think about your experiences last semester in 490H.

Did you usually choose to work on tasks with your team or by yourself? If you had to assign a percentage, what percent of the time did you choose to work by yourself?

Were you more likely to work with people with similar academic backgrounds or different academic backgrounds? Again, if you had to assign a percentage, what percent of the time do you think you chose to work with people coming from different backgrounds?

Besides academic background, did anything else about your teammates influence your likelihood of choosing to collaborate with them? For example, their gender? Did your likelihood of working with people based on major, gender, etc. change as the project progressed or was it consistent from the start?

Sometimes, after you work with someone you feel like you really accomplished something. Other times, you feel like you were just spinning your wheels. Let's call those times when you were very productive "successful". Did you notice differences by teammates in how often your interactions were successful? Can you talk about that, including what about your particular teammates may have caused those differences? If you had to assign a percentage, what percent of total interactions would you guess were successful? And if you had to assign a range by teammate, what would be the lowest and what would be the highest percent of successful interactions?

Functional Background Diversity

As the project progressed, your interactions with your teammates may have changed. For example, you came in with unique knowledge. As the semester went on, did you notice if your knowledge became more specialized (you knew what your teammates knew so you learned different stuff) or more similar (you shared information and, therefore, had more similar knowledge)? Or maybe it stayed the same? Can you talk about that?

Knowledge of Who Knows What

After you interacted with your teammates, you may have learned more about what they knew coming into the project. Do you think this was the case? If you had to assign a number, by what percent did your knowledge of your teammate's knowledge increase each time you worked together? Was this constant throughout the project or greater at the start and less at the end? Was this consistent across each teammate or different? Can you talk through your ideas here?

Desire to Work with Others

You may have had different reactions after working with different people on your team. I'm wondering how interactions with specific teammates influenced future interactions with those people. For example, if you had a successful interaction with someone (using "successful" as we talked about earlier), how much more likely were you to choose to interact with them in the future? How about after a negative interaction? Again, I know it is difficult but if you could assign percentages that would be very helpful.

Belief in Credibility of Knowledge of Others

Again, thinking about successful versus unsuccessful interactions, how did your perception of your teammates' credibility change after these interactions? Did it increase after you had a really great experience collaborating? Did it go down after a negative experience? By about what percent did it increase or decrease after these experiences?

Thinking back to your 490H teams, were you aware of how your team members perceived each other's credibility? For example, let's say you worked with Ann and Ben. Did you know how credible Ann believed Ben to be and vice-versa? A simple "yes, I knew" or "no I didn't know" will suffice. If you want to give more details that is fine, too.

Processes to Assist Coordination

Sometimes, the more you work with people, the easier it becomes to work with them. Did your ability to coordinate with your teammates increase after each interaction? Or maybe it stayed the same or decreased. What do you think? Were there any variations in that increase or decrease by teammate? If you had to assign a percentage, by about how much did it increase or decrease across teammates?

Shared Goals

You signed a charter and completed a scoping document at the start of the semester, but sometimes team goals continue to evolve as the semester progresses. You can think about your project goals as individual if you each wanted to accomplish different things or shared if you were all working toward a common goal. I'm wondering how the "sharedness" of your goals changed after you interacted with your team? Was it different by team member? What, do you think, caused those differences? By about how much did your goal "sharedness" increase or decrease, on average?

Tangible Goals

490H projects come in a wide variety; they range in how tangible success may be, how much creativity is required, and in other ways. How tangible (using a scale of 1-10 with 10 being very tangible) do you think the goals of your project were? Did that change at all throughout the semester?

Creativity Required (Parrotta, Pozzoli, Pytlikova, 2014; Williams & O'Reilly, 1998)

Similar to that question, how much creativity was required by your project (scale of 1-10 with 10 being highly creative). Did that change at all during the semester?

Disruptions

Sometimes, teams gain new knowledge through the course of the project, through their faculty advisors, through class, or in other ways. If your team gained new knowledge, how did that influence your behaviors? Were you more or less likely to interact with teammates after receiving “new to all” information? By about how much more or less?

Appendix G

Proposed TMS Network Inventory Items

1. This person held distinct expertise in a unique area.
2. I thoroughly enjoyed working with this person.
3. I was absolutely confident in relying on the information provided by this person.
4. I was able to accomplish tasks with complete efficiency when working with this person.

Appendix H

Evaluation of Teamwork and Professionalism

Teamwork

Please use the below form to evaluate the team's ability to understand different roles and negotiate conflict.

Role Identification and Delegation

(4) Team clearly defines their roles; these roles are interdependent but not overlapping or redundant; team members are accountable for the completion of all team tasks
(3) Team mostly defines roles; roles have some interdependence but some overlap and redundancies; team members are accountable for the completion of tasks within individual roles
(2) Team defines some of their roles; roles are not interdependent and overlap, redundancies are present; team members are somewhat accountable for the completion of tasks within individual roles
(1) Team does not define roles and is not accountable

Coordination of Tasks

(4) Tasks are well documented and clear interfaces are used for the successful transfer of information between team members
(3) Tasks are generally well documented and reasonably clear interfaces are used to transfer information between team members
(2) Tasks are documented, however, documentation could be more clear; information transfer between the team
(1) Tasks are not documented; information transfer has no coordination

Conflict Resolution

(4) Team is able to identify and address conflict in a timely manner; team develops appropriate methods to resolve conflict
(3) Team is able to identify and address conflict, however conflict remains unresolved
(2) Team is able to identify conflict but is unable to appropriately address or resolve the conflict
(1) Team has conflict but is unable to identify, address, or resolve it

Coherence Around Common Mission

(4) Clear and consistent definition of a common mission or objective by all team members; commitment by team members to help accomplish shared goal
(3) Mostly clear definition of team's mission with consistency from most team members; commitment by most team members to accomplish shared goal
(2) Vague definition of team's mission with inconsistent views amongst team; varying levels of commitment to mission within team
(1) Team does not have a common mission or objective

Professionalism

Please use the below form to evaluate the team's ability to use business etiquette skills and behave in a professional and ethical manner.

Listening
(4) Listens to verbal and visual communication to fully understand a message and reflect the message back to the speaker with 100% agreement
(3) Listens to verbal and visual communication to mostly comprehend a message; is able to reflect the message back to the speaker with general agreement
(2) Listens to verbal and visual communication to comprehend some elements of a message; attempts to reflect the message back to the speaker with some difficulty
(1) Does not listen to verbal or visual communication to sufficiently understand most messages; does not attempt to reflect messages back to speaker

Communication
(4) Verbal and non-verbal communication skills in one-on-one, group, and professional settings demonstrate respect for an audience and convey content so that the audience may fully understand the message
(3) Verbal and non-verbal communication skills in one-on-one, group, and professional settings mostly demonstrate respect for an audience; content is conveyed so that the audience understands the main points of the message
(2) Verbal and non-verbal communication skills in one-on-one, group, and professional settings demonstrate some respect for an audience; content is conveyed so that the audience understands some of the main points of the message
(1) Verbal and non-verbal communication skills in one-on-one, group, and professional settings do not demonstrate respect for an audience and do not convey content so that the audience may understand the message

Attire
(4) Personal appearance is appropriate for the setting and demonstrates care and respect for others
(3) Personal appearance is mostly appropriate for the setting and demonstrates respect for others
(2) Personal appearance is somewhat appropriate for the setting and demonstrates consideration of others
(1) Personal appearance is inappropriate for the setting

Ethics
(4) Ability to fully recognize ethical issues; clear understanding of one's personal ethics and values; clear ability to act on ethical principles
(3) Ability to recognize most ethical issues; general understanding of one's personal ethics and values; ability to act on ethical principles in most situations
(2) Ability to recognize some, but not all, ethical issues; some understanding of one's personal ethics and values; ability to act on ethical principles in some situations
(1) Unable to recognize most ethical issues; limited understanding of one's personal ethics and values; unable to act on ethical principles in most situations

Appendix I

Average team scores including and excluding unit non-response data

	Including Non-Response	Excluding Non-Response
Specialization	4.25	4.19
	3.56	3.75
	3.19	3.25
	3.88	4.00
	3.85	3.92
	3.67	4.33
	3.92	3.60
	3.83	3.92
	3.80	4.13
Enjoyment	4.55	4.44
	3.31	3.92
	3.81	3.50
	3.81	4.00
	4.06	4.17
	3.67	4.17
	3.46	3.50
	4.13	4.24
	3.80	4.13
Trust	4.40	4.31
	3.56	4.00
	3.81	3.58
	4.00	4.25
	3.69	3.58
	3.67	4.33
	3.62	3.50
	4.27	4.40
	4.05	4.19
Coordination	4.35	4.25
	3.31	4.00
	3.69	3.67
	3.81	4.00
	3.69	3.58
	3.89	4.50
	3.23	3.30
	3.97	4.12
	4.15	4.31
Percent Female	0.00	0.00
	0.25	0.33
	0.40	0.50
	0.40	0.50

0.60	0.75
0.25	0.33
0.40	0.20
0.14	0.00
0.40	0.50

Appendix J

Individual Responses to Original and Follow-Up Surveys

Evaluator	Subject	Specialization	Specialization Re-Test	Enjoyment	Enjoyment Re-Test	Credibility	Credibility Re-Test	Coordination	Coordination Re-Test
1	1	4	4	4	5	2	5	4	4
1	2	3	4	4	5	3	4	2	3
1	3	4	5	4	5	4	5	5	5
2	1	4	3	5	5	4	4	4	5
2	2	3	3	4	4	4	4	4	4
2	3	5	5	3	1	4	3	3	1
2	4	4	4	5	5	5	5	4	5
3	1	5	5	5	5	5	5	5	5
3	2	5	5	5	5	5	5	5	5
3	3	5	5	5	5	5	5	5	5
3	4	5	5	5	5	5	5	5	5
4	1	3	4	5	5	5	5	5	5
4	2	3	4	5	5	5	5	5	5
4	3	4	4	5	5	5	5	5	4
5	1	4	4	4	4	4	4	3	4
5	2	4	4	4	4	5	4	5	4
5	3	5	4	2	4	4	4	3	4
5	4	4	4	5	4	5	4	5	4
5	5	4	4	3	4	3	4	3	4
5	6	5	4	5	4	3	4	4	4
6	1	3	4	4	4	4	4	4	3
6	2	4	4	3	4	4	4	4	4
6	3	4	3	4	4	4	4	4	3
6	4	3	3	3	3	4	4	3	3
6	5	4	3	4	3	4	4	4	3
6	6	3	2	4	4	4	4	4	4
7	1	2	2	4	4	3	3	3	3
7	2	3	3	5	5	5	5	5	5
7	3	3	3	2	2	5	5	3	3
7	4	2	2	5	5	5	5	4	4
8	1	4	4	3	3	4	4	2	2
8	2	4	4	4	4	4	4	3	3
8	3	4	4	4	4	4	4	4	4
8	4	4	5	4	5	4	5	5	5
9	1	5	5	4	4	4	4	3	3
9	2	5	5	5	5	4	4	3	3
9	3	5	5	5	5	4	4	4	4
9	4	5	5	4	4	4	4	4	4
10	1	3	3	4	4	4	4	4	4
10	2	3	3	5	5	5	5	4	4
10	3	3	3	4	4	4	4	4	4
10	4	3	3	5	5	4	4	4	4
11	1	5	5	5	5	3	3	4	3
11	2	4	4	3	3	4	4	3	3

11	3	3	3	4	5	5	5	3	3
11	4	4	4	4	4	3	3	3	3
12	1	4	4	3	3	3	3	3	3
12	2	4	4	4	4	4	4	4	4
12	3	4	4	3	3	3	3	4	4
12	4	4	4	4	4	4	4	4	4

References

Adams, K. L., Galanes, G. J., & Brilhart, J. K. (2009). *Communicating in groups: Applications and skills*. New York, NY: McGraw-Hill.

Austin, J. R. 2003. Transactive memory in organizational groups: The effects of content, consensus, specialization, and accuracy on group performance. *Journal of Applied Psychology, 88,* 866–878.

Bormann, E. G., & Bormann, N. C. (1972). *Effective small group communication*. Minneapolis, MN: Burgess Publishing Company.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (Eds.) (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C.: National Academy Press.

Brass, D. J. (2011). A social network perspective on industrial/organizational psychology. *Handbook of industrial and organizational psychology*, *1*, 107-117.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research.* New York, NY: Guilford Press.

Burt, R. S. (2001). Structural holes versus network closure as social capital. In N. Lin, K.S. Cook, & R.S. Burt, (Eds.), *Social capital: Theory and research* (31-55). Piscataway, NJ: Transaction Publishers.

Cannon□Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior*, *22*(2), 195-202.

Caudle, S. L. (2004). Qualitative data analysis. *Handbook of practical program evaluation*, *2*, 417-438.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of applied psychology*, *83*(2), 234-246.

Chen, G., Farh, J. L., Campbell-Bush, E. M., Wu, Z., & Wu, X. (2013). Teams as innovative systems: Multilevel motivational antecedents of innovation in R&D teams. *Journal of applied psychology*, *98*(6), 1018-1027.

Chen, G., & Kanfer, R. (2006). Toward a systems theory of motivated behavior in work teams. *Research in organizational behavior*, *27*, 223-267.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, *20*(1), 37-46.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American journal of sociology*, S95-S120.

Contractor, N., & Su, C. (2011). Understanding groups from a network perspective. In A. Hollingshead & M.S. Poole (Eds.), *Research methods for studying groups and teams: A Guide to approaches, tools, and technologies* (284-310). Abingdon, UK: Routledge.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. Washington, DC: Sage..

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* Toronto: Holt, RineHart, and Winston, Inc.

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

DeShon, R. P., Kozlowski, S. W., Schmidt, A. M., Milner, K. R., & Wiechmann, D.

(2004). A multiple-goal, multilevel model of feedback effects on the regulation of individual and team performance. *Journal of Applied Psychology*, *89*(6), 1035-1056.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. Hoboken, NJ: John Wiley & Sons.

Fang, R., Landis, B., Zhang, Z., Anderson, M. H., Shaw, J. D., & Kilduff, M. (2015). Integrating personality and social networks: A meta-analysis of personality, network position, and work outcomes in organizations. *Organization Science*, *26*(4), 1243-1260.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215-239.

George, J. M. (1990). Personality, affect, and behavior in groups. *Journal of Applied Psychology, 75,* 107–116.

Gibbs, J.P., Martin, W.T. (1962). Urbanization, technology and the division of labor. *American Sociological Review*. *27*, 667–677.

Gino, F., Argote, L., Miron-Spektor, E., & Todorova, G. (2010). First, get your feet wet: The effects of learning from direct and indirect experience on team creativity. *Organizational Behavior and Human Decision Processes, 111*(2), 102–115.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*

78(6), 1360-1380.

Granovetter, M. (1985). Economic action and social structure: the problem of embeddedness. *American Journal of Sociology, 91*(3), 481-510.

Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, *47*(1), 307-338.

Handcock, M., Hunter, D., Butts, C., Goodreau, S., Krivitsky, P., Bender-deMoll, S. & Morris, M. (2016). *statnet: Software Tools for the Statistical Analysis of Network Data*. The Statnet Project. R package version 2016.4.

Hirst, G., Van Knippenberg, D., Zhou, J., Quintane, E., & Zhu, C. (2015). Heard it through the grapevine: Indirect networks and employee creativity. *Journal of Applied Psychology*, *100*(2), 567-574.

Hollingshead, A. B. (1998a). Communication, learning, and retrieval in transactive memory systems. *Journal of Experimental Social Psychology, 34*, 423–442.

Hollingshead, A. B. (1998b). Retrieval processes in transactive memory systems. *Journal of Personality and Social Psychology, 74*(3), 659–671.

Hollingshead, A. B. (2001). Cognitive interdependence and convergent expectations in transactive memory. *Journal of Personality and Social Psychology*, *81*(6), 1080.

Homans, G. C. (1958). Social behavior as exchange. *American journal of sociology*, 597-606.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.

Huisman, M. (2009). Imputation of missing network data: Some simple procedures.

Journal of Social Structure, *10*(1), 1-29.

Jobe, J. B., & Mingay, D. J. (1989). Cognitive research improves questionnaires. *American Journal of Public Health*, *79*(8), 1053-1055.

Johnson, D.W., Johnson, R.T., & Smith, K.A. (2006). *Active learning: Cooperation in the university classroom*. Edina, MN: Interaction.

Katz, R. (1982). The effects of group longevity on project communication and performance. *Administrative Science Quarterly (27)*1, 81-104.

Katz, N., Lazer, D., Arrow, H., & Contractor, N. (2004). Network theory and small groups. *Small Group Research*, *35*(3), 307-332.

Katz, N., Lazer, D., Arrow, H., & Contractor, N. (2005). The network perspective on small groups. *Theories of Small Groups: Interdisciplinary Perspectives*, 277-312.

Katzenbach, J. R., & Smith, D. K. (1993). *The wisdom of teams: Creating the high-performance organization*. Boston, MA: Harvard Business Press.

King, K.G. & Herrmann, J.W. (2015) Learning outcomes for a multidisciplinary undergraduate honors program: Development, measurement, and continuous improvement. *Quality Approaches in Higher Education, 6*(1), 4-16.

Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt Brace.

Kozlowski, S. W., & Bell, B. S. (2003). Work groups and teams in organizations. In W.C. Borman, D.R. Ilgen & E.J. Klimoski (Eds.), *Handbook of psychology*: *Industrial and organizational psychology* (333-375). London, UK: Wiley.

Kozlowski, S. W., & Klein, K. J. (Eds.) (2000). *A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes*. Hoboken, NJ: Wiley.

Kozlowski, S. W., Chao, G. T., Grand, J. A., Braun, M. T., & Kuljanin, G. (2013). Advancing multilevel research design capturing the dynamics of emergence. *Organizational Research Methods, 16*, 581-615.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education*, *79*(5), 540-563.

Kuljanin, G. (2011). A computational study of team collaboration and performance (Unpublished doctoral dissertation). Michigan State University, East Lansing.

Landis, J.R.; Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*. *33*(1), 159–174.

Lee, J. Y., Bachrach, D. G., & Lewis, K. (2014). Social network ties, transactive memory, and performance in groups. *Organization Science*, *25*(3), 951-967.

Lewin, K. (1935). *A dynamic theory of personality: Selected papers*. NewYork: McGraw-Hill.

Lewis, K. (2003). Measuring transactive memory systems in the field: scale development and validation. *Journal of Applied Psychology*, *88*(4), 587.

Lewis, K., & Herndon, B. (2011). Transactive memory systems: Current issues and future research directions. *Organization Science*, *22*(5), 1254-1265.

Liang, D. W., Moreland, R., & Argote, L. (1995). Group versus individual training and group performance: The mediating role of transactive memory. *Personality and Social Psychology Bulletin*, *21*(4), 384-393.

Lincoln, Y. S., & Guba, E. G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Program Evaluation*, *1986*(30), 73 –

84.

Marsden, P. V. (1990). Network data and measurement. *Annual Review of Sociology*, *16*, 435-463.

Marshall, C., & Rossman, G. B. (2014). *Designing qualitative research*. Sage: Washington, DC.

Mell, J. N., Van Knippenberg, D., & van Ginkel, W. P. (2014). The catalyst effect: The impact of transactive memory system structure on team performance. *Academy of Management Journal*, *57*(4), 1154-1173.

Messick, S. (1989). Validity. In R.L. Linn (Ed.) *Educational Measurement* (13-103). New York, NY: Macmillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741-749.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey-Bass.

Moreland, R. L. (1999). Transactive memory: Learning who knows what in work groups and organizations. In L. L. Thompson, J. M. Levine, & D. M. Messick (Eds.), *Shared cognition in organizations: The management of knowledge*. Mahwah, NJ: Erlbaum.

Moreland, R. L., & Myaskovsky, L. (2000). Exploring the performance benefits of group training: Transactive memory or improved communication?. *Organizational Behavior and Human Decision Processes*, *82*(1), 117-133.

Morgan, D. L., & Krueger, R. A. (1998). *Planning focus groups*. Washington, DC: Sage.

Muthén, L. K., & Muthén, B. O. (2012). Mplus. *The comprehensive modelling program for applied researchers: User's guide*, *5*.

Napier, R. W., & Gershenfeld, M. K. (1973). *Groups: Theory and experience*. Boston, MA: Houghton Mifflin.

Palazzolo, E. T., Serb, D. A., She, Y., Su, C., & Contractor, N. S. (2006). Coevolution of communication and knowledge networks in transactive memory systems: Using computational models for theoretical development. *Communication Theory*, *16*(2), 223-250.

Park, G., Spitzmuller, M., & DeShon, R. P. (2013). Advancing our understanding of team motivation integrating conceptual approaches and content areas. *Journal of Management*, *39*(5), 1339-1379.

Parrotta, P., Pozzoli, D., & Pytlikova, M. (2014). The nexus between labor diversity and firm's innovation. *Journal of Population Economics*, *27*(2), 303-364.

Patton, M. Q. (2002). Qualitative interviewing. *Qualitative research and evaluation methods*, *3*, 344-347.

Pearsall, M. J., & Venkataramani, V. (2015). Overcoming asymmetric goals in teams: The interactive roles of team learning orientation and team identification. *Journal of Applied Psychology*, *100*(3), 735-748.

Prell, C. (2012). *Social network analysis: History, theory and methodology*. Washington, DC: Sage.

Qualtrics (2016). Qualtrics. Provo, UT. http://www.qualtrics.com.

R Core Team (2015). R: A language and environment for statistical computing. R

Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Washington, DC: Sage.

Ren, Y., & Argote, L. (2011). Transactive memory systems 1985–2010: An integrative framework of key dimensions, antecedents, and consequences. *The Academy of Management Annals*, *5*(1), 189-229.

Richter, A. W., Hirst, G., Van Knippenberg, D., & Baer, M. (2012). Creative self-efficacy and individual creativity in team contexts: Cross-level interactions with team informational resources. *Journal of Applied Psychology*, *97*(6), 1282-1290.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36.

Samuelson, P. A. (1954). The pure theory of public expenditure. *The review of economics and statistics*, 387-389.

Schuh, J. H., & Upcraft, M. L. (2001). *Assessment in student affairs: A guide for practitioners*. San Francisco, CA: Jossey-Bass.

Simone, S., Campbell, C. M., & Newhart, D. (2012). Measuring opinion and behavior. In B. Knight, G. McLaughlin, R. Howard (Eds.), *Handbook on Institutional Research* (502-522). San Francisco, CA: Jossey-Bass.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research, 96*(1), 21-51.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*.

Cambridge, UK: Cambridge University Press.

Tröster, C., Mehra, A., & van Knippenberg, D. (2014). Structuring for team success: The interactive effects of network structure and cultural diversity on team potency and performance. *Organizational Behavior and Human Decision Processes*, *124*(2), 245-255.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review (61)*4, 674-698.

Van Knippenberg, D., De Dreu, C. K., & Homan, A. C. (2004). Work group diversity and group performance: an integrative model and research agenda. *Journal of Applied Psychology*, *89*(6), 1008-1022.

Van Knippenberg, D., Kooij-de Bode, H. J., & van Ginkel, W. P. (2010). The interactive effects of mood and trait negative affect in group decision making. *Organization Science*, *21*(3), 731-744.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.

Weber, J. M., & Murnighan, J. K. (2008). Suckers or saviors? Consistent contributors in social dilemmas. *Journal of Personality and Social Psychology*, *95*(6), 1340.

Wegner, D. M., Giuliano, T., & Hertel, P. T. (1985). Cognitive interdependence in close relationships. In *Compatible and incompatible relationships* (pp. 253-276). Springer New York.

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G.R. Goethals (Eds.) *Theories of group behavior* (185-208). New

York, NY: Springer New York.

Wegner, D. M., Erber, R., & Raymond, P. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology*, *61*(6), 923.

Williams, K. Y., & O'Reilly III, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research. *Research in Organizational Behavior*, *20*, 77-140.

Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review*, *27*(2), 185-203.