8-5-1999

The Effect of Transactive Memory and Collective Efficacy on Aircrew Performance

Daryl R. Smith
Cedarville University, darylsmith@cedarville.edu

The Effect of Transactive Memory and Collective Efficacy on Aircrew Performance

by

Daryl Raymond Smith

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

1999

Program Authorized to Offer Degree: Business Administration

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
15.Oct.99

3. REPORT TYPE AND DATES COVERED
DISSERTATION

4. TITLE AND SUBTITLE
THE EFFECT OF TRANSACTIVE MEMORY AND COLLECTIVE EFFICACY ON AIRCREW PERFORMANCE

5. FUNDING NUMBERS

6. AUTHOR(S)
MAJ SMITH DARYL R

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
WRIGHT STATE UNIVERSITY

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
FY99-308

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

14. SUBJECT TERMS

15. NUMBER OF PAGES
120

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Daryl Raymond Smith

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the
final examining committee have been made.

Chair of Supervisory Committee:

Terence R. Mitchell

Reading Committee:

Terence R. Mitchell

Sally R. Fuller

John Miyamoto

Date: August 5, 1999

University of Washington

Abstract

The Effect of Transactive Memory and Collective Efficacy on Aircrew Performance

by

Daryl Raymond Smith

Chairperson of the Supervisory Committee:
Professor Terence R. Mitchell
Department of Management and Organization

The use of teams is becoming prevalent in American organizations. The United States Air Force for example, employs aircrew teams on the majority of their aircraft. This thesis focuses on system and motivational variables that influence the performance of aircraft teams. Two potentially important team variables are identified and examined in three research studies. Transactive memory is a system which combines the knowledge possessed by individual team members with a shared awareness of who knows what, who is good at what, and who does what. Collective efficacy is the group's collective belief that it can perform a specific task. This research tests these two constructs as competing constructs in explaining team performance. A laboratory and two field studies are conducted to determine the effects of transactive memory and collective efficacy on team performance. The results indicate that transactive memory has a consistent and positive relationship with performance across studies. However, the relationship failed to reach statistical significance due to small sample sizes. Change in the composition of the team due to turnover is shown to be detrimental to transactive memory. In addition, transactive memory makes important contributions to the team's collective efficacy. In operational environments, collective efficacy is significantly related to higher performance. A confident team is a more effective team. These results are discussed in terms of their theoretical and practical significance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENT

The Book of James says that "every good and perfect gift comes down from above, from the Father of lights in whom is no variableness or shadow of turning". So I must first thank Almighty God who has given me the ability to think and reason and act in a purposeful manner and given me the grace to complete this project. I would like to thank Terry Mitchell whose input and conscientious efforts allowed this project to be completed in my three year time allotment. Terry you are a gifted individual and have been very helpful to me, thanks for caring. To Sally Fuller who is a breath of fresh air in the Business School. I appreciate how well you work with others and for your unique perspectives in this work. To John Miyamoto, John I always enjoyed your classes and you are equally a breath of fresh air for the Psychology Department. Thanks for caring for students and for me. I want to acknowledge the time that was sacrificed by my three wonderful children Ryan, Stephanie, and Andrew. My times with you at Ft. Lawton were the most carefree and enjoyable of my stint in school. I hope I didn't miss too many times that I shouldn't have. Finally, to Laura, one of the most wonderful women to ever walk this earth. Thanks for your support and patience in yet another project. I hope that soon, we will have less career obstacles to climb and more time for each other. That is far more important. I love you.

Chapter 1

Introduction

The use of teams is becoming more prevalent in United States businesses. Osterman (1994) found that more than 50% of U.S. business organizations use teams. Within those, 40% report that over half of the company's employees work in teams. The military employs teams widely as well. The basic Army unit is the platoon. The Navy has SEAL teams and submarine crews, while the Air Force has multiple flight crews. Commensurate with this increased use of teams has been an increase of research on teams. The key words "teams" and "teamwork" resulted in 29 references in the *Academy of Management Journal and Review* since 1990. The number of references increased to 54 when *Administrative Science Quarterly* and *Journal of Applied Psychology* were added. With the increased popularity of the use of teams, it is important as organizational behaviorists to understand what makes an effective team. To this end, theorists have identified several potentially important variables, including transactive memory and collective efficacy, which may separate the exceptional team from the average team.

Transactive memory (Wegner, 1987) is a system which combines the knowledge possessed by individual team members with a shared awareness of who knows what, who is good at what, and who does what. Collective efficacy is the group's collective belief that it can perform a specific task (Lindsley, Mathieu, Heffner, Brass, 1994). Both of these constructs have been shown to be positively related to team performance. However, the majority of the studies conducted have been confined to the laboratory. The purpose of this paper is to examine the role of both transactive memory and collective efficacy in the performance of both nominal and natural occurring teams in both the laboratory and the field.

In the following sections, I define key terms and variables and then present a brief history of both transactive memory and collective efficacy research. I review both theoretical and empirical work. The major focus is upon transactive memory with collective efficacy (the more well established construct) serving as a gauge of the effectiveness of transactive memory in the team - performance relationship. I then propose several hypotheses to extend our knowledge of the importance of transactive memory and collective efficacy as contributors to team performance. Next, I review a series of three studies to investigate these hypotheses. The first is a lab study using a PC-based simulator and actual military aviators working as a team to successfully complete a mission. Studies 2 and 3 are field tests of these constructs using KC-135 air refeuling aircraft crews at two different Air Force bases in order to determine if these constructs help to separate the exceptional from the average aircrew. The results and discussions of these three studies are presented in turn. In conclusion, the results of the three studies are compared and summarized and practical implications, as well as directions for future research, are discussed.

Chapter 2

Literature Review

Teams

Teams Defined. What constitutes a team and does a team differ from a group? There are many definitions of a team, but I choose to use the following: "...a distinguishable set of two or more individuals who interact dynamically, interdependently and adaptively to achieve specified, shared, and valued objectives" (Brannick, Roach, and Salas, 1993, p. 287). The emphasis on interdependence and interaction is what separates a team from a group. For example, a life insurance organization may have a sales group. There may be seven salespersons in the group who each independently seek out new insurance policies from individuals. If each of these persons independently acts in accomplishing his or her tasks, then this is best known as a "group" and not a "team". This label is used regardless of the fact that each individual may share the same building, have a cubicle in the same room, report to the same supervisor, and be referred to as the "sales team" by the CEO. The fact that salespeople are *not* interdependent in accomplishing their tasks or objectives makes this a group. Contrast the sales group with the top management team (TMT) on a strategic planning retreat. This group including the Vice President of Sales, Finance, Accounting, Marketing, and the CEO. All act dynamically and interdependently in pursing valued and shared objectives (in this case setting strategic goals for the company). Probably the clearest example of a team is a coronary surgical unit. Each set of surgeons, doctors, and nurses acts dynamically and interdependently to accomplish an open heart surgery. Interdependence is the key distinguishing feature of a team. Having distinguished the terms, they will be used interchangeably from this point forward. However, the use of the term group for this present study will be synonymous with the term "team" as defined above.

While there has been much written on teams, Organizational Behaviorists have failed to give users of teams direction when assembling and monitoring teams (Church, 1996). One reason for this oversight is the fact that teams, like diamonds, have many facets (e.g. demographics, size, leader vs. leaderless, processes, structure, development, boundaries, context). Theoretical work on teams has offered models of team effectiveness with associated categories of variables important for teams (Sundstrom, DeMeuse, Futrell, 1990; Goodman, Ravlin, Schminke, 1987). For the purposes of this proposal I will briefly cover some of the categories of variables examined in the literature in order to provide a context for the constructs of interest.

Categories of Team Variables

There have been many attempts to theoretically model team effectiveness (Goodman, et al., 1987; Sundstrom et al., 1990; Hackman, 1983). Campion, Medsker, and Higgs (1993) have gone a step further than others by not only presenting a theoretical framework for understanding team effectiveness, but by also testing the framework on diverse samples. They hypothesized that work team characteristics would be related to effectiveness. They found support for this hypothesis with both nonexempt administrative support jobs in a large financial services company (Campion et al., 1993) and exempt professional (knowledge worker) jobs in an insurance company (Campion, Papper, and Medsker, 1996). Their work is helpful in establishing a framework of team variables and thereby providing a context in which to understand the studies that follow. The categories of work team characteristics they studied are a follows: Job Design, Interdependence, Composition, Context, and Process. In the first study Campion et al. (1993) found that Job Design (e.g. self-management, participation, task variety, etc.) was the most important category of variables associated with team effectiveness. The second most important category of variables were Process variables (e.g. collective efficacy, and communication/ cooperation within the team). The Process category reflects those things that go on in the team to influence effectiveness.

Interestingly, in the second study (Campion et al., 1996) the Process variables were found to be the most potent predictor of effectiveness of any category of variables (followed closely by Job Design). Relationships between Process characteristics and effectiveness criteria were even higher with the professional workers than they were for the nonexempt administrative workers. Campion et al. (1996) conclude by vigorously arguing that managers monitor and encourage positive team processes to increase team effectiveness. The manager must concentrate on the Process variables, especially if Job Design is not malleable.

Campion et al. (1996) highlight the importance of the Process category of team variables. The manager who wishes to enhance effectiveness should take heed. Within this category there seems to be at least two different types of variables at work. There is the idea of a network or a system (social support, communication, coordination) and the idea of motivation (group potency or collective efficacy). I will examine the idea that team motivation and team systems/networks are both important in team performance. I have selected one construct from each of these types for examination; transactive memory (a systems type construct) and collective efficacy (a motivational construct). These constructs are expanded upon in the following sections.

Transactive Memory - Theoretical Work

The roots of the construct transactive memory go back to an early and influential theory of group behavior known as the *group mind* (Wegner, 1987). Rousseau (1767) and Hegel (1807) assumed that groups, like individuals, had a form of mental activity that guides action. Many of the early pioneers (e.g. Wundt) who contributed to modern social psychology held this viewpoint (Wegner, 1987). Whereas group mind theorists had emphasized the *similarity* of individual minds as a hallmark of the group mind, transactive memory describes a social network of individual minds that transcends such uniform agreement. A transactive memory system connects disparate minds. It places direct emphasis on the social organization of diversity rather than on the social destruction of diversity (Wegner, 1987). It is important to emphasize that transactive memory is not a group mind in the sense of groupthink (Janis, 1972) or perhaps more accurately stated an overemphasis on agreement and lack of divergent opinions (Fuller & Aldag, 1998). Rather it is a linking of disparate minds in an attempt to process and structure information (Wegner, 1987).

Wegner (1987) was the first to formally propose the construct of transactive memory, which was a fresh approach to the idea of a group mind. The group mind idea had been buried during the behavioral revolution of the 1930s (Wegner, 1987) and is beginning to return to favor (Klimoski & Mohammed, 1994). The study of transactive memory has as its goal the prediction of group (and individual) behavior through an understanding of the manner in which groups process and structure information (Wegner, 1987). As first defined, a transactive memory system is a set of individual memory systems in combination with the communication that takes place between individuals (Wegner, Giuliano, and Hertel, 1985). In sum, transactive memory is concerned with how a team acquires, stores, and handles information.

The term memory is employed because the team must encode, store, and retrieve information as a group, just as in individual memory (Wegner, 1987). But, why the term transactive?

> The transactive quality of memory in a group is evident in the transactions that take place during encoding and retrieval. In transactive encoding, people discuss incoming information, determining where and in what form it is to be stored in the group...the very nature of incoming information can be changed, translated into a form that the group can store. Transactive retrieval, in turn requires determining the location of information and sometimes entails the combination or interplay of items coming from multiple locations (Wegner, 1987, p. 190).

Transactive memory uses an external memory system (Wegner, 1987). We store as much outside our minds as within them. There are two requirements for external memory. First there must be a *label* or retrieval cue (e.g. "Tom's phone number"). The second requirement of external memory is the *location* of the item (e.g. "in the rolodex"). In contrast, internal encoding requires a *label* and the *item*. Therefore, other people can be locations of external storage for the individual. Another person can be used much like a library book. The book can be accessed for the information located within the volume. The interdependence produced by a transactive memory system produces knowledge-holding system that is larger and more complex than either of the individual's own memory systems (Wegner, 1987).

Wegner (1987) goes on to elaborate on the features of transactive memory. First, transactive memory is a property of the team. It is a constructed system which is built up over time by individual constituents. This construction is a fairly automatic consequence of social perception and interaction. Once in place, it can impact what the group as a whole can remember. In a nutshell, transactive memory is a group information-processing system.

A number of researchers have begun to study and theorize about transactive memory, led by Moreland (Moreland, Argote, Krishnan, 1998; Rulke & Rau, 1997; Moreland, Argote, Krishnan, 1996; Liang, Moreland, Argote, 1995; Wegner, Erber, Raymond, 1991). Moreland et al. (1998) have continued to develop this idea that people supplant their own memories with external aids. Members of a team can use other members as a memory aid so that important information will not be forgotten (Wegner, 1987). When individual members cannot remember a certain piece of information or are uncertain about the accuracy of such information, they can turn to another member for help. This transactive memory system combines the knowledge possessed by individual group members with a shared awareness of who knows what.

The definition of transactive memory began to develop more fully with additional theoretical and experimental work. Liang et al. (1995) define transactive memory as "...a combination of the knowledge possessed by particular group members and an awareness of who knows what" (p. 385). Moreland et al. (1996) use a similar definition, where the system combines the knowledge possessed by particular group members with a shared awareness of who knows what. However, they go a step further. They trace the development of transactive memory as a subset of a larger construct known as *socially shared cognition* (to be discussed later). They then infer that the shared knowledge of other group members would include "who is good at what" as well as "who knows what". Rulke and Rau (1997) make it explicit when they define transactive memory as the combination of the knowledge of particular group members combined with a shared awareness of *who knows what* and *who is good at what.*

The above represents the present state of the construct definitionally. I would extend the definition one step further. Since transactive memory deals with the information between team members I would say it also includes "who does what". This information is vital to the effective operating of a team. Recall that the term "team" demands that the members be involved in interdependent tasks. For this interdependent cooperation to occur the team members must realize who does what. Task roles may be clearly defined (e.g. surgeon and nurse) in which case "who does what" is very clear. On the other hand, tasks may be ill-defined and situationally determined (e.g. a Navy SEAL team) whereby team members must determine "who does what". That brings us to the final definition of transactive memory which will be used in the remainder of this proposal. Transactive memory is a system which combines the knowledge possessed by individual team members with a shared awareness of who knows what, who is good at what, and who does what in a team.

Note the words, "combines the knowledge possessed by individual team members". The actual knowledge possessed by individual team members has been shown to be key in identifying expertise in the group. Transactive memory is more than a shared awareness; it also encompasses the individuals' knowledge, much like a library book on a shelf. It is important to know which library book to pull down, but it is also important for the library book to contain the knowledge.

Related Terms

Now that I have defined transactive memory, it may be helpful to differentiate it from related terms. In other words, what is transactive memory not? Transactive memory deals with knowledge among group members, however transactive memory is not the same as tacit knowledge. Tacit knowledge is knowledge that cannot be articulated or codified very easily. This knowledge (usually) comes from direct experience (e.g. riding a bicycle) (Polanyi, 1958, 1962). The idea began in the evolutionary economics field and has been adapted by the organization theorists. It is a useful metaphor which has been applied across organizational levels (Nelson & Winter, 1982).

I see two key relations between tacit knowledge and transactive memory. First, the degree to which a transactive memory system is in place, may dictate how easily an organization can articulate or codify its routines. Routines are very important to organizations and most people discuss tacit knowledge in organizations in the framework of routines. If a good transactive memory system is in place, it may be easier to codify the routine because people can identify where the information lies. Second, the fact that some knowledge is tacit makes transactive memory all the more important. If I can only know some things through direct experience, it is important then that I know "who knows what". I can't know everything, especially things learned through direct experience, but my teammates may possess that knowledge.

As alluded to earlier, transactive memory is not some sort of group mind in the sense of groupthink (Janis, 1972). Groupthink in this context would refer to group unanimity, solidarity in thought and too strong a desire to avoid contradictory ideas or evidence. Recently, Fuller and Aldag (1998) have critiqued the concept of groupthink as it has been articulated by Janis (1972). They cite a lack of empirical support for the concept and propose that 'groupthink' as defined by Janis may not necessarily have negative outcomes. Regardless of the validity of the construct of groupthink, it should be clear that transactive memory is not referring to some type of groupthink construct. Transactive memory does not refer to *similarity* of individual minds, but rather connections of *disparate* minds (Wegner, 1987).

There are a slough of similar terms to transactive memory: shared mental models, team mental models, common cause maps, shared frames, teamwork schemas, and sociocognition (Klimoski & Mohammed, 1994). The work in this area is not well developed and therefore terms are often used interchangeable or without great precision. I will attempt to address many of these terms, while attempting to not get bogged down. The goal is a clear delineation of transactive memory from related terms.

Weick and Roberts (1993) use the term collective mind. They focus on the collective mind as a *system of behaviors* which are *heedfully interrelated.* Their classic example is the aircraft carrier where sailors must work in coordination to conduct successful flight operations. "People act heedfully when they act more or less carefully, critically, consistently, purposefully, attentively, studiously, vigilantly, conscientiously, perniciously" (Weick & Roberts, 1993, p. 361). People can vary to the degree in which they are heedful which directly relates to smooth operations or disasters. This is different from transactive memory which focuses on knowledge and memory location and retention, while the collective mind focuses on behaviors and coordination.

Gruenfeld and Hollingshead (1993) assert that cognition is beginning to be viewed as a collective rather than an individual phenomenon. They go on to say that transactive memory is evidence of this change in approach to cognition. Moreland, et al. (1996) also conclude that cognition can occur as a collective phenomenon, they use the term *socially shared cognition* and place transactive memory as a particular type of socially shared cognition. Gruenfeld and Hollingshead (1993) employ the term, *group sociocognition*, which is "social interaction that leads to an emergence of unique, collectively produced conceptualizations that no individual has to begin with." Group sociocognition differs from transactive memory in that transactive memory is not so concerned about the *production* of information, but rather how the group *stores* and *processes* information among its members. However, it is clear that both of these constructs deal with information and knowledge.

One of the most popular team research topics is team mental models (Klimoski & Mohammed, 1994). Different meanings are attached to this term and the seminal treatment of the topic by Klimoski and Mohammed (1994) does not even attempt to define it. However, Cannon-Bowers, Salas, and Converse (1993) have worked extensively with Navy teams and provide the following definition. Team mental models require that team members hold common or overlapping cognitive representations of task requirements, procedures, and role responsibilities. Klimoski and Mohammed (1994) do define the generic term *mental model* as "a psychological representation of the environment and its expect behavior" (p. 405). So from these definitions we may deduce that transactive memory is a subset of team mental models, but not the same construct.

Team mental models deal with not only with knowledge among team members, but also representations of tasks, situations, response patterns, the environment, and the environment's expected behavior (Klimoski & Mohammed, 1994). Klimoski and Mohammed (1994) suggest that the content of mental models are theories of situations or of actions. So team mental models are focused on *task behaviors* and the *environment* much more so than transactive memory. Transactive memory deals primarily with who knows what, who is good at what, and who does what. It concerns itself with knowledge (or information) held *within* the group and is not focused primarily outside of the group. The environment does supply incoming information into the group that the group must process, but again transactive memory focuses on information inside the group. Since team mental models are focused on certain tasks and how they interact with the environment, it would seem logical that an effective transactive memory system would be critical to an effective team mental model. Transactive memory would store the knowledge which would allow the team to interact with the task and environment. But again, team mental models encompass more than transactive memory systems, the environment and tasks. Admittedly, the lines between transactive memory and team mental models become blurred, but I am taking a "first cut" at separating the two. The field as a whole has not yet clearly delineated the two (or made an attempt to do so).

Shared mental model is the final term to be discussed. Shared mental models may be defined as follows, "...group members typically have some sort of organized knowledge structures relating to various aspects of the group's situation, such as their task, their environment, and their fellow group members" (Peterson, Mitchell, Thompson, Burr, 1996, p. 5). Again we see the notion that shared mental models deal not only with knowledge about fellow group members, but also about the task and the environment. This is what separates transactive memory from shared mental models. Transactive memory deals primarily with knowledge and information among group members and is focused within the group. Now that transactive memory has been clearly delineated, I return to the discussion of theoretical treatment of transactive memory.

Transactive memory features

How exactly is a transactive memory system constructed? It begins when individuals learn something about each others' domains of expertise (Wegner, 1987). Stereotypes are the default (Wegner, et al., 1991) and serve as the first building block, though they are not necessarily accurate. Over time conversation and observation allow members to discern with precision who is expert in what domain. This perception of relative expertise of self and others requires self-disclosure (Wegner et al., 1991). The individual with the relevant expertise can become the storage tank for that knowledge. Knowledge of team member's *access* to information is also an important tool in building transactive memory. Access to information can lead to responsibility for the information. If no one individual seems to "be the expert" and access to information is not clear, *negotiated entries* occur (Wegner et al, 1991). In negotiated entries one person agrees to accept responsibility for the information. Responsibility is key in building the transactive memory system, "the system can be built because individuals in a group accept responsibility for knowledge" (Wegner, 1987, p. 194).

Rulke and Rau (1997) found support for the theoretical explanation for how transactive memory systems form. They found that early in group interactions on a new task, group members spent a great deal of communication on declaring expertise and coordinating / planning. The information encoding process seems to consist of small spiral encoding cycles of question-expertise-coordination.

In sum, Wegner (1987) asserts that the person who will be the acknowledged location of a set of labeled knowledge will primarily established through expertise. Rulke and Rau's (1997) work supports this contention. If no expertise is established the group will rely on circumstantial knowledge responsibility (or individual access to information). Or, the group will fall back on how the knowledge has been encountered in the group (e.g. the finance reports seem to come through Dave). Finally, if expertise and access are equal, the group must negotiate the entry to decide which individual will hold responsibility. "An effective team will not leave responsibility for information to chance" (Wegner, 1987, p. 192)

Work on antecedents and consequences of transactive memory is in its infancy. Antecedents that have been proposed are mutual self-disclosure, frequent transaction and dialogue, time, communication, and checks and interventions (e.g. teacher to student, "are you with me") (Wegner, 1987; Wegner, et al., 1991). Wegner (1987) proposed that satisfaction may be one consequence of a transactive memory. He reasoned that clear differentiation of expertise should in turn lead to satisfaction. Furthermore, a mature transactive memory system may be a sign of a successful team or relationship.

Liang, et al. (1995) theorized that a developed transactive memory system should be exhibited by three key behavioral features. In other words, the knowledge of "who knows what" could be exhibited through three behaviors. From the behaviors, it can be inferred that a strong transactive memory system exists (Moreland, et al., 1998). The first feature is *memory differentiation* which is the tendency of group members to specialize in remembering distinct aspects of the task. Secondly, a developed transactive memory system should result in *task coordination* where group members work smoothly together on a task (e.g. greater cooperation, less confusion, etc.) Finally developed transactive memory systems should exhibit high *task credibility* where group members trust one another's knowledge about the task (e.g. few challenges to declared expertise). Two studies have shown support for these key features (Liang et al, 1995; Moreland et al, 1996).

Moreland et al. (1998) also theorized that teams with developed transactive memory systems will reflect three (knowledge) indices concerning their belief system about group member expertise. They assert that teams with transactive memory will exhibit a *complexity* of beliefs about expertise. These teams will also show a higher level of *accuracy* of those beliefs and a high level of *agreement* among team members concerning those beliefs. Conversely a team with little or no transactive memory will have a simple and incomplete set of beliefs concerning expertise. These beliefs will not be accurate nor will they exhibit high levels of agreement. Moreland et al. (1998) provide one study as empirical support for these knowledge indices.

Advantages and Disadvantages

Wegner (1987) enumerates several potential benefits of transactive memory. First he points out the integration of knowledge: useful creative products can be produced by transaction--they help manufacture new knowledge for the group. This is similar to Gruenfeld and Hollingshead's (1993) idea of group sociocognition. Transactive memory also allows individual team members to gain access to new areas of expertise. Finally, others in the group may 'catch', (i.e. note, acquire, and store) incoming information that any one individual may miss, ensuring it is available for future team use.

There are also several drawbacks to a strong transactive memory system (Wegner, 1987). Errors can occur at all three stages of group information processing; encoding, storage, and recall. Obviously this occurs at an individual level as well, but can be compounded by group level errors. There is also a potential new source of error, incomplete specification of knowledge path responsibility (i.e. where this knowledge is to be stored and who is responsible for it). Information may be inadvertently channeled away from the expert, however this is more likely under a developing and immature transactive memory system. Overconfidence can lead to an over estimation of its capability by group members (i.e. trust that someone else has the information, when they do not). Finally, and I think most importantly for organizations, a strong transactive memory system infers that turnover can have grave effects. When a person leaves the team, they depart with that store of knowledge leaving a gap in the transactive memory system.

Transactive Memory - Empirical Work

Since transactive memory is a relatively new construct, little empirical work has been done. For this reason the entire body of empirical work to date will be summarized.

Wegner (1987) began transactive memory work on dating couples. He studied recall in dating couples. Subjects were asked for area of expertise for self and partner. The individuals then viewed items for either one minute or 30 seconds. This resulted in a 2 X 2 design (Expertise--partner/self and Circumstantial Responsibility--partner/self). Self expertise led to the greater number of remembered items, which is no surprise. However, when subjects were circumstantially responsible for a topic (i.e. allowed greater viewing time) they remembered more when they believed their partner was not an expert in that topic. When they considered their partner an expert they "let the information pass by" assuming that the partner would pick it up. This provides evidence that expertise seems to indicate where knowledge will be stored in the transactive memory system.

Wegner's (1987) first study dealt with intact couples. He and his colleagues (Wegner et al., 1991) extended the study by comparing recall in intact (or natural) couples versus those put together simply for the study (i.e. impromptu couples). They also manipulated expertise by either assigning expertise (or responsibility, without regard to actual expertise) or not assigning expertise. This resulted in a 2 X 2 design as well (Pairs--natural or impromptu and Expertise--assigned or none). They found an effect of expertise, natural couples remembered more than impromptu couples with no assignment. Impromptu couples remembered more than natural couples with assignment. There was also an interaction, natural couples without assignment remembered more than with assignment. In impromptu couples there was no difference in assignment mode (assignment was non-significantly greater).

Wegner et al. (1991) drew several conclusions from this study. First, for assignment to improve memory performance in teams, it may require time and practice. The assignment did help improve memory in impromptu couples, but it did not reach statistical significance. The finding that assignment of expertise or responsibility (regardless of expertise) was interesting. Natural couples without assignment remembered significantly more than with assignment. It seems that items which fall within the domain of expertise where given less than the usual attention, why? Perhaps new assignment introduces uncertainty, or new assignment may introduce overconfidence to ignore usual items. Perhaps, assignment interrupts the flow of normally fluid cognitive processes. Wegner et al. (1991) called for further research into the time and course of transactive memory development. It does seem clear that imposing artificial structure into a natural team is counterproductive. While assignment of structure to newly formed teams may be helpful.

Rulke and Rau (1997) answered this call for further research. They used a laboratory study of undergraduates who were trained as individuals or in groups. The task was the construction of an AM radio. One week after initial training, the subjects again assembled a radio for the criterion trials. The training groups remained intact and the individually trained subjects were put into groups for the radio assembly. Subject's interaction was videotaped and later coded and analyzed. Expertise was found to be key to transactive memory systems. Declaring expertise and coordinating / planning were found to be the largest categories of sentences spoken. Declaring expertise and acquiring information about a domain of expertise took place during the earlier rather than later periods of group interaction. Category of sentences spoken was influenced by category spoken in previous periods. The encoding process seems to consist of small spiral encoding cycles of question-expertise-coordination. The results of the study "suggest that transactive memory is developed when shared experience is present and through an transactive encoding process" (Rulke & Rau, 1997). The results also give direct support to Wegner's (1987) hypothesis that discovery of expertise is a primary vehicle to the encoding of transactive memory.

Moreland and his colleagues have been involved with a serious of laboratory studies to investigate transactive memory (Liang et al, 1995; Moreland et al, 1996; Moreland et al., 1998). Their studies have been centered around two general hypotheses (Moreland et al, 1998). One is that groups will perform better when their members are trained together rather than apart. Two, the benefits of such training will depend largely on the operation of transactive memory systems. Training people in groups allows the interaction, self-disclosure, and communication necessary to build transactive memory systems (Wegner et al., 1991).

All three of the Moreland studies utilized the AM radio assembly task employed by Rulke and Rau (1997). Liang et al. (1995) compared teams originally trained as individuals versus teams that were trained as teams. Through direct performance measures (e.g. number of errors in assembly, procedural recall) it was clear that teams trained as a group performed better. Meanwhile, videotape recording and analysis revealed that teams trained together exhibited better memory differentiation, task coordination, and task credibility. From these behavioral features, it was concluded that such teams had stronger transactive memory systems. Regression analysis revealed that group training improved group performance primarily through transactive memory systems. Control variables included task motivation, group cohesion, and social identity. These control variables did not effect the training - performance relationship. This was the first study to demonstrate transactive memory with more than two people.

The next study in the series was nearly identical to the first (Moreland et al., 1996) and was designed to rule out alternative explanations (i.e. those other than transactive memory) for the results in study one. The second study employed four groups of training: individual, team-building, re-assignment, and integral team. In the individual condition, subjects were trained on the radio task individually and tested in a group one week later. The second condition was identical to the individual condition, except that after the individual training session, groups were formed and participated in a short team building exercise (to develop a mentoring quiz for seniors to use during freshman orientation). This exercise was used to encourage group development without providing information about "who knows what".

The group training condition (integral team) was identical to study one (Liang et al., 1995). A new condition was identical to the group training condition except that subject teams were unexpectedly scrambled one week later. The researchers again found that the group training condition produced the highest performance (the other conditions did not differ) and that this performance difference was due to transactive memory. For example, assembly errors were reduced by 50% in the group training condition. Memory differentiation, task coordination, and task credibility were all higher for the group training condition.

Evidence for transactive memory in the first two Moreland studies relied upon the inference of transactive memory as a result of the three behavioral indications of memory differentiation, task coordination, and task credibility. Study 3 (Moreland et al., 1998) was designed to more directly tap the knowledge inherent in a developed transactive memory system. The researchers wanted to get at exactly what team members know about one another. This study used similar procedures to the previous studies and employed two training conditions (group or individual). Subjects were then brought back the next week. Subjects were then given a knowledge questionnaire (discussed below) to assess their knowledge of other group members. Next subjects completed a procedural recall sheet and then assembled a radio. However, contrary to what they had been told in week one (and was done in the other two studies) this recall sheet and subsequent assembly of the radio was done individually rather than in a group. This was done in order to actually assess individual member's knowledge of radio assembly procedures.

The knowledge questionnaire was used to produce three indices for each group. The *complexity* of group member's beliefs about one another's radio expertise; the *accuracy* of those beliefs; and the level of *agreement* within a group about the distribution of expertise. As hypothesized, members whose groups were trained together rather than apart had significantly greater complexity, accuracy, and agreement in their knowledge of other group members. This is direct evidence that group training helps to develop transactive memory systems. Furthermore, the researchers found that these direct measures (i.e. knowledge indices) of transactive memory were positively and significantly correlated with the behavioral (or indirect) evidence of transactive memory (i.e. memory differentiation, task coordination, and task credibility).

Though technically it was not transactive memory as defined in this present proposal, Peterson et al.'s (1996) study on shared mental models (and group efficacy) sheds some light on transactive memory. The researchers studied student groups working on quarter-long research projects. They found that both group efficacy and shared mental models predicted performance (measured as the final grade). Specifically it was early group efficacy and shared mental models late in the quarter which predicted performance in the longitudinal study. Furthermore, the predictors were not independent of each other. Early group efficacy predicted later shared mental models which in turn predicted performance. What is interesting is that the researcher's operationalization of shared mental models was information-centered, making it very similar to transactive memory.

Peterson et al. (1996) assessed shared mental models through a set of knowledge type questions. Individuals in each team were asked to distribute points to each group member on how much they contributed to the five task components of the project. This *Disagreement over contributions* score can be thought of as "who does what". *Egotism* was measured for each group as well to determine the extent to which group members inflated their contributions. Again, this is a measure of "who does what". One would expect that teams with better transactive memory systems would have higher agreement on "who does what". A second shared mental model measure was used to rate the importance of each of the five task components in order to produce an outstanding final project. High level of agreement on this measure would indicate a shared awareness of "who knows what", in this case, do we all know what is important? Groups with well developed transactive memory systems should have high agreement on this score as well.

Again, groups with better shared mental models (similar to transactive memory) late in the quarter, produced better final projects. Shared mental models (or transactive memory) was hypothesized to improve performance by improving group coordination. This idea is supported by the work of Liang et al. (1995). One of their behavioral indices of the presence of transactive memory, task coordination, was found to be related to higher performance.

The closest that transactive memory empirical work has come to examining natural teams in a field setting is Hollingshead's (1998) recent work with clerical office workers. Hollingshead used clerical workers from a large university in a word memorization task of work related words in a lab study. She found that subjects learn and recall more information in their own areas of expertise when their partner had different, rather than similar, work-related expertise. Furthermore, this effect reverses for recall of information outside work-related expertise. These findings are similar to Wegner et al.'s (1991) work with dating couples.

Though all the subjects were clerical workers, the workers were placed in nominal teams by the researcher. This was done to manipulate the expertise of the others on the team. This work was conducted as a lab study. However in addition to the lab study, Hollingshead also surveyed workers concerning their own natural work groups. The survey measured various items including group tenure and group size. The survey also included a self report of the extent to which participants knew about each of their coworker's areas of work-related knowledge and job responsibilities (seven-point scale); and the extent to which there was shared agreement in their work group about members' work-related expertise and job responsibilities (seven-point scale). Hollingshead then correlated the self-reports with the predictor variables of group tenure and size. Results indicate that group size is negatively correlated with knowing about other group member's expertise as hypothesized. In contradiction to another hypothesis, it was found that tenure was negatively correlated with perceived agreement in their work group about members' expertise and job responsibilities. This was the first attempt to examine natural teams. Though a step in the right direction, the study suffered from common method variance, with the participant rating both the predictor and criterion variables. Furthermore, these findings were based on the reports of only one person from the natural team.

Summary of Empirical Evidence

So in sum, what do we know empirically about transactive memory? We know that it has been exhibited in laboratory studies (Wegner, 1987; Wegner et al., 1991; Liang et al., 1995; Moreland et al., 1996; Rulke & Rau, 1997; Moreland et al., 1998). All of the studies have used undergraduate students and have been limited to memory recall or completion of the same complex task (i.e. radio assembly). Transactive memory has been exhibited in pairs (Wegner, 1987; Wegner et al., 1991) or in teams of three (Moreland et al., 1998). Transactive memory has been exhibited directly by memory recall in dating couples (Wegner, 1987), indirectly through behavioral indices (Liang et al., 1995; Moreland et al., 1996), and directly through knowledge indices (Moreland et al., 1998). Rulke and Rau (1997) and Wegner et al. (1991) have demonstrated that the recognition of expertise seems key to the transactive memory process and that transactive memory seems to develop through interactions over time.

There has been little or no research in the following areas. There has been no field studies or simulations used to test transactive memory. Transactive memory has not been tested on groups of more than three people and little has been done with natural teams. Transactive memory has not been tested when roles are well-defined, though Wegner et al. (1991) did impose artificial structure that could be remotely related to roles. The effect of turnover was examined in passing (Moreland et al., 1996), but nothing is known of its effects on transactive memory in well-developed teams over time. A related turnover issue is one of teams that are reconstituted (formed and reformed) on a regular basis. Clearly there are many issues to be explored in the transactive memory field.

Collective Efficacy - Theoretical Work

The decade of the 1990s has brought considerable interest in the construct of collective efficacy. The term has roots in Bandura's work in self efficacy (Bandura, 1977,1986). Bandura (1977) originally defined self efficacy as, "the conviction that one can successfully execute the behavior required to produce (particular) outcomes." Wood and Bandura (1989) later added that self efficacy refers to beliefs in one's capabilities to mobilize the motivation, cognitive resources, and course of action needed to meet given situational demands. Bandura (1982) proposed that efficacy may also operate at the group level and since that time researchers have strived to understand its relationship to group performance. By 1997, Bandura asserted that indeed efficacy does operate at the group level and has similar sources, serves similar functions, and operates through similar processes as does self efficacy.

Researchers have used a number of terms centered around this idea of member's beliefs about the group: collective efficacy, group efficacy, collective or group esteem, group potency, and group aspiration level (Guzzo, Yost, Campbell, 1993; Lindsley et al. 1994; Mischel & Northcraft, 1997; Little & Madigan, 1994; Gibson, 1996). Unfortunately these terms have often been used loosely, with little precision. Lindsley et al. (1994) do a good job in delineating between the terms. They define collective efficacy as the group's collective belief that it can successfully perform a specific task. Note, they assert that collective efficacy is task specific. In contrast, with group aspiration level, group members unanimously agree on a specific performance target. The aspirations are exact statements of performance goals rather than cognitive beliefs about the group's capability to accomplish particular levels of performance. Group esteem is the extent to which individual's generally evaluate their social group positively. It is a more global concept than task-specific collective efficacy, and refers to the value of the group rather than the group's expected effectiveness in performing a task. Group potency is a group's shared belief that it can be effective. It is a more generalized belief in effectiveness than collective efficacy. Potency reflects a general assessment of the likely effectiveness of the team across situations, whereas team efficacy reflects shared performance expectations for a relatively specific situation. Potency is meant to refer to a shared belief about general effectiveness across multiple tasks encountered by groups in complex environments.

This study is concerned with the construct of collective efficacy. Though not always done in the literature, I wish to delineate between collective and group efficacy. Mischel and Northcraft (1997) define the term *collective efficacy* as an individual's belief that his/her group or team can execute a task successfully. Individual beliefs are aggregated to determine collective efficacy. Gibson (1996) defines *group efficacy* as the group's collective estimate (as a consensus) regarding the group's ability to perform a task objective. I will follow their lead and use the terms in this manner.

There has been great debate over whether efficacy should be measured at the individual or group level (Gibson, 1996). In other words, should individual member's efficacy beliefs concerning the group be aggregated or should the group reach a consensus on efficacy level? Some researchers assert that either method is fine because both predict performance (Peterson et al, 1996; Guzzo et al., 1993). Others (Mischel & Northcraft, 1997) urge adoption of the aggregate collection of individual beliefs concerning group efficacy. The ratings are done individually and the referent is the group. They prefer this method because individual beliefs drive and direct individual effort and they assert that aggregation is a better predictor than consensus. Gibson (1996) counters that consensus is a more appropriate method because it more accurately reflects an attribute of the group, and group efficacy is a group level construct.

The danger in using the group efficacy consensus approach is the possibility of the group arriving at a "politically correct" answer (Guzzo et al., 1993). In other words, the group tendency to present a socially desirable answer in response to a set of demand characteristics. The subject pool in this proposed study could be particularly prone to these socially desirable answers. Air Force members are taught to present a confident air and bravado. High goals are expected as reflected in the Air Force motto, "Aim High". I feel that this may compromise the validity of a group efficacy measurement. An anonymous individual collective efficacy estimate is one way to avoid this pitfall. Furthermore, collective efficacy has been shown to be more predictive of performance than group efficacy (Mischel & Northcraft, 1997). In light of these issues, collective efficacy will be used in this research and the method of aggregation of individual responses will be used to determine it.

In sum, collective efficacy gets at the task confidence of the group, while transactive memory gets at the knowledge within the group. Collective efficacy is the group's collective belief that it can successfully perform a specific task. Transactive memory is a system which combines the knowledge possessed by individual team members with a shared awareness of who know what, who is good at what, and who does what. These are clearly different constructs. However, it may be that high transactive memory could contribute to higher collective efficacy. A team where individuals realize who knows what, who does what, and who is good at what may feel more confident about task accomplishment. This idea is alluded to by Mischel and Northcraft (1997).

Collective Efficacy - Empirical Work

Collective efficacy research has demonstrated that collective efficacy exists as a group attribute and that it predicts performance (Bandura, 1997). Collective efficacy has been shown to be a predictor of performance in large organizations such as elementary schools (Bandura, 1993). Staff member's collective efficacy concerning their ability to motivate and educate students was strongly related to the school's academic performance. The staff's collective efficacy was more important in predicting academic performance than was the student body composition (socioeconomic status and racial composition).

At the team level collective efficacy has also been shown to be an powerful predictor of performance for both sports (Hodges and Carron, 1992) and work (Little & Madigan, 1994) teams. Little and Madigan (1994) studied eight manufacturing teams of twelve employees each over an 18 month period. These were self-managed work teams in a continuous manufacturing plant. Through surveys and structured interviews, it was found that collective efficacy was highly correlated with mean performance ratings made by independent line leaders.

These studies have clearly demonstrated a collective efficacy - performance linkage. Others (Knight, Durham, Locke, 1996; Gibson, 1996) have examined possible mediators or moderators of this relationship. One of more established constructs in the prediction of performance is goal setting. Knight et al. (1996) studied the relationship of goal setting, strategic risk, and collective efficacy (they use the term team efficacy) using 88 3-person teams in a computer tank simulation. Controlling for ability, teams with higher collective efficacy chose harder goals, which led to riskier strategies and higher performance. Goals set by the team were strongly influenced by collective efficacy. So it seems that collective efficacy may translate into higher performance through higher set team goals. In addition, performance feedback is an important component of goal setting. Prussia and Kinicki (1996) demonstrated that the impact of performance feedback on group brainstorming performance operated entirely through its effect on collective efficacy and affective reactions.

Gibson (1996) cleverly examined cultural differences in the collective efficacy - performance relationship by studying U.S. and Indonesian nursing teams and simulations using U.S. and Hong Kong management teams. Gibson found that task interdependence, collectivism and differentiation (how we seek for information - either self-reliance or looking to others) moderate the collective efficacy - performance relationship. When task interdependence is high, higher collective efficacy was related to higher performance. This relationship disappeared under conditions of low task interdependence (where the task does not require teamwork). Collective efficacy seems to be more effective for those high in collectivism. Furthermore, for those that tend to look to others for information, higher collective efficacy led to higher performance. Whereas, for those that look to self for information, higher collective efficacy led to lower performance. Gibson was the first to show that under certain conditions higher collective efficacy can actually lead to lower performance. She hypothesizes that this occurs when the team high in collective efficacy refuses to look to others for information.

Perhaps the study which is most germane to the work presented here was conducted by Lindsley et al. (1994). They had 54 two-member teams complete a series of six preprogrammed ten-minute missions on a PC-based combat jet simulator. Collective efficacy and potency were assessed using survey measures along with objective indices of team task performance. They found that collective efficacy and potency (a more global confidence score as described earlier) were distinguishable constructs and that performance related more significantly with efficacy than potency. Collective efficacy was found to have a significant positive influence on performance development. Additionally, performance related significantly to subsequent collective efficacy levels.

In sum, empirical research has supported the idea that collective efficacy predicts performance (Bandura, 1997, 1993; Little & Madigan, 1994; Hodges & Carron, 1992; Lindsley et al. 1994). Under certain conditions, goal setting may mediate the collective efficacy -performance relationship. Additionally, cultural factors such as collectivism can moderate the relationship. Collective efficacy predicts performance, but is also effected by previous performance (Lindsley et al., 1994) and performance feedback (Prussia & Kinicki, 1996)

The Relationship of Transactive memory and Collective efficacy.

What is the potential relationship between transactive memory, collective efficacy, and performance? One potential link is that collective efficacy may lead to transactive memory. This seems to be what Peterson et al. (1996) found in their work on collective efficacy and shared mental models. Early group efficacy led to a higher level of mental models later in the quarter which in turn predicted group performance. Though not exactly the same construct as transactive memory, Peterson et al.'s (1996) measurement of mental models shared many similar characteristics with transactive memory. Specifically the group's shared mental models focused on items like "who knows what" and "who does what". Peterson et al.'s (1996) teams had no previous history with one another before the beginning of the study (i.e. the academic quarter) and little history before the first measurement of efficacy and mental models. It seems unlikely that a team with little or no prior history could have developed mental models (or transactive memory), which would predict collective efficacy later in the quarter.

More likely it seems that transactive memory and other factors lead to collective efficacy. First, intuitively we suspect that experience and ability lead to greater collective efficacy. Theoretically, Mischel and Northcraft (1997) assert that there are two components of collective efficacy: 1) Collective Task Efficacy and 2) Collective Interdependence efficacy. Collective task efficacy is the team member's estimation of whether they have the KSAs necessary to perform the task. Certainly, if a group does not feel they have the ability and experience necessary to provide the KSAs for the task, collective efficacy would suffer. The Air Force focuses a great deal of effort and attention on proper flight training. Anecdotally, the airline industry prefers to hire ex-military aviators, primarily due to their outstanding training, experience, and discipline. It would seem clear that most aircrew teams would feel that they have the necessary KSAs to perform the task (i.e. collective task efficacy). Mischel and Northcraft's second component, Collective Interdependence efficacy, concerns the issue of the team possessing the KSAs necessary to work together to accomplish the task. Though not as concentrated as the aviation skills training, the Air Force has established a program (referred to as Crew Resource Management or CRM) in order to provide aircrews' with the KSAs necessary to work effectively as a team.

The seeds of collective efficacy lie in Bandura's (1982) work on self-efficacy. It is reasonable to assume that many principles of self efficacy translate directly to the group level as Bandura (1997) asserts. Bandura asserts that four categories of experience help to develop self efficacy. The primary and most important is what Bandura refers to as enactive mastery. Enactive mastery consists of previous personal attainments. It is clear that experience and ability are an integral part of those previous personal attainments.

Enactive mastery and other categories of experience provide informational cues for the three types of assessment processes used to form self efficacy (Gist and Mitchell, 1992). One of the assessments involved is the assessment of personal (or team) and situational resources/constraints. Again, assuming self efficacy processes operate at the group level, the team assesses whether they have the resources (including the KSAs) necessary to successfully complete the task in view of the constraints of the situation. A team that is lacking the necessary abilities and experience is unlikely to have a positive assessment of their resources and, therefore, will likely have a lower state of collective efficacy. In sum, it seems that experience and ability have an important theoretical role in the development of collective efficacy.

What other theoretical processes help to develop collective efficacy besides ability and experience? It appears that a well developed transactive memory system will provide a solid basis for formation of high collective efficacy. A case may be made for this assertion, using the theoretical approaches just touched upon.

Transactive memory is a knowledge system of who knows what, who does what, and who is good at what. This knowledge system would be very helpful in assessing Mischel and Northcraft's (1997) two components of collective efficacy. Does the team have the KSAs to perform the task (collective task efficacy) and more importantly does the team possess the KSAs to work together effectively (interdependence efficacy)? Being aware of who knows what, who does what, and who is good at what would allow team members to more confidently assess these two components.

Furthermore, it would seem that a higher level of transactive memory would allow the team to have a better sense of Bandura's (1982) enactive mastery. A team that has been together, worked together, and thus developed a transactive memory system should have more enactive mastery to draw upon.

Finally, and perhaps most importantly, a well developed system of transactive memory would greatly enhance a team's ability to assess team and situational resources/constraints. A team that has been together and worked on a task together in the past, would have a more accurate gauge with which to judge team and situational resources/constraints. Transactive memory should be invaluable in this respect. If these arguments hold true, there should be a high relationship between transactive memory and collective efficacy.

Chapter 3

Research Model and Hypotheses

Given the above literature review, I now propose an overall theoretical model which will serve as the impetus for the hypotheses. Figure 1 depicts the model. The model is to be tested on both nominal and actual teams with clearly defined roles. The teams will be composed of two or more members. The relationships depicted should be evident in both the lab and the field.

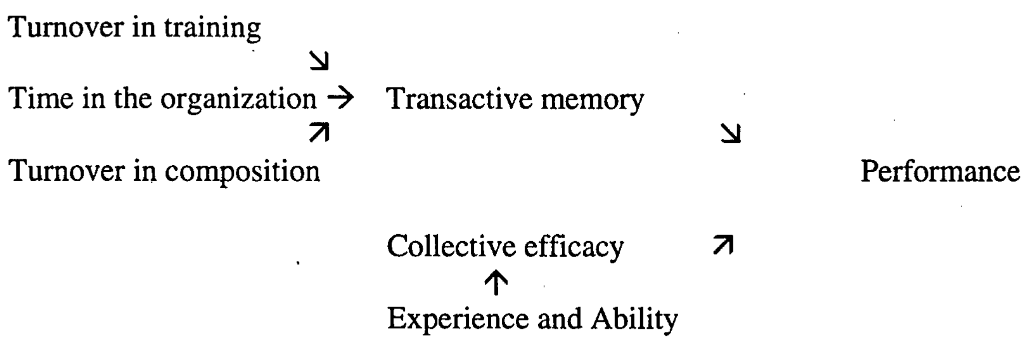

Figure 1. Transactive memory and collective efficacy; dual influences on performance.

The model depicts how two types of Process variables (Campion et al., 1996) can influence team performance. The first is a systems/network variable, transactive memory. Important inputs that influence the development of transactive memory are shown. Turnover in team training and/or team composition should negatively influence the development of transactive memory. Additionally, time in an organization should enhance the development of transactive memory systems. Transactive memory in turn should positively influence team performance. The second Process variable is a motivational one, collective efficacy. Experience and ability are important inputs into collective efficacy, as well as team performance. Collective efficacy and transactive memory are both thought to positively influence team performance.

The specific background for the particular studies conducted in this current dissertation come from the work of Moreland et al. (1998). These scholars call for additional research on natural work groups to confirm laboratory transactive memory findings. They suggest that research on natural work groups might involve archival or field studies and could be experimental or correlational in form. Information on transactive memory systems could be obtained through self-report or observational data and those data could then be correlated with various measures of group performance. They go on to suggest that particularly fruitful work may be done in organizations that vary in group training techniques by examining what effects these differences have on transactive memory systems and group performance.

The present studies are patterned after many of Moreland et al.'s (1998) suggestions. The present studies seek to extend empirical transactive memory research into several unexplored areas. This exploration should in turn either support or question the theoretical work to date. Thus far, theoretical work has exceeded empirical work. Furthermore, the influence of transactive memory will be compared to the more established construct of collective efficacy in an attempt to set up competing models to explain team performance. The model in Figure 1 suggests that both transactive memory and collective efficacy will influence performance separately. Results will shed light on which of the two constructs is the most powerful predictor. In order to extend Moreland et al.'s (1998) empirical work, the following hypothesis will be tested.

Hypothesis 1: Teams with higher levels of transactive memory will outperform teams with lower levels of transactive memory.

A number of team characteristics have not been examined in relation to transactive memory. These characteristics were touched upon in the beginning of this chapter. One issue to be addressed is the degree to which transactive memory systems have an effect on performance when roles are more clearly defined. In all of the studies using the AM radio assembly task, roles were left undefined by the experimenters. Each group defined its own roles during assembly. Only one transactive memory study has attempted to address anything related to team roles. Wegner et al. (1991) imposed a structure (by assigning expertise) as one of their experimental conditions. In this case, the imposed structure harmed performance in natural couples who already had an established expertise structure in place. The imposed structure improved performance in impromptu couples, but not significantly. Imposing structure on the team (by assigning expertise) could be viewed as defined roles. It would seem that well defined roles would lessen the need for a mature transactive memory system. If a transactive memory system is "who knows what, who is good at what, and who does what" then roles answer these questions to a degree. For example in a police SWAT team, the sniper is the one who knows about long range shooting, he is the one who is good at long range shooting, and he does the long range shooting. However, even with well defined positions, certain tasks may not be "position specific". Returning to the SWAT team, one of several officers may be the expert in weapon maintenance.

Do transactive memory systems operate in teams of greater than three individuals? It seems clear that teams of two and three do develop such systems. It would seem logical that with increased group membership size, the importance of transactive memory would increase as well. As more members are added it may be harder to determine "who knows what, who does what, and who is good at what". On the other hand, there is much more information available to the group by the sheer presence of more disparate minds. It would be helpful to identify expertise so that information could be smoothly funneled to the expert. The greater numbers also bring a greater challenge to group coordination and cohesion. Well developed transactive memory systems should improve these areas. It would also seem to take greater time and effort to build such systems with larger membership. Finally, task credibility (i.e. recognizing expertise and not challenging others) may be harder to obtain with larger teams. Therefore, the following hypothesis will be tested.

Hypothesis 2: In teams of greater than three individuals, transactive memory will affect performance in a positive manner.

Do transactive memory systems operate outside the laboratory, for instance in the field or in simulations? It would seem that they do. Wegner's (1987, 1991) natural couples developed their transactive memory system outside of the laboratory and it was simply tested inside the lab. However, Moreland et al.'s (1998) radio groups were all undergraduates and the groups were nominal. Tests of actual teams in the field are needed.

Hypothesis 3: Transactive memory effects found in laboratory groups, will be present outside the laboratory.

A number of variables should influence the development of transactive memory in the team (see Figure 1). Will consistent turnover or the reconstituting of teams cripple transactive memory and in turn lower performance? It seems likely, especially if transactive memory systems take time to develop as Wegner (1987) hypothesized. Moreland et al. (1996) demonstrated that teams that experienced turnover after training did not perform as well as teams that were trained and tested together. It also is conceivable that time in an organization may be related to the development of transactive memory. The longer an individual has been in an organization, the greater the opportunities to know other's strengths and weaknesses. This greater knowledge of other organizational members should translate into higher transactive memory levels among the teams, even if these teams are consistently reconstituted.

Hypothesis 4: Turnover, via the reconstitution of teams, will negatively affect transactive memory and performance in turn.

Hypothesis 5: Teams trained together will foster development of transactive memory systems which will result in increased performance over others trained in separate teams.

Hypothesis 6: Interaction over time in an organization should increase the level of transactive memory among teams.

Finally, collective efficacy has been shown to be a predictor of piloting performance in the laboratory (Lindsley et al. 1994). By extension it seems that collective efficacy should also predict flight performance in the air (see Fig. 1). Mischel and Northcraft (1997) have hypothesized that a portion of an individual's collective efficacy estimate is a determination of the KSAs of the group members in both task knowledge and interpersonal group skills. This is related to the idea of transactive memory (who knows what, who does what, who is good at what). However, as seen earlier, transactive memory consists of more than knowledge of interpersonal and task skills. As pictured in Figure 1, transactive memory and collective efficacy are both hypothesized to influence performance. Based on the previous discussion, the following hypotheses will be tested:

Hypothesis 7: Collective efficacy will be a predictor of flight crew performance.
Hypothesis 8: Transactive memory will explain variance in flight crew performance, above and beyond that of collective efficacy.

Chapter 4

Study 1: Flight Simulation

Methods

Overview

Study 1 was a flight simulation using a pilot and advisor team. The pilot flew a PC based simulator mission with the aid of the advisor in an effort to engage and destroy enemy aircraft. Transactive memory was examined through an experimental design that manipulated the composition of the pilot/advisor teams. More specifically, in one condition the pilots had the same advisor during their second set of trials while in the other condition the advisor for the second set of trials was new. This study tests Hypothesis 1, that teams with higher levels of transactive memory will outperform teams with lower levels of transactive memory.

Subjects

Subjects for this study were eight current or former military fighter pilots or Weapons Systems Operators from the local area. Each subject had been a rated military fighter aviator with at least 1,250 hours of military jet fighter time. An earlier pilot study indicated that actual military pilots were required to realistically complete the task. Due to the rigorous participation criteria, only eight aviators could be recruited. Subjects participated on a volunteer basis and all were males. In all, the subjects possessed over 18,550 hours in fighter aircraft. Seven of the eight had actual combat time in either Vietnam or the Persian Gulf. Table 1 summarizes the subjects' experience in particular aircraft. Each box indicates a type of aircraft, the number of individuals who flew that aircraft and the total hours in that aircraft. Several subjects had time in multiple aircraft, so the total does not sum to eight. The subject pool was very impressive and provides excellent external validity.

Table 1. Subject's experience in various fighter aircraft.

F-4 3 @ 3875	**F-15** 3 @ 2810	**F-5** 1 @ 200
A-4 2 @ 1325	**F-14** 2 @ 2300	**F-18** 1 @ 1000
F-105 1 @ 750	**F-111** 1 @ 1300	**A-10** 1 @ 1400

(number of subjects @ total hours)

Design

This study was a 2X2X2 mixed factor design. The first factor was Advisor (Same / Different) which was a between subject factor. The second factor was Week (One / Two) which was a within subject factor. The third factor was Technical (Tech) Condition (High / Low), also a within subject factor. The Tech condition was created for an alternative research project. Under the Tech condition, the advisor's access to the pilot was manipulated. In the High Tech condition, the advisor was seated next to the pilot with visual access to the pilot and the pilot's displays. In the Low Tech condition, a curtain separated the advisor and the pilot, restricting visual access between the two.

The primary factor of interest for this study was Advisor (Same / Different) during Week 2. It is in this condition that transactive memory was manipulated. During Week 2 the pilot either worked with the same advisor he had trained with during Week 1 (transactive memory condition) or a different advisor (no transactive memory condition). Therefore, only the results of the Week 2 Advisor manipulation will be reported with one exception. When there was a significant interaction with one of the other factors, then that interaction will be reported.

Pilots flew an F-22 PC based simulator on a simulated combat mission. They were assisted by one of two advisors who helped them navigate, evade enemy threats, and engage and destroy hostile aircraft. Pilots were trained by the advisor during the first week and then performed two actual trials. One of the trials was in the High Tech condition where the advisor had greater visual access to the pilot and the pilot's information and view. The other trial was in the Low Tech condition where the advisor's view of the pilot was restricted, as was information about the pilot's situation. The pilots returned one week later. They worked with either the Same Advisor as Week 1 or a Different Advisor. They received a short refresher session on the simulation and then performed two criteria trials, High and Low Tech. Tech condition was counterbalanced.

Procedure and Task

Advisors were trained over several weeks on the characteristics of the PC based simulator, the simulation, and the specific mission to be accomplished. Training ensured that both advisors were equal in ability.

When pilots arrived they were greeted by the experimenter. They received a short introduction to the task and then were introduced to the advisor. The advisor then took them to the simulator and conducted a training session with the pilot using a three page checklist (see Appendix A). The training session covered the mechanics of the control inputs, the symbology of the simulator displays, and the handling characteristics of the simulator. The pilot was then instructed on weapons use, evading threats, employing wingmen, and finding targets. The training session concluded with a criterion trial. In the criterion trial the pilot had to successfully down three enemy aircraft without being shot-down. Once criterion was reached, the actual trials began. Introduction, training, and criterion trials required approximately 45 to 50 minutes to complete.

The actual mission consisted of a simulated Combat Air Patrol (CAP) mission over Bosnia. These were routine missions where friendly fighter aircraft orbit and wait for enemy aircraft to come across a designated line. Once enemy aircraft were detected they were to be engaged and destroyed. The simulation began with the pilot in the air over Italy. The simulation was then "fast forwarded" as the pilot was taken across the Adriatic Sea to Bosnia. Just before leaving the Adriatic Sea airspace, the simulator was frozen and the actual trial began. The advisor aided the pilot in navigating the aircraft, identify enemy targets, engaging the enemy, avoiding air to air and ground threats, and employing weapons. The advisor's role was a cross between the role of an onboard Weapons System Operator in older fighter aircraft and an Airborne Warning and

Command System (AWACS) advisor who is responsible for directing friendly air forces in a coordinated manner against threats. The Air Force is currently investigating ways to make the AWACS advisor more helpful and this study was conducted in conjunction with that investigation.

The pilot and advisor had access to similar information, however, only the advisor could view situations more than 25 miles from the pilot's aircraft. In the High Tech condition the advisor had a much better idea of what was actually occurring in the "cockpit" of the pilot.

Every three minutes an alarm sounded and the simulator was frozen. Displays were then covered and both pilot and advisor filled out questionnaires on situational awareness measures (see Appendix B) and their perception of the value of communication between themselves (Appendix C). The advisor, using information from the displays, immediately graded only answers to the situational awareness questions. Once those questions had been graded, the simulation was continued. Pilots were not advised of their performance on the situational awareness measures. Questionnaires were completed four times for both the High and Low Tech trials. The actual mission portion lasted for 12 minutes for both the High and Low Tech conditions unless the aircraft was shot down or crashed (at which time that trial terminated). At the end of the trial, several performance measures were recorded by the advisor on the front of the Situational Awareness questionnaires (see Appendix D). Sessions were videotaped for later analysis of behavioral transactive memory indices (see Appendix E).

The Week 2 trials were identical to Week 1 with few exceptions. The advisor either remained the same as the first week (Same) or was replaced by the second advisor (Different). Instead of an extensive training period, pilots were given a quick (10 minute) review session of the symbology, switchology, control inputs, utilization of wingmen, and a short practice dogfight. Performance measures, situational awareness measures, and videotape all remained the same.

Measures

Performance. The simulation computed several performance measures for each mission and printed it out following the mission. The result was a categorical summary (five categories from Poor to Outstanding) computed by the simulator using the other performance measures. Bomb Damage Assessment (BDA) was a measure of weapon employment (100% reflects total accuracy with all weapons). Effectiveness considered targets destroyed as well as friendly losses (0-100%). Mission duration was calculated from the time the simulator came off of freeze, until the final Situational Awareness buzzer sounded (a maximum of 12 minutes). Number of aircraft shotdown was self-explanatory. How Ended was a dichotomous variable, either flew away (2) or shot down/crashed (1).

Situational Awareness. Situational awareness (SA) of both the pilot and advisor was measured using the form shown in Appendix B. This form was based on the SAGAT (Situational Awareness Global Assessment Technique) used to assess military pilot's SA (Endsley, 1995). The items addressed location of wingmen, location of bogeys (i.e. enemy aircraft), aircraft information (e.g. aircraft attitude), and number, type, and result of weapons fired. All this information is important to combat success. Situational Awareness was a percentage score. A score of 1.0 indicated completely accurate SA answers.

Perception of Communication. The value of communication between pilot and advisor was assessed on this questionnaire (see Appendix C) developed at the Human Interface Technology (HIT) Laboratory at the University of Washington with the help of the Investigator.

Behavioral Analysis. Videotapes were made of both Week 1 and Week 2 sessions. A trained observer from the Speech Communications Department, who was blind to the experiment, later analyzed the videotapes for the three behavioral indices of transactive memory (memory differentiation, task coordination, and task credibility). Additionally, the Investigator also analyzed the tapes using the same rating forms (see Appendix E). Ratings between observers were compared and revealed a high correlation ($r = .972$). Therefore, both observer's ratings were combined for statistical analyses. Only ratings of Week 2 were used in the analyses. Videotape rating training to ensure consistency between raters, was conducted with Week 1 tapes. Week 2 tapes were the actual tapes coded for transactive memory.

The rating form (see Appendix E) was an expanded version of the form used by Liang et al. (1995) in their behavioral coding. The rating form consisted of nine scales designed to evaluate the three components of Transactive Memory as defined by Moreland and colleagues (1998). Eight of the scales were 7 point lickert scales with a neutral middle position. The final measure was a behavioral count of incidences of confusion between the pilot and the advisor. The three components of TM and the measures are listed in Figure 2.

Memory Differentiation	Task Coordination	Task Credibility
Remember Different Elements	Task Coordination	Level of Criticism
Responsibility for Different tasks	Confusion (count)	Level of Frustration
	Smoothness of Commun.	Accept Suggestions
	Level of Cooperation	

Figure 2. Chart of Behavioral Analysis Items for Transactive Memory Videotape Rating Form.

Results

Determination of Significance. The individual sample size (n=8) of Study 1 was small. Furthermore, when dealing with group level phenomenon it is difficult to reach the "classic" alpha significance level of .05. With these considerations in mind, an alpha significance level of .10 was chosen as a reasonable level (Cascio & Zedeck, 1983). All p values reported will be one-tailed unless noted otherwise.

Descriptive Statistics. A visual inspection of the data was conducted. Two of the performance dependent measures (Result and BDA) resulted in very high performance with little variance. This is likely due to the high abilities of the subjects. No other departures from normality were noted.

Equivalence of Advisors. Advisors worked closely together during training in order to ensure equivalent ability and techniques. Analyses of the first week's performance data indicated no difference between advisors' performance with pilots.

Hypothesis 1. Hypothesis 1 asserts that higher performance will be attained by pilots in the transactive memory condition (same advisor both weeks). Study 1 endeavored to measure performance through a variety of means; each of these measures will be covered in turn.

Pilot Performance. It was hypothesized that pilots working with the same advisor in both Weeks 1 and 2 (Same Ad, T.M. condition) would show higher performance than those changing advisors on Week 2 (Different Ad, no T.M. condition). Only Week 2 is of interest for this transactive memory analysis as that is where the advisor change occurred (see Table 2). A MANOVA was performed on five of the dependent variables (the dichotomous variable was excluded to satisfy MANOVA requirements). The MANOVA for the performance data indicates that the advisor condition does not significantly effect the performance of the pilot in the second week (Wilks' lamba p = .289). The two groups do not differ across the vector of DVs.

Table 2. Performance Data by Transactive memory Condition

	Same Advisor (T.M.)	**Different Advisor (no T.M.)**
Duration of Flight	11.00	9.64
# Aircraft shot down	2.75	3.38
Bomb Damage Assess.	90.36	83.93
Mission Effectiveness	90.00	85.88
Downed 1 or Flew Away 2	1.75	1.62
Result (6=outstanding)	4.88	4.13

The majority of the dependent measures were in the hypothesized direction (five of the six). However, none of the dependent measures reached statistical significance. This is likely attributable to the small sample size (n = 8).

Situational Awareness. The transactive memory analyses revealed that when the advisor was changed for Week 2, that situational awareness on the part of the pilot was decreased (see Table 3). This is likely due to the fact that the advisor was not familiar with what type of information that the particular pilot required to build high SA. The Same Advisor pilots had higher SA than the Different Advisor pilots (p = .028). Furthermore, the Same Advisor Condition in period 3 also resulted in higher SA (p = .037), however, measures from this period should be taken with caution. Debriefs with the pilots after the simulator mission revealed that they were periodically taken into scenarios under which they would not have entered under actual combat conditions. These "furballs" often occurred when several enemy aircraft engaged the pilot while he was near a surface to air missile (SAM) site. The pilots reported that they normally would have egressed (exited) the situation, but were not given this option in our simulation. These scenarios (or "furballs") occurred only during periods 2 and 3 when the fighting was at its most intense. Furthermore, the values for periods 3 and 4 suffer from a smaller sample size as there are no values for pilots shot down in previous periods.

Table 3. Situational Awareness by Transactive memory Condition for Pilots

	Same Advisor	Different Advisor
Overall SA	.75**	.62**
SA period 1	.74	.73
SA period 2	.62	.40
SA period 3	.85**	.61**
SA period 4	.81	.60

**p < .05

The situational awareness of the advisor was also examined. Just as from the Pilot's perspective, we see that a change in advisor – pilot pairings also adversely affected the advisor's SA (see Table 4). When the advisor – pilot paring was changed, the advisor's SA decreased ($p = .039$). There are similar results for periods 2 ($p = .096$) and period 4 ($p = .073$). It is suspected that the pairs were not as cognizant of the information that the other member required to build high levels of SA.

Table 4. Situational Awareness by Transactive memory Condition for Advisors

	Same Advisor	Different Advisor
Average SA	.92**	.77**
SA period 1	.94	.89
SA period 2	.88*	.69*
SA period 3	.85	.72
SA period 4	.95*	.73*

**p < .05; *p < .10

Perceptual Data. Reliabilities for the perceptual scales were assessed. For the pilots' ratings $\alpha = .76$, and for the advisors' ratings $\alpha = .89$. The perceptual (or subjective) ratings made by both the pilot and advisor at the end of each flying period revealed few significant differences. A MANOVA was performed on the seven dependent variables for the pilot's ratings. The MANOVA for the perceptual data indicates that the advisor condition (same of different) does not significantly effect the perceptual ratings of the pilot in the second week (Wilks' lamba $p = .935$). The two groups do not differ across the vector of DVs. However when the ratings are examined individually there are indications that there may be some underlying differences. From the pilot's perspective, when the advisor changed, all subjective ratings dropped or stayed the same. Only the pilot's rating of the advisor's performance approached significance ($p = .116$). When the advisor changed, the pilot tended to rate the advisor's performance lower.

Things were a bit stronger from the advisor's perspective (see Table 5). All ratings made by the advisor were lower when the advisor was working with a new pilot (Different Ad) in Week 2. The advisor rated his own performance lower when working with a new pilot ($p=.053$) and the advisor rated the new pilot's performance lower as well ($p= .092$). Interestingly, the advisor's confidence in the correctness of his ratings of the pilot's SA drops when working with a new pilot ($p= .074$)

Table 5. Perceptual Data from the Advisor's perspective

Measure	Same Advisor (T.M.)	Different Advisor (no T.M.)
Amt of Communication	.75	.72
Value of Communication	.81	.80
Instigator of Communication	.86	.85
Advisor's Performance	*.87	*.81
Pilot's Performance	*.89	*.82
Other's SA	.85	.79
Confidence in Rating	*.81	*.65

*p < .10

Behavioral Analysis. Reliabilities for the sub-scales were assessed. For the Memory Differentiation items $\alpha = .11$, for the Task Coordination items $\alpha = .63$, and for the Task Credibility items $\alpha = .51$. The Memory Differentiation component is particularly low. This may be due to hierarchical roles as discussed later. The Memory Differentiation Component of transactive memory revealed no significant differences. This was surprising as the series of studies by Moreland and colleagues (1998) consistently showed a difference between teams that stayed together versus teams that were reconstituted.

Analyses of the Task Coordination Component of transactive memory revealed some surprising trends. Generally, the Different Advisor Condition resulted in smoother task coordination, less confusion, and more effective sequencing of communication. In a nutshell, it seems that the Different Advisor Condition resulted in better communication between pilot and advisor.

Specifics of the Task Coordination Component are below (see Table 6). There was a significant interaction between Tech and Advisor in relationship to Confusion. This interaction is seen in Figure 3. High Tech seems to mitigate increased confusion in the Different Advisor condition, but not in the Same Advisor Condition. Perhaps there is an increased vigilance when working in the Different Advisor condition which would explain this interaction (see Fig. 3). Interaction aside, Same Advisor showed significantly more incidences of confusion than did the Different Advisor Condition. Different Advisor also showed higher task coordination and smoother sequencing of communication.

Table 6. Task Coordination Component by Transactive memory Condition

	Same Advisor	**Different Advisor**
Amt of Confusion	4.50***	1.50***
Task Coordination	6.25*	6.69*
Sequence Communication	5.56***	6.31***
Cooperation	6.50	6.50

***$p < .01$; *$p < .10$

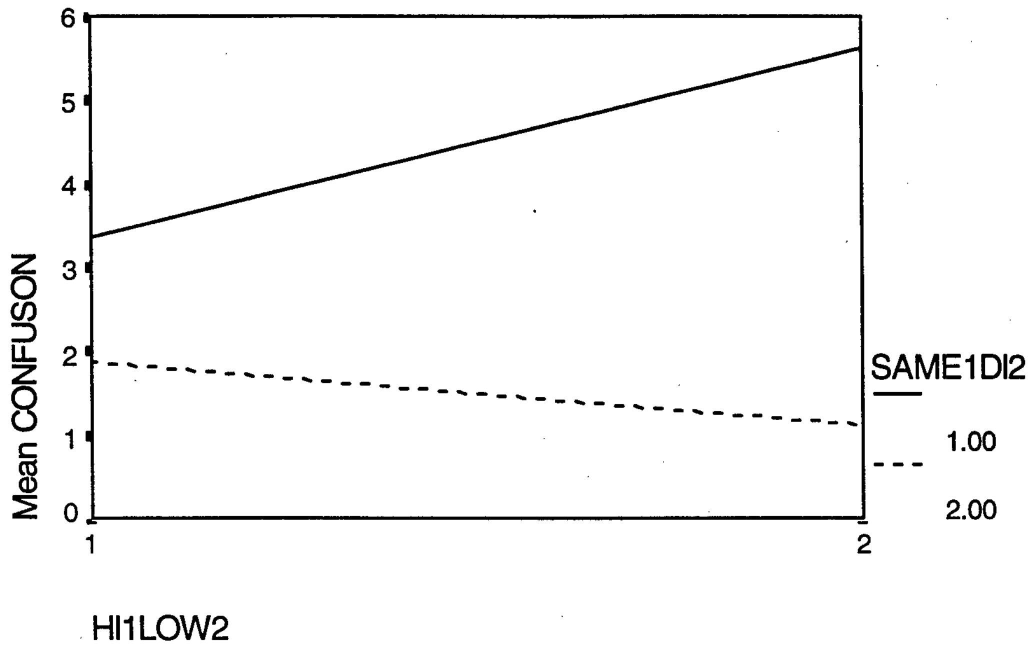

Figure 3. Interaction of Tech and Advisor Condition for the Confusion Variable

There were no specific trends noted in the Task Credibility Component of transactive memory. It is likely that the computer simulation of combat was unable to induce any differences in Frustration, Criticism, or Accepting Suggestions that could result in an actual flight under hostile flying conditions.

In summary of the Behavioral Analysis, the assignment of roles may affect the Memory Differentiation Component of transactive memory. The increased arousal of working with a new Advisor/Pilot may have led to increased emphasis on communication between the two. This communication was not able to overcome the advantages of working in the same teams that was evident under the Performance and Situational Awareness measures. Finally, it is unlikely that a simulation can induce increased levels of criticism, frustration, and resistance to suggestions (especially with clearly delineated roles) due to the more powerful effects of such things as Social Desirability (e.g. pilot and advisor maintaining a cordial relationship).

Taken as a whole these results provide only partial support for Hypotheses 1.

Discussion of Study 1

It seems that when Advisor / Pilot teams are kept in tact (the transactive memory condition), performance measures have a tendency to be higher. Larger sample sizes are required to confirm this trend. Both pilots and advisors showed higher SA when they remained in their set team. Interestingly, when the Advisor / Pilot team is changed, the advisor rates both the pilot's and the advisor's (his own) performance as lower. Perhaps this is a reflection of a perceived increase in difficulty in working with a new pilot on the part of the advisor. Finally, video behavioral analyses indicate that memory differentiation and responsibility may become overlapped in a situation where team members are assigned roles that encompass a teacher/pupil relationship. It also seems that new teams put more effort into smoothly communicating and coordinating with one another in an effort to effectively work with the new team member.

Surprisingly, the Memory Differentiation component of transactive memory revealed no significant differences in this study. One explanation for the lack of significance is the fact that team members were assigned clear roles in this study. In Moreland's studies team members were never assigned roles, rather roles emerged as persons with specialized knowledge migrated towards certain positions. Furthermore, in this study there was almost a hierarchical relationship between the two positions. The advisor not only assisted the pilot, but also trained the pilot in the use of the simulator (e.g. use of controls, symbology, etc.). These roles may have dictated that both pilot and advisor would maintain similar types of information. It also may have led to an overlap in responsibility for the tasks dictated by the scenario. This idea that assigned roles may be a boundary condition for the Memory Differentiation Component requires further research.

The Different Advisor Condition (the no transactive memory condition) produced significantly higher results for the Task Coordination Component of Transactive memory. These results were not expected, rather the Same Advisor Condition was hypothesized to result in better task coordination. It seems likely that a change in advisors resulted in greater arousal on the part of the pilot and the advisor. When the advisor changed it may have caused both the pilot and the advisor to put more effort into their communication since they were unfamiliar with one another. Conversely, perhaps the same advisor and pilot teams took more of a "laissez faire" attitude ("ok, same ole, same ole, same mission, same advisor") as little had changed from the first week's mission. In the Same Advisor condition, advisors and pilots were familiar with one another and perhaps not as vigilant in their communication patterns. This would explain less effective levels of communication, but yet an increase in performance and SA measures.

The greatest strength and weakness of this study are both related to the sample. The sample size was extremely small. This made it difficult to find any significant differences in the results. However, there were some significant differences in performance indicating the robust value of transactive memory in teams. Taken as a whole, several trends were clear even with the small sample size.

The reason for the small sample size was the extremely high criteria for participation in the study. The vast majority of studies to date have used college students as subjects. The use of combat jet pilots increases the generalizability of these findings outside of the lab into the cockpit. Subjects were professional and expert in the flying task. Earlier pilot studies of this research indicated that novices were not able to accomplish the task of combating enemy aircraft due to their inability to fly the simulator, operate the weaponry, and performing the combat maneuvers required to defeat the enemy. Anecdotally, several of the participants commented on the high fidelity and realism of the simulator.

Study 1 was one of the most realistic lab tests of transactive memory to date. However, field testing is required in transactive memory research. No field tests of transactive memory have appeared in the literature. Furthermore, Wegner et al. (1991) employed natural teams in their sample and Hollingshead (1998) recently surveyed employees at the University of Illinois on their natural office staff teams. However, in the lab portion of her study, Hollingshead used nominal teams. In short, there have been few tests of transactive memory using natural teams and no tests in field settings. Study 2 and 3 provide field tests of transactive memory and collective efficacy on actual Air

Force aircrews in the KC-135. These are important tests examining the question: is transactive memory a construct strictly for the lab, or is it important in the real world?

Chapter 5

Study 2: Fairchild AFB

Methods

Overview

Study 2 was a field study using Air Force flight teams at Fairchild AFB. Crew levels of transactive memory and collective efficacy were collected prior to an operational aircraft check flight. The flight was graded by independent evaluators to provide performance measures.

Subjects and Setting

Subjects were actual Air Force KC-135 aircrews stationed at Fairchild AFB in Spokane, Washington. Fairchild is the largest KC-135 base in the world. The Boeing KC-135 Stratotanker is used for aerial refueling and to transport cargo. The aircraft is based on the Boeing 707 airframe. There are four crewmembers on board the aircraft. The Aircraft Commander (AC) is the senior pilot on board and is the final authority for aircrew decisions. The AC generally has at least three years of experience in the aircraft and has flown for a minimum of four years. The Copilot (Co) is the second pilot on board. The Copilot assists the AC in piloting the aircraft. He or she has access to identical controls and nearly all the same switches and instruments available to the AC. The experience range of a Copilot varies between several months to three years. The Navigator (Nav) is seated behind the pilots and is responsible for navigating the aircraft using both electronic and celestial equipment. The navigator also assumes some communication duties and generally is responsible for any classified material aboard the aircraft. The experience range of a navigator can vary from a couple of months to over ten years. The KC-135 is beginning to undergo modifications in order to replace the Navigator with updated electronic equipment to be managed by the pilots.

The Boom Operator (Boom) is the only non-commissioned officer aboard the aircraft. The Boom is responsible for operating the boom apparatus in order to refuel other aircraft from the KC-135. The Boom is also responsible for handling the cargo and dealing with any passengers on board. Boom Operators vary in experience from several months to over twenty years. In summary, the KC-135 aircrew is composed of an Aircraft Commander, Copilot, Navigator, and Boom Operator.

A total of 87 crewmembers participated. However, 12 of those individuals were involved with flights with no checkrides (and thus no performance scores) and were therefore dropped from the sample. Another two individuals responded to the survey, but the other two crewmembers on board did not and therefore, the crew surveys were unusable. This resulted in a sample of 73 (of a possible 76) individuals and a total of 19 aircrews. Of the three missing surveys, one individual failed to respond to the survey. One individual (from a Day 1 flight; crew surveyed on Day 2, see below) was dispatched from the base before they could be surveyed and one individual was assigned to the evaluation team, but acting as a crewmember on a flight, and therefore was asked not to fill out a survey. Twenty crews were evaluated during the inspection. Usable surveys were returned by nineteen of the crews for a response rate of 95%. Subjects served on a volunteer basis. There were 66 males and 7 females. Experience in the KC-135 ranged from several months to over 20 years.

Procedure and Task

Air Force crews are subject to a major inspection (known as ASEV, or Aircrew Standardization Evaluation Visit) once every two years. During this visit, 25% of the crews on base are planned to be evaluated by Headquarters personnel for flying efficiency and effectiveness during an operational checkride flight. A checkride is a flight where one or more of the crewmembers is evaluated by an experienced check crewmember in order to determine the readiness and proficiency of the evaluatee and crew.

During mission planning the day prior, or the morning of the checkride, the experimenter administered crews a questionnaire to assess transactive memory and collective efficacy. (see Appendix F). Each crew received a survey for each of the four crewmembers. Prior to the flight, the crews dropped completed surveys into a designated box in the mission planning room of Base Operations. At the end of the day the experimenter retrieved the surveys which were sealed in an envelope. Two of the crews, which were evaluated during day one of the evaluation, received their surveys after the flight. It was not possible to administer their surveys prior to the flight.

After the checkride was complete, the evaluator completed a questionnaire concerning the aircrew's performance on the checkride. This questionnaire focused on overall aircrew performance (see Appendix G). Both questionnaires were accompanied by a cover letter (see Appendix H).

Measures

Collective Efficacy. Collective efficacy was measured using questions 44, 47, 49, and 50 of the aircrew questionnaire (see Appendix F). Question 44 was adapted from Peterson et al.'s (1996) measure of team efficacy. Following the suggestion of Mischel and Northcraft (1997), the efficacy questions refer to both task skills and knowledge. Items 47 and 49 were used by Lindsley et al. (1994) in their study of collective efficacy in flight simulation tasks. They point out that using confidence measures at different levels of performance has been shown to show greater variation than use of yes/no questions for each performance level (c.f. Gist and Mitchell, 1992). Therefore, confidence at each level was used for items 47 and 49. Item 50, the final collective efficacy item, was a measure developed by the author. There is a corresponding item in the Evaluator's questionnaire (see Appendix G). The collective efficacy measure was calculated by averaging across the four items. Since the items employed different scales, item scores were converted to standardized z-scores. The z-scores were then averaged for the four items across individuals and then teams to arrive at the collective efficacy score for each crew.

Transactive memory. Following the recommendation of Moreland et al. (1998), transactive memory was evaluated through three direct indices: the *agreement* within a group about who knows what, who does what, and who is good at what; the *accuracy* of those beliefs, and the complexity (or *strength*) of those beliefs.

Agreement

Crew responsibilities (items 11-18 of Appendix F) were developed by the author in coordination with experienced KC-135 crewmembers. These items are flight tasks with ill-defined responsibilities. In other words, there is no standard procedure for who must accomplish these tasks. Therefore, they give a good indication of how well the crew knows "who does what". Each item was scored by comparing answers across the four crewmembers. If all four crewmembers agreed on that answer, the score was 100% for that item. If only two members agreed the score was 50% and so on. If all four of the crewmembers disagreed, the score was 0% for that item. The scores were then averaged across the items. Similarly, the Crew Skills section (items 19-22) gives a good indication of "who is good at what" and "who knows what". Level of agreement was calculated as was done for the Crew Responsibilities section. Finally, items 26-29 of the Crew Strengths section indicate "who is good at what". Level of agreement was calculated as before. The final level of agreement score was derived by using a weighted mean of all three agreement sections.

Accuracy

Accuracy of beliefs was measured using the Crew Strengths section (items 26-29 of Appendix F). Crew members were asked to rate each other's greatest strength in a forced choice method from among three major flight areas: crew coordination, technical proficiency, and systems / emergency knowledge options. Each crewmember also rated their own greatest strength (item 29, Appendix F) and the evaluator rated the crewmember's greatest strength (item 5 Individual Evaluation, Appendix G). An accuracy score was determined using both the self-reported strength and the evaluator rated strength. For example, the copilot self reports that crew coordination is his greatest strength. Two of the other three crewmembers also report that crew coordination is the copilot's greatest strength. The self-report accuracy is then 67%. Meanwhile the evaluator rates technical proficiency as the copilot's greatest strength as does one of the three crewmembers. The evaluator accuracy rating is then 33%. The overall accuracy score for the copilot is then the average of the two, in this case 50%. The accuracy rating would be determined in like manner for the other three crewmembers. The accuracy scores for the four crewmembers would then be averaged to produce one accuracy score for the crew.

Strengths

Moreland et al. (1998) suggested using complexity of group members' beliefs about one another's skills as a direct measure of knowledge in transactive memory. Their measure of complexity consisted in how well subjects could describe each other's strengths and weaknesses in knowledge and technical proficiency. A more direct approach to this issue is a self-report measure. The Crew Strengths section (items 23-25 of Appendix F) assessed strength of belief level for each crewmember. A five-point scale was used from: 1-"very limited knowledge" to 5- "as well as you can know them". Strength of beliefs was simply the mean of each crewmember's ratings. Crews with high transactive memory should report higher knowledge of crew strengths.

The overall transactive memory score was a mean of the agreement, accuracy, and strength scores. Analyses were also conducted using each of the three components to investigate which component seemed to be the most important.

Crew Hardness. A measure of "crew hardness" was used to measure a construct that captures how often crewmembers fly as an integral crew. Crew hardness was determined on a percentage basis over the 90-day period prior to the evaluation. Crew hardness was determined by dividing the number of times individual crewmembers flew together over the total number of times flown over a three month period. These figures were obtained from the Flight Records Division of the flying unit. For example, during the ASEV checkride the AC flew with Co B, Nav C, and Boom D. In the previous 90 days the AC once flew with just Co B (.50), once with Co B and Nav C (.75), and once with B, C, and D (1.0). The AC also flew with other crewmembers on one flight (.25). The AC's overall hardness figure would be .50. Meanwhile the Copilot had five flights and his hardness score was .40 and the Nav had 3 flights and a hardness score of .60. Finally the Boom had 6 flights and a hardness score of .70. The crew hardness score would be

$$[(.50) + (.40) + (.60) + (.70)] \;/\; 4 = .55$$

Crew hardness was also calculated by weighting the individual hardness score by the number of flights over the last 90 days. Using the numbers above the crew hardness score would be as follows:

$[4(.50) + 5(.40) + 3(.60) + 6(.70)] / 4 = .56$

Similar results were obtained using both procedures. In addition, number of flights over the last 90 days was a separate control variable. For these reasons it was determined to use only the first method of crew hardness calculation.

As mentioned previously, two crews completed their surveys after their checkride flight. The crew hardness scores for these two crews were adjusted to include the checkride flight in the crew hardness scores. The other evaluated crews' hardness score did not include the checkride flight, but only flights in the previous 90 days. Analyses revealed that adjusting the crew hardness scores in this manner did not alter overall results.

Control Variables. A number of control variables were included to account for individual differences in ability and experience. The first control variable was total flight hours, which is an indication of experience and, to a lesser extent, ability. Total flight hours was a self-report measure and the flight hours were summed for the entire crew.

Instructor status was also a control variable. After a period of roughly three years a person is selected for instructor school after demonstrating sound ability in their position. Instructors receive special training on the aircraft and aircraft procedures and in turn train younger crewmembers. Thus, Instructor status reflects both experience and ability and the number of instructors aboard the aircraft was calculated for each crew. Instructor status was determined via the Flight Records Section of the flying unit.

Duration of mission qualification, is the time that a crewmember has been mission qualified (i.e. declared competent) in that crew position. This figure can range from several months to over twenty years. This figure was obtained from Flights Records and totaled for each crew providing another indication of crew experience.

Distinguished graduate status indicates those crewmembers who were distinguished graduates in a flying or mission related training program. This is a measure of ability and a self-report item. This variable was dummy coded for the crew, with a 1 indicating that one or more crewmembers were distinguished graduates.

Flights in last 90 days was the last control variable used. Number of flights in the last ninety days is a common check the Air Force conducts to determine the "currency" of air crewmembers. Currency helps to ensure that crewmembers remain proficient in the aircraft. It is assumed that the more flights a crewmember has, the more proficient they will become (up to a point). There are a minimum and maximum number of flights allowed every ninety days. Number of flights can also be a proxy for experience or ability. The instructors in a flying unit (those more experienced and with higher ability) generally have additional duties which preclude them from flying a great deal. In other words, the younger crewmembers do most of the flying, while the experienced crewmembers supervise, administrate, and fly a little bit as well. In addition to flying less, the instructors also tend to fly with a variety of different crewmembers when they do fly. The number of flights in the last 90 days was obtained from the Flight Records Division of the flying unit.

Criterion Variables. There were two criterion variables in this study. The first was Critiques per Crew. For each checkride, individuals were rated on multiple items (e.g. mission planning, weather avoidance, takeoff, situational awareness, etc.). On each item an individual can receive a critique for not accomplishing the item in an error free or proper manner. The Crew Critiques per Crew score was determined by adding up the total number of critiques that were related to crew effectiveness. For example, if the copilot received a critique for landings this would not be related to crew effectiveness. A poor landing is an individual deficiency. However, a critique for forgetting to lower the gear on final approach would be a crew deficiency. The crewmember flying the aircraft is primarily responsible for having the gear lowered, however each crewmember is responsible for ensuring the gear is lowered on final approach to landing.

Evaluators annotate the number of critiques onto a standard Air Force checkride form. Evaluators were asked to transfer the number of critiques related to crew coordination onto the Evaluator's survey (see Appendix G).

The number of crew critiques for each individual crewmember was totaled and divided by the number of crewmembers checked, to arrive at the Critiques per Crew score. After all Critiques per Crew scores were computed, the variable was then reverse scored so that a higher total indicates a better Critique score for the crew. This could be considered the more objective of the two criterion variables since it was tied to specific performance parameters.

The second Criterion variable was Overall Crew Evaluation (item 3 of Appendix G). Each evaluator on the flight was asked to rate the overall crew performance. A mean score was taken from the individual evaluator ratings. This measure could be considered the more subjective of the two criterion variables.

Results

Determination of Significance. The individual sample size (n=76) of study 2 was favorable. However, because there were four members per crew, Study 2 had a lower crew sample size (n=19) than desired, despite enjoying an extremely high response rate (95%). Furthermore, when dealing with group level phenomenon, it is difficult to reach the "classic" alpha significance level of .05. Additionally, this was the first field study of its kind and therefore was somewhat exploratory in nature. With these considerations in mind, an alpha significance level of .10 was chosen and will be used for this study (Cascio & Zedeck, 1983). All p values reported will be one-tailed unless noted otherwise.

Descriptive Statistics. The variables of interest in this study were either measured at the crew level or measured at the individual level and then aggregated to the crew level. Therefore, checks for normality and other descriptive statistics were performed on the crew level variables. A visual inspection of the data revealed no serious departures from normality.

Control Variables. When it is stated that experience and ability were controlled, the five control variables listed in Table 7 are the reference. Table 7 depicts the correlations between the control variables and performance.

Table 7. Correlations between control and performance variables

	Critiques per Crew#	**Overall Crew Evaluation**
Total Hours	.16	.32*
Mission Qualified Duration	.17	.18
Distinguished Graduate (Dummy Variable)	-.20	-.07
Instructors Onboard	.32*	.34*
Flights (Last 90 days)	-.18	-.58**

Reverse scored so that more critiques per crew relates to higher performance
*$p < .10$, ** $p < .05$

Somewhat surprisingly, the distinguished graduate correlation was a negative one. However, the correlation was quite low and does not approach significance. Of greater concern, is the negative relationship between flights in the last ninety days and performance. Recall from the earlier discussion of control variables that number of flights can also be a proxy for experience or ability. Instructors (those more experienced and with higher ability) generally have additional duties which preclude them from flying a great deal. It is clear from Table 7 that a greater number of instructors on board were highly positively related to Overall Crew evaluation. Therefore, it is not surprising that the number of flights - performance relationship was a negative one. Stated another way, having lots of flights in the last ninety days was an indication of less experience and ability. Support for this claim can be seen in Table 8.

Table 8. Correlations between flights and experience indicators

	Flights (Last 90 days)
Total Hours	-.34*
Mission Qualified Duration	-.35*
Instructors Onboard	-.39*

*$p < .10$

<u>Transactive memory</u>. The transactive memory variable was computed using the means of the Agreement, Accuracy, and Strength components for each crew. Table 9 shows the correlations between the components and the transactive memory variable. Strength showed the highest correlation with transactive memory and Accuracy the lowest. While the three components showed good correlations with transactive memory, the relationships among the components themselves were interesting. None of the components was significantly related to another. Moreland and colleagues (1998) reported highly correlated components, but this may have been due to common method variance. Additionally, Strength and Accuracy exhibit a negative (but nonsignificant) relationship with one another (see Table 9).

Table 9. Correlations between transactive memory components

	Agreement	**Accuracy**	**Strength**	**T.M.**
Agreement				
Accuracy	.19			
Strength	.29	-.19		
T.M.	.58***	.55***	.68***	

***$p < .01$

Due to small crew sample size, factor analyses was not appropriate (Byrne, 1994). However, it seems from the relationships among the components that, as measured, the transactive memory variable may consist of more than one factor. A look at the descriptive statistics on Accuracy was revealing. Accuracy was very low among all the crews. The mean score on Accuracy was .321; the median was .333. The lowest crew Accuracy score possible was 0, indicating no accuracy. The highest is 1.0, indicating perfect accuracy among crewmembers. This mean and median indicate that the crews were very near the point of having only one out of four of the crewmembers being accurate on who had what strength. In other words, less than a third of the time were crewmembers accurate about crew strengths.

Theoretically, it seems clear that Accuracy is a legitimate component of Transactive memory. It may be that its measurement needs to be improved. Crewmembers were asked to delineate between three closely related strengths: crew coordination, technical ability, and systems knowledge. This discrimination task may have been too difficult. These results indicate that was indeed the case.

In light of the above discussion, the Accuracy component was eliminated from the analyses. Due to the elimination of the Accuracy component and because the remaining components were not significantly correlated, each component of transactive memory was used separately to identify its effects on performance. This use of individual components (i.e. no combining of components) is in accord with recommendations of previous research (Moreland et al., 1998). Analyses indicated that only the Strength component was of value as a separate and individual variable, while Agreement showed little relation to the criterion variables. Strength was the self-reported measure of how well crewmembers know one another's strengths. It is much easier to collect, measure, and judge than are the other two components of transactive memory.

Collective efficacy. Collective efficacy was measured from four items on the survey using three different scales. The scales were then combined and averaged through z scores. The scales showed good interater reliability between individuals; Cronbach's (1957) alpha was .68.

Criterion Variables. The two criterion variables, Critiques per Crew and Overall Crew Evaluation exhibited a highly significant positive relationship ($r = .63$, $p = .002$). Keep in mind, that Critiques per Crew were reverse scored so that higher critiques per crew indicated better crew performance.

Hypothesis 1,2,3. After the data were examined for normality and the two aggregate variables were analyzed, the hypotheses were tested in turn. The first three hypotheses, state that teams with higher levels of transactive memory will outperform teams with lower levels of transactive memory with greater than three team members and in the field. Therefore, crews with higher transactive memory scores should have higher performance scores. After controlling for experience and ability (hours, mission qualified duration, distinguished graduate status, instructors on board, and flights in the last 90 days), there were indications of a positive relationship between one transactive memory component and performance (see Table 10). Strength demonstrated a strong positive relationship with the criterion variables. The relationship failed to reach significance primarily due to the small sample size that resulted in only 12 degrees of freedom after controlling for the five control variables. Agreement surprisingly revealed a negative, but nonsignificant relationship with the criterion variables. This is likely due to Agreement's negative relationship with number of instructors on board (r = .49, p = .018). Recall from Table 7 that instructors was significantly related to the criterion variables. However, since instructors tend to fly with a variety of crews, they exhibited lower Agreement scores. These data provide partial support for Hypotheses 1, 2, and 3.

Table 10. Partial correlations between T.M. components and performance

	Critiques per Crew#	**Overall Crew Evaluation**
Agreement	-.10	-.23
Strength	.33	.21

Reverse scored so that more critiques per crew relates to higher performance

Hypothesis 4. According to Hypothesis 4, turnover via the reconstitution of teams, will negatively affect transactive memory and performance in turn. As previously shown, there is evidence that transactive memory does positively affect performance. To test whether turnover affects transactive memory, the crew hardness variable was used. Greater crew hardness indicates crews that fly as an integral unit more often. Alternatively, a low crew hardness score indicates crewmembers that frequently change crews for each flight. To support Hypothesis 4, crew hardness should be highly correlated with components of transactive memory, and it was (see Table 11). Strength exhibits this relationship. Agreement failed to reach statistical significance, due to small sample size. This is another indication that Strength may was the more valuable predictor variable. Therefore, Hypothesis 4 is strongly supported.

Table 11. Correlations between transactive memory components and crew hardness

	Crew Hardness
Agreement	.29
Strength	.57***

***$p < .01$

Hypothesis 5. This hypothesis states that teams trained together will foster development of transactive memory systems which will result in increased performance over others trained in separate teams. Since Fairchild is an operational base and not a training base, this hypothesis could not be tested.

Hypothesis 6. Being assigned to the same flying squadron allows crewmembers to interact in a flying environment without necessarily flying with one another on a frequent basis. Hypothesis 6 suggests that this interaction over time will increase the crew's transactive memory. To test this hypothesis, data was collected on how long individuals had been assigned to the same squadron. Crews composed of members of the same squadron (n=9) were examined to discover a possible correlation between time in the squadron and transactive memory components. I hypothesized this would be a positive relationship and it was for Strength (see Table 12). Again, Agreement was not a useful predictor.

Table 12. Correlations between squadron time and transactive memory components

	Time assigned to Squadron
Agreement	.03
Strength	.59**

**$p < .05$

Besides increased opportunity to interact and observe, being in the squadron together increases member's chances of flying together (see Table 13). There are four refueling squadrons located at Fairchild, but members of a squadron generally fly together. However, even after controlling for Crew Hardness, time in the squadron is still highly related to transactive memory (see Table 13). The relationship failed to reach statistical significance for the components of transactive memory, however this can be attributed to the small sample size (n=9) of crews from the same squadron. The correlations themselves are positive and strong (see Table 13). Overall the results provide some support Hypothesis 6, longer time assigned to the squadron is related to development of transactive memory.

Table 13. Correlations of squadron time with crew hardness and transactive memory

Controlling for:	**Relationship**
	Squadron Time - Flights r = .55*
Crew Hardness	Squadron Time – Agreement r = .27
Crew Hardness	Squadron Time – Strength r = .36

*$p < .10$

Hypothesis 7. The seventh hypothesis states that collective efficacy will be a predictor of flight crew performance. To test this hypothesis, collective efficacy was correlated with both of the performance criterion variables. After controlling for experience and ability collective efficacy demonstrates a strong relationship with performance (see Table 14).

Table 14. Partial Correlations between collective efficacy and performance

	Critiques per Crew#	**Overall Crew Evaluation**
Collective efficacy	.62***	.47**

Reverse scored so that more critiques per crew relates to higher performance
** $p < .05$, ***$p < .01$

Two multiple regressions were also conducted, regressing each performance variable onto the five control variables and collective efficacy. Both overall F-tests were significant (for Critiques per Crew: F=2.53, p<.10, 2-tailed; for Overall Crew Evaluation: F=2.44, p<.10, 2-tailed). The lower p values can be attributed to a large number of control variables in comparison to a smaller sample size. Of greater importance is the change in R squared of the regressions after controlling for experience and ability. The values are displayed in Table 15.

Table 15. Results of the analyses for performance regressed onto collective efficacy after controlling for experience and ability

	Critiques per Crew#	**Overall Crew Evaluation**
Collective efficacy	Change in Rsq = .28 Fchange = 7.64**	Change in Rsq = .12 Fchange = 3.33*

Reverse scored so that more critiques per crew relates to higher performance
*$p < .10$, ** $p < .05$

The table highlights the fact that crew collective efficacy is explaining a healthy portion of the variance in performance, particularly in Critiques per Crew, after controlling for experience and ability variables. These finding strongly support Hypothesis 7; collective efficacy predicts crew performance.

Hypothesis 8. Transactive memory will explain variance in flight crew performance, above and beyond that of collective efficacy. The preceding results clearly indicate that collective efficacy explains variance in performance. Previously, evidence was presented indicating that the Strength component of transactive memory was positively related to performance, but not at a statistically significant level. Partial correlations were conducted to examine if the transactive memory components explain variance above and beyond that of collective efficacy. When the effects of the five control variables and collective efficacy were partialed out, Strength showed a nearly zero relationship to performance (p>.10). Hypothesis 8 is not supported, transactive memory does not predict performance above and beyond that of collective efficacy.

Further Analyses After reviewing the results above, additional analyses were conducted to further examine the powerful variable, collective efficacy. Mischel and Northcraft (1997) suggest that greater experience and ability should lead to higher collective efficacy. Bivariate correlations between the five control factors and collective efficacy support this contention (see Table 16). Correlations are clearly positive, except for Flights in the last 90 days which is an indication of less experience as discussed earlier.

Table 16. Correlations between control variables and collective efficacy.

	Collective efficacy
Total Hours	.33*
Mission Qualified Duration	.27
Distinguished Graduate (Dummy Variable)	.22
Instructors Onboard	.43*
Flights (Last 90 days)	-.16

*$p < .10$

It seems likely that higher levels of transactive memory will lead to higher collective efficacy as discussed in Chapter 2 (Mischel & Northcraft, 1997; Gist & Mitchell, 1992; Bandura, 1982). Partialing out the effects of experience and ability, correlations were conducted between transactive memory components and collective efficacy. There is a clear and strong relationship between the two as exhibited in Table 17 for Strength. Agreement exhibited no relationship with collective efficacy, providing further evidence that Strength is the most important predictor variable.

Table 17. Correlations between transactive memory and collective efficacy controlling for experience and ability

	Collective efficacy
Agreement	.00
Strength	.64***

***$p < .01$

Is it possible to show that Crew Hardness is related to collective efficacy? Intuitively it seems reasonable that a crew that flies together often would develop higher collective efficacy. Further analyses were conducted to determine if crew hardness would directly predict collective efficacy. Crew hardness was positively correlated with collective efficacy, however, this relationship was mediated by Strength (see Table 18). This supports the contention that the mechanism that translates crew hardness into collective efficacy is transactive memory. Agreement did not mediate this relationship. This is not surprising given the fact that there was no relationship between Agreement and collective efficacy (see Table 17).

Table 18. Correlations between collective efficacy and crew hardness and transactive memory components

Controlling for: Five Control Variables	**Crew Hardness - Collective efficacy**
Alone	.46**
and Agreement	.49**
and Strength	-.01

** $p < .05$

Discussion of Study 2

In the second study, collective efficacy was the key variable in explaining aircrew performance. As hypothesized, collective efficacy was clearly related to both criterion variables. A component of transactive memory, Strength, was positively related to performance as well, but did not reach statistical significance. With increased sample size, transactive memory may indeed be predictive of performance. However, the results of Study 2 indicate that the transactive memory was strongly related to collective efficacy, which was not hypothesized, but is in accord with theory on collective efficacy (Mischel & Northcraft, 1997).

Of the transactive memory components, it is clear that Strength was the most potent predictor of both performance and collective efficacy. Moreland and colleagues (1998) reported highly correlated transactive memory components in their lab studies using radio assembly. Similar to this study, they had planned to examine the effect of each component separately. However, because the components were highly correlated a combined measure of transactive memory was used by Moreland. This high correlation may have been due to common method variance. What is clear from Study 2 is that

Strength was the "star performer". This is a positive finding. Strength is much easier to collect and calculate than are the other two components. It is simply a self-report measure of how well each crewmember knows the strengths of other crewmembers using a five-point lickert scale. Whereas Agreement requires a series of questions concerning "who does what and who knows what". The Agreement score must be calculated by comparing responses across crewmembers for each question which is time consuming. Furthermore, it requires the researcher to generate the series of questions for each type of task. Similarly, a researcher would be required to generate a list of items to be scored individually for the Accuracy component. As was seen in Study 2, finding the right series of questions may be difficult. Floor and ceiling effects must be avoided. Study 2 suffered from floor effects.

The major variable of interest, transactive memory, was found to be effected negatively by increased turnover, as hypothesized. Consistently substituting crewmembers onto crews adversely affects transactive memory. The more a crew flies together as a unit, the higher their transactive memory. Additionally, transactive memory clearly mediated a strong crew hardness - collective efficacy relationship. Transactive memory seems to be the mechanism through which crew hardness is translated into collective efficacy.

Finally, it seems that flying crewmembers from the same squadron as a crew is associated with higher levels of transactive memory. The evidence presented supports this contention. Because crewmembers from the same squadron have greater opportunity to fly together, this finding was not surprising. However, even after controlling for these increased flying opportunities; increased time in the squadron was associated with higher levels of transactive memory. This finding is not as intuitive. Overall, these results suggest the series of relationships exhibited in Figure 4.

Crew Hardness ↘ Transactive memory (Strength) → Experience & Ability ↓ ↘ Collective efficacy → Performance

Time in Squadron ↗

Figure 4. A revised model of the transactive memory / collective efficacy process.

Transactive memory is positioned in such a way that indicates that it is predictive of collective efficacy. This seems to be the logical sequence. Intuitively, it would seem that a crew that better knows who knows what, who does what, and who is good at what, would be a more confident crew. It is less likely that because a crew is more confident they would then acquire such a knowledge system. This model conflicts slightly with the model proposed by Peterson et al. (1996). They found that early collective efficacy predicted later shared mental models. There are key differences between Study 2 and the Peterson et al. (1996) study. First, the teams in Peterson et al. had little history as a team prior to the first measurement of collective efficacy and mental models. Whereas, in many cases the teams in Study 2 had a well established history before the study began. Second, shared mental models and transactive memory are similar, but not identical constructs. It seems reasonable that it would take time to establish shared mental models, and with no prior history it was unlikely that such models would emerge early in a team's tenure. Alternatively, there could be clues early in a team's development about who knows what, who does what, and who is good at what, and therefore transactive memory could perhaps be established more quickly than shared mental models.

It should be noted that these results do not infer that crews with lower collective efficacy perform poorly on flights. Nor do they infer that lower transactive memory leads to poor performance. Indeed the majority of the crews evaluated were satisfactory performers. Only one crew of nineteen received an unsatisfactory evaluation rating. Additionally, 14 of the 19 crews were rated as "average" or better by the evaluators. It seems that the majority of the crews were good performers.

So it seems that higher collective efficacy separates the "exceptional" crew from the "average" crew. Variables which improve collective efficacy are crew hardness (lack of turnover), transactive memory, and of course experience and ability. So it seems that whatever the squadron leadership can do to keep crews together, would be beneficial to their crew confidence and performance in turn. Experience and ability are also indicative of the "exceptional crew". Increasing the number of instructors on board increases both collective efficacy and performance. However, during this period when the Air Force is facing nearly a crisis situation in failing to retain highly experienced aircrewmembers (Wall Street Journal, 2 June 1999), squadron leadership has a dwindling capability to just "put more instructors on board". They must find other means to increase crew confidence. Flying crews together and thereby increasing transactive memory seems to be one way the leadership can influence confidence and performance.

The data suggest that instructors tend to fly together less often than the "line" flyers and this is as it should be. However, I propose that if instructors flew as hard crews more often, they would improve their performance scores to an even greater degree. It is also possible that higher transactive memory levels of less experienced crews may help to keep their performance from dipping to unsatisfactory levels.

This study also informs theory on the development of collective efficacy. Results seem clear that one way to increase the team's collective efficacy is through transactive memory. Developing the team member's ability to determine who knows what, who does what, and who is good at what should increase team confidence. Previous theoretical work has pointed to the important role of enactive mastery (Bandura, 1982) in developing collective efficacy. It seems that at least one vehicle used by enactive mastery to translate experience into efficacy is transactive memory. This has important implications for efficacy theory.

Theoretically, there seem to be two determinants of transactive memory. First, as Moreland and colleagues (1998) discovered, a team working on a task together over time increases transactive memory. This study supports this previous theoretical and empirical work by showing that greater crew hardness leads to higher transactive memory.

Second, it seems that a team may not be required to be engaged in their primary task to increase their transactive memory. It seems that other conditions besides flying together may help build transactive memory. One such variable is time together in an organization; in this case a flying squadron. People talk, you hear rumors, you see people interact during mission planning, in flying safety meetings, etc., and you get some idea of how competent they are around an airplane. You get some idea (by observing and / or listening) what their strengths and weaknesses are as well. Simply interacting and observing over time while engaged in auxiliary activities related to the primary task may be sufficient to increase the team's transactive memory. A practical application of this finding is to encourage organizational member interactions in discussions and or exercises related to the task or task components. This would be especially important if a team cannot be frequently engaged in the task due to expense, time, etc. Group techniques such as "pulling part success" via the sharing of "war stories" may be one vehicle to increase transactive memory. Another practical application is to encourage leadership to compose crews from the same squadron whenever possible.

The major shortcoming of this study was the small number of crews evaluated. The inspectors hoped to have a sample size of 25 crews. However, real world commitments such as the Kosovo Crisis, drastically reduced the number of crews available to be evaluated. This smaller sample size likely kept the transactive memory - performance relationship from reaching significance. However, with the negative also comes with a positive. Several strong relationships were found even with a small sample size. This is indicative of some robust relationships among the variables; in particular those related to collective efficacy and performance.

Study 2 provided a rigorous field test of transactive memory and collective efficacy in an operational Air Force environment. It is one of the few tests of such constructs outside of the laboratory. In an effort to extend the understanding of the role of transactive memory and collective efficacy in aircrew performance, a second study of Air Force KC-135 crews was conducted at Altus AFB, OK. The site provided an opportunity to replicate the results of study 2 on a similar sample. However, there were important differences between the two samples. Fairchild AFB is an operational base while Altus AFB is a training base for new KC-135 crewmembers. This results in two different atmospheres: a mission first versus a training first focus. Furthermore, there was a variation in experience, Fairchild providing the "average" crewmember, while Altus has the "new" KC-135 crewmember.

Chapter 6

Study 3: Altus AFB

Methods

Overview

Study 3 was similar to Study 2. Study 3 used Air Force flight teams at Altus AFB, Oklahoma. Altus AFB primarily differs from Fairchild AFB in that it is a training base, where KC-135 crews receive initial qualification training in the KC-135. Levels of transactive memory and collective efficacy were evaluated prior to an operational aircraft check flight. Independent evaluators provided performance measures for the flight.

Subjects and Setting

Subjects were actual Air Force KC-135 aircrews undergoing qualification training at Altus AFB in Altus, Oklahoma. Altus is the primary KC-135 training base for the United States Air Force. The aircraft and aircrew positions are similar to those described in Study 2. Crew positions are identical with one exception. Altus is beginning a transition to the KC-135 with a Pacer Crag modification. This modification includes advanced navigation equipment available to the pilots and this eliminates the need for a Navigator. Crews flying the Pacer Crag do not employ a Navigator. Six of the fifteen aircraft used in this sample were Pacer Crag modified.

A total of 44 crewmembers participated. However, one of those individuals was on an aircraft with only one trainee and was dropped from the sample. This resulted in a sample of 43 (of a possible 45) individuals and a total of 15 aircrews. Only trainees were surveyed on the checkflight; of the 15 crews only two trainees failed to respond. On several flights, permanent party instructors (those assigned to instruct at Altus, n=9) served as crewmembers for positions with no trainees and were not surveyed. Subjects served on a volunteer basis. No gender data were available.

Procedure and Task

At Altus, the AC, Co, Nav, and Boom are undergoing initial qualification in the aircraft. Therefore, each training flight is conducted under the supervision of an instructor in that respective position. The Instructor Pilot (IP) supervises both AC and Copilot training. The AC in training is generally upgrading from the copilot position, but is sometimes transitioning from another aircraft. Other than the AC in training, the crewmembers are new to the aircraft. At the end of the training program (after approximately eight training flights, depending on weather and training continuity) the trainees receive a checkride similar to the one described in Study 2 to determine their suitability to be declared proficient in the KC-135. During the checkride, the trainees do not fly with an instructor, but rather with an evaluator who is expert in that crew position.

During the training program, trainees were briefed on and encouraged to participate in the research process by squadron leadership. As part of their mission planning package prior to their checkride, aircrew members received a package of surveys similar to the one used at Fairchild AFB (see Appendix I). The surveys were filled out anonymously by the crewmembers, returned to the crew envelope, and sealed. The envelopes were then mailed to the investigator by the squadron administration section. The survey assessed transactive memory, collective efficacy, total flight hours, and whether the crewmember was a distinguished graduate from flight training. After the checkride was completed, the evaluator filled out a questionnaire regarding the aircrew's performance on the checkride. This questionnaire focused on transactive memory, and overall aircrew performance (see Appendix G). Both questionnaires were accompanied by a cover letter (see Appendix H). The same portions of the survey were analyzed in Studies 2 and 3 so that comparisons could be drawn between the samples.

Measures

Collective Efficacy. Collective efficacy was measured exactly as described in Study 2.

Transactive memory. Transactive memory was measured as described in Study 2. As will be highlighted in the Results section, Accuracy again demonstrated a low mean. Due to this fact, coupled with a desire to remain consistent across studies, analyses were only conducted using the two transactive memory components, Agreement and Strength.

Crew Training Hardness. Originally Study 3 was designed to have two training conditions comparing crews trained as integral crews versus those trained in a random crew pattern. Until recently the two squadrons at Altus practiced these two training methods. Recently, the second squadron began to shift to an integral or "hard crew" training philosophy because they felt that it provided better training for the crews. Due to these changes in crew scheduling, a different measure of "crew training hardness" was used to determine how transactive memory, performance, collective efficacy, and crew training hardness relate.

Crew training hardness is a construct that captures how often the trainees fly as an integral crew. Crew training hardness was calculated by summing the number of times during the program that the trainees flew with other crewmembers of the checkflight crew. This number was then divided by the total number of dyads possible. For example, see Figure 5 below:

AC-Co	7 flights	Co-Nav	0 flights
AC-Nav	0 flights	Co_Bo	4 flights
AC-Bo	4 flights	Nav-Boom	0 flights

The Crew Training Hardness score would be: [7 + 0 + 4 + 0 + 4 + 0] / 6 = 2.5

Figure 5. An illustration of how Crew Training Hardness totals were calculated.

Crew Training Hardness was also calculated between only trainees (ignoring dyads with permanent party instructors) on board the flight. Results of analyses using the Crew Training Hardness variable did not change. Therefore, crew training hardness as calculated in the example will be reported. Number of flights together was a self-report item.

Control Variables. Two control variables were included to account for individual differences in ability and experience. The first was total hours per trainee. This was a self-report measure. Hours for the trainees were added for the crew and then divided by the number of trainees on board. The total hours did not include those of the permanent party instructors flying as crewmembers on the flight. These data were excluded primarily so as not to skew the hours data for the trainees, who were the subjects of interest. Additionally, these data were not available and since the instructors were not being evaluated they were considered to be of no importance. As a footnote, all the instructors at Altus have a similar number of total flying hours.

The second control variable was number of instructors on board. In Study 2 this was an important variable. It should be noted that the number of instructors on board has a different meaning in Study 3. At Fairchild, an instructor was either being evaluated or vulnerable to be evaluated. This was not the case at Altus. Permanent party instructors were not evaluated on the trainees' checkrides. Rather, they simply acted as competent crewmembers filling a position on the aircraft and were treated as peers by the trainees. Therefore, each instructor on board represented "one less person to worry about" for the trainees. While they still interacted as fellow crewmembers, the trainees knew that the instructors would not intentionally ignore proper procedures and techniques. Due to their vast experience, the instructors could be counted on by the trainees to do a proficient job without close monitoring and scrutiny. This potentially allowed the trainees to primarily focus monitoring and attention towards the other trainee crewmembers on board. It also provided less opportunities for critiques of crew coordination by the evaluators. While the trainees could be critiqued for improper coordination with the instructors, it was highly unlikely that the instructor would engage in such actions or communication so as to be critiqued for his or her interaction with the trainee. The presence of instructors also provided a more stable environment, which should enhance the Overall Crew Evaluation rating. Number of instructors on board was determined by trainee answers to item 6 of the student survey (see Appendix I).

Initially, Distinguished graduate status from undergraduate flying programs was to be a third control variable. However, only 3 of 43 respondents reported being a distinguished graduate. As a result, this variable was dropped as a control variable.

The other two control variables in Study 2 were not suitable for Study 3. Duration of mission qualification was not applicable as these trainees were not yet mission qualified. The number of flights in the last 90 days was also not applicable. These trainees were at Altus for the previous 90-day period undergoing training, therefore, the number of flights was very similar across subjects.

Results

Determination of Significance. Similar to Study 2 and for the same reasons, an alpha significance level of .10 was chosen for Study 3. All p values reported will be one-tailed unless noted otherwise.

Descriptive Statistics. The variables of interest in this study were either measured at the crew level or measured at the individual level and then aggregated to the crew level. Therefore, checks for normality and other descriptive statistics were performed on the crew level variables. A visual inspection of the data revealed no serious departures from normality.

Control Variables. The two control variables in Study 3 were the hours per trainee on the aircraft and the number of permanent party instructors on board. When it is stated that experience and ability were controlled for, these two control variables are the reference. Table 19 depicts the correlations between the control variables and performance.

Table 19. Correlations between control and performance variables

	Critiques per Crew# (Objective Performance)	**Overall Crew Evaluation** (Subjective Performance)
Total Hours per trainee	.04	.06
Instructors Onboard	.17	.31

Reverse scored so that more critiques per crew relates to higher performance

Not surprisingly, hours per trainee were not significantly related to performance. This is a training environment where the aircraft is new and unfamiliar to all but the AC (and sometimes the AC can be new to the aircraft). Those with more previous flying time generally tend to possess more "airsense" on the first couple of training flights. However, by the time of the checkride evaluation, crewmembers are generally on equal footing in regards to time in the aircraft and so effects of previous flying experience are diminished. This is in contrast to the Fairchild study where previous flying experience almost always translated into previous KC-135 experience. A t-test revealed that the Altus sample was significantly lower ($p<.001$) in terms of flying hours per crewmember than was the Fairchild sample.

As was seen at Fairchild, number of instructors on board had a positive correlation with crew evaluations (see Table 19). This relationship failed to reach significance due to small sample size. This finding was not surprising due to the positive effect of the presence of instructors discussed earlier.

Transactive memory. The transactive memory variable was initially computed using the means of the Agreement, Accuracy, and Strength components for each crew. Table 20 shows the correlations between the components and the transactive memory variable. Accuracy showed the highest correlation with transactive memory and Agreement the lowest. This is in contrast to Fairchild, where Strength showed the highest and Accuracy the lowest correlation with transactive memory. Why the difference between Study 1 and 2? At Altus only trainees completed surveys, so on most crews there were fewer individuals completing the accuracy rating on one another, increasing the likelihood of greater accuracy. For example the highest Accuracy score recorded by a crew was a perfect 1.0. However this crew only contained two trainees. Therefore, the two trainees had to only agree on one another's greatest strength judgment and the evaluator's judgment (a total of 4 agreements) to produce the perfect Accuracy score. In contrast, a crew of four trainees would have to agree on three others' greatest strength judgments plus the evaluator's judgments (a total of 24 agreements) to produce a perfect 1.0 score. Furthermore, the crew recording the second highest Accuracy score also contained just two trainees. It is clear that the Accuracy score could be easily skewed by number of trainees on board. Indeed, Accuracy was negatively correlated with number of trainees on board (r = -.38, p = .085).

As in Study 2, the Altus crews also scored very low on the Accuracy component. This is surprising given that possibility for greater accuracy with fewer crewmembers completing ratings. Indeed, crews did score slightly higher on Accuracy at Altus (mean .348) over Fairchild (mean .321), but the difference was not statistically significant. Similar to Study 2, only about a third of the time were crewmembers accurate about crew strengths

As in Study 2, none of the transactive memory components were significantly related to another. Because of the suspect nature of the Accuracy scores, their low means, and in order to remain consistent across studies, Accuracy was dropped from the analyses in Study 3. Therefore, the components of transactive memory used in Study 2 (Agreement and Strength) will also be reported for Study 3.

Table 20. Correlations between transactive memory components

	Agreement	**Accuracy**	**Strength**	**T.M.**
Agreement				
Accuracy	.19			
Strength	.19	.07		
T.M.	.49**	.85***	.52**	

**$p < .05$. ***$p < .01$

Collective efficacy. Collective efficacy was measured from four items on the survey using three different scales. The scales were then combined and averaged through z scores. The scales showed good interater reliability between individuals; Cronbach's (1957) alpha was .73.

Criterion Variables. Unlike Study 2, the two criterion variables, Critiques per Crew and Overall Crew Evaluation did not exhibit a significant positive relationship ($r = .14$, n.s.). Keep in mind, that Critiques per Crew were reverse scored so that higher critiques per crew indicated better crew performance. There are two possible explanations for this result. First, the Overall Crew Evaluation rating could have been influenced by the presence of permanent party instructors acting as crewmembers during the trainee's flight. This presence could have created "a halo effect" for the trainees by giving the impression that the trainee's crew coordination skills were stronger than they actually were. The smooth crew effectiveness could actually have been facilitated by the presence of the instructors, not the abilities of the trainees. Support for this argument is derived from the positive correlations between number of instructors on board and Overall Crew Evaluation shown in Table 19. The correlation is higher than that of the Instructor - Critiques per Crew value.

On the other hand, the Critiques per Crew rating could be in error. Because it is a training environment, there tend to be a greater number of critiques on a checkride at Altus compared to an operational base such as Fairchild. It may be more difficult to separate the critiques related to effective crew functioning and other types of critiques. As a result, the Critiques per Crew value may not be as accurate as it would be in an environment with less total critiques. Likely due to the inexperience of the ratees, the mean number of Critiques per Crewmember was nearly double at Altus (mean = 1.625) of what it was for Fairchild (mean = .886). These greater number of critiques, leave open the possibility that it was more difficult to separate crew functioning critiques from individual critiques.

In sum, it is difficult to identify the most appropriate Criterion variable in Study 3. Therefore, results using both criterion variables will be reported, keeping in mind that any disparity in results must be questioned.

Hypothesis 1,2,3. After the data were examined for normality and the two aggregate variables were analyzed, the hypotheses were tested in turn. The first three hypotheses, state that teams with higher levels of transactive memory will outperform teams with lower levels of transactive memory with greater than three team members and in the field. Therefore, crews with higher transactive memory scores should have higher performance scores. After partialing out the effects of experience and ability, transactive memory components do have a positive relationship with the performance variables (see Table 21)

Table 21. Partial correlations between T.M. components and performance

	Critiques per Crew#	**Overall Crew Evaluation**
Agreement	-.13	.21
Strength	.34	.20

Reverse scored so that more critiques per crew relates to higher performance
***$p < .01$

Strength demonstrated positive relationships with both criterion variables. It is noteworthy that these correlations are very similar to those in Study 2 (see Table 10), indicating high consistency in these relationships across samples. Unfortunately, due to small sample size, none of these correlations reached a statistically significant level. Agreement demonstrated a positive relationship with Overall Crew Evaluation, but a nonsignificant negative relationship with Critiques per Crew. Recall, that the two criterion variables were not highly correlated. Taken as a whole, there is some support for Hypotheses 1-3, that transactive memory positively influences performance.

Hypothesis 4. Hypothesis 4 deals with the effect of turnover on transactive memory and performance in turn. This idea is considered in Hypothesis 5, by using a measure of consistency in crew composition. See the Hypothesis 5 section for the results.

Hypothesis 5. This hypothesis concerns the effects of training on performance. It was hypothesized that teams trained together will foster development of transactive memory systems which will result in increased performance over others trained in separate teams.

Previously, we saw that there was evidence that transactive memory components predict performance. Bivariate correlations were conducted to determine the relationship between crew training hardness and transactive memory. Strength showed strong positive correlations with crew training hardness (see Table 22). Similar to Study 2 (see Table 11), Agreement was not related to crew hardness. Again, this is an indication that Strength is the most effective predictor of the transactive memory components. The strong relationship between crew training hardness and Strength supports the contention that training crews together positively effects transactive memory. These results also support Hypothesis 4, that increased turnover will negatively effect transactive memory and performance in turn.

Table 22. Correlations between transactive memory and crew hardness

	Crew Hardness
Agreement	.10
Strength	.56**

$^{**}p < .05$

In further analyses, partial correlations were conducted to test whether Crew Training Hardness predicted performance after controlling for experience and ability. Results indicate that Crew Training Hardness had practically no relationship with either Critiques per Crew (r=.03, n.s.) or Overall Crew Evaluation (r=-.07, n.s.) after controlling for experience and ability. In sum, Hypothesis 5 is supported.

Hypothesis 6. Hypothesis 6 deals with the duration of squadron assignment being positively related to transactive memory. Aircrews trained and evaluated at Altus are kept within the same squadron, therefore, this hypothesis could not be tested.

Hypothesis 7. The seventh hypothesis states that collective efficacy will be a valid predictor of flight crew performance. To test this hypothesis, collective efficacy was correlated with both the performance criterion variables. After controlling for experience and ability, collective efficacy demonstrated no significant relationship with performance (see Table 23).

Table 23. Partial Correlations between collective efficacy and performance

	Critiques per Crew#	Overall Crew Evaluation
Collective efficacy	-.09	.18

Reverse scored so that more critiques per crew relates to higher performance

These results were surprising in light of previous research and the results of Study 2. It seems that the young training crews at Altus may have been suffering from a naive sense of overconfidence. When their confidence levels are compared with those of the more mature and experienced crews at Fairchild, the Altus crews show higher efficacy levels ($t32=1.43$, $p=.081$). The results are more starkly contrasted in one of the four collective efficacy components; confidence level that at least one member of the crew will receive an "Exceptional Performance" rating on the checkride. Exceptional Performance (EP) ratings are rare (given to approximately 10% of all evaluatees) and therefore confidence of its occurrence should be lower. However, the Altus crews showed significantly higher levels ($t32=2.36$, $p=.012$) of efficacy on this component (see Table 24)

Table 24. Collective efficacy levels at Altus and Fairchild

	Altus	Fairchild
% Confident of an Exceptional Performance Rating	75.39**	62.46

**$p < .05$

These results could indicate a failure to predict due to inflated efficacy levels or perhaps efficacy's inability to predict performance with a lack of experience and ability. Regardless of the reasoning, Hypothesis 7 was not supported.

Hypothesis 8. Hypothesis 8 asserts that transactive memory will explain variance in flight crew performance, above and beyond that of collective efficacy. Previously we saw indications that transactive memory was predictive of performance after controlling for experience and ability. Does this relationship change when we control for the effects of collective efficacy? Correlations between transactive memory components and performance were conducted controlling for the effects of experience, ability, and collective efficacy. The results are seen in Table 25.

Table 25. Correlations between T.M. components and performance controlling for collective efficacy

	Critiques per Crew#	Overall Crew Evaluation
Agreement	-.12	.19
Strength	.36	.17

Reverse scored so that more critiques per crew relates to higher performance
***$p < .01$

When compared with the Results of Table 21 where only experience and ability were controlled for, it is clear that the relationships between transactive memory components and performance showed minimal change. This indicates that collective efficacy did not mediate the transactive memory - performance relationship. Due to the small sample size the relationships did not reach statistical significance. However, the relationships are clearly in the positive direction for Strength and in once case for Agreement. On the whole, these results provide partial support for Hypothesis 8.

Further Analyses As was done in Study 2, further analyses were conducted to investigate the relationships of other variables with the key variables of transactive memory and collective efficacy. Bivariate correlations between the two control factors and collective efficacy were examined first (see Table 26). Correlations are lower than in Study 2, and surprisingly collective efficacy showed no relationship with hours per trainee or number of instructors on board. This provides evidence that collective efficacy scores were naively inflated as will be explored in the Discussion section.

Table 26. Correlations between control variables and collective efficacy

	Collective efficacy
Total Hours	-.02
Instructors Onboard	.11

Also as in Study 2, partial correlations were conducted between transactive memory components and collective efficacy (see Table 27). The relationship between Strength and collective efficacy was not nearly as strong (about one third as strong, see Table 17 for contrast) as it was in Study 2. However, the relationship between the two was positive. Again results failed to attain statistical significance due to small crew sample size. Agreement showed little relationship with collective efficacy in either study.

Table 27. Correlations between transactive memory and collective efficacy controlling for experience and ability

	Collective efficacy
Agreement	.10
Strength	.18

Discussion of Study 3

Two findings stand out in Study 3. First are the encouraging results indicating a positive relationship between transactive memory and performance. While failing to reach statistical significance, these results clearly indicated a positive relationship between transactive memory components and performance after controlling for experience and ability. Strength again was the most potent transactive memory component, while Agreement offered little explanatory power. The second finding was the diminished importance of collective efficacy in predicting performance at Altus AFB (vis-à-vis Fairchild). Both of the findings need to be explored in the light of higher collective efficacy levels and the different environments involved.

The training crews at Altus clearly exhibited higher overall efficacy levels than the Fairchild crews. In particular they were much more confident of attaining the highest (and rare) performance rating of "Exceptional Performance". It would seem that the crews were a bit naive on this point. It could be that the training crews were unaware (due to lower experience with checkrides) of the difficulties involved with doing well on the checkride. Alternatively, these high collective efficacy levels could be a form of impression management, where the crews know that it is socially desirable for aircrew members to appear confident, or even "cocky". It could be that the instructors failed to give the trainees a realistic expectation concerning checkride results, in particular the chance of an EP rating. However, these warnings could have been received from the instructors, but not heeded by the trainees. Perhaps in an effort to increase efficacy through verbal persuasion (Bandura, 1982), the instructors could have intentionally convinced the trainees that they had a high probability of maximum success on the checkride.

Whatever the explanation, it is clear that the trainees' reported higher efficacy levels. It is incredulous that a younger, more inexperienced crewmember would have higher efficacy scores than a mature KC-135 crewmember. Perhaps this false bravado explains the lack of significance of the collective efficacy score with performance. There could be a ceiling effect on efficacy and as a result a lack of variance in the variable to exhibit predictive power. Evidence for this claim is as follows. Altus crews had higher mean scores than Fairchild crews on all four efficacy components (only one reached statistical significance). Two of the efficacy components had a mean rating near the highest possible level for crews at both Fairchild and Altus. The two remaining components were "confidence of an EP rating" (with Altus scoring significantly higher) and "chances of three critiques or less on the checkride" (with Altus scoring higher, but not reaching statistical significance). If these higher means indicate unrealistic collective efficacy judgments, these values would increase random error and result in lower correlation values.

Though it is clear that Altus had (unjustifiably) higher collective efficacy levels, there may be another reason that collective efficacy lost its predictive power at Altus. Altus is a training base where crewmembers have less experience and ability in the KC-135 aircraft than do crews at an operational base such as Fairchild. Perhaps in the absence of high levels of experience and ability, collective efficacy may not separate the average crew from the exceptional crew. It may only be at higher levels of ability that collective efficacy makes a difference. As an analogy, in professional sports we often hear the expression "at this level it becomes a mental game, do you think you can win, when the other guy doesn't". Collective efficacy may be a separator at these greater experience and ability levels. But during training other constructs may be more important.

Transactive memory enjoyed a positive relationship with performance with or without the presence of collective efficacy. Training together as a team and learning "who knows what, who does what, and who is good at what" may be more important when becoming proficient at a complex task. It could be that this knowledge allows one to focus on other aspects of the task and less on the teamwork component. Additionally, as hypothesized in the Discussion of Study 2, the higher transactive memory scores of less experienced crews may keep their performance from dipping to unsatisfactory levels.

How do the findings of Study 3 compare with the revised model presented in Study 2? The answer is: very favorably. Figure 4 is presented again below as Figure 6 for convenience.

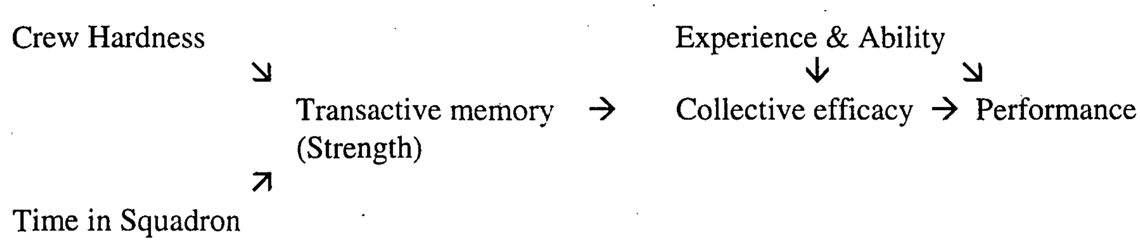

Figure 6. A revised model of the transactive memory / collective efficacy process.

The results of Study 3 clearly indicate that Crew Hardness was highly related to the transactive memory components. Time in the Squadron was not a variable in Study 3. Transactive memory and its components were related to collective efficacy, but not to the degree seen in Study 2. This could be due to the inflation of collective efficacy or the difference in flying environments.

The control variables of experience and ability were not as important for collective efficacy at Altus. This can be explained by remembering that experience and ability were much lower at Altus as compared to Fairchild, while collective efficacy was higher at Altus than at Fairchild. The artificial inflation of collective efficacy is likely to blame for the lack of relationship between experience, ability, and efficacy.

Furthermore, it was found that collective efficacy was not predictive of performance. This result was somewhat surprising after collective efficacy's strong role in performance in Study 2. As was stated earlier, collective efficacy may have been naively inflated at Altus. It is also equally likely that collective efficacy's role in performance may not be as important when experience and ability are low. Other factors (including transactive memory) may play a larger part under these conditions.

The relationship between transactive memory and performance was quite consistent between samples. There are good indications that transactive memory and its components are predictive of performance. Unfortunately, due to small sample size these correlations did not reach statistical significance. Similarly, it was found that the relationship between transactive memory and collective efficacy was positive. However, this relationship was not nearly as strong as it was in Study 2. Again, this could be due to the collective efficacy inflation at Altus or it could be due to the difference in flying environments. Perhaps the role of transactive memory in collective efficacy is weaker in situations where experience and ability are not well developed

The shortcomings of Study 3 are similar to those of Study 2; a small sample size. The number of individuals surveyed was 43. However, once the individuals were placed into their crews (which is the variable of interest) the sample size falls to 15. Efforts were made to increase the sample size. However two forces acted against this effort. First, the investigator could not be on site to encourage and monitor participation. Second, a portion of the permanent party instructors at Altus AFB were dispatched to participate in the real world conflict in Kosovo. Therefore, those instructors left behind were left with an even higher workload than normal. Prior to the conflict, the squadrons were already below 100% manning, so the Kosovo deployment worsened an already high workload situation. Therefore, instructors did not always make time for returning surveys. Additionally, those instructors charged with ensuring the student crews received surveys were distracted by other primary duties.

The flip side of the small sample size issue is the fact that significant findings can be viewed as robust relationships among the variables that were not a product of "just more data". In particular, strong relationships were found between crew training composition and transactive memory.

The practical applications of these findings are to encourage the leadership at Altus AFB to make crew consistency a top priority in scheduling training sorties. This should pay off with better performance at the end of the training program. In turn, operational squadrons, such as Fairchild, will then be receiving crewmembers able to perform at a higher level. Transactive memory was clearly more important to performance at Altus than was collective efficacy or even experience and ability.

Theoretically an examination needs to be made on the relationship of collective efficacy and performance with experience and ability as a potential moderator. There is some evidence from Study 3 that in a learning or training environment of a complex task, collective efficacy may not be important in task performance. Perhaps more accurately put, collective efficacy seems to distinguish exceptional and average performance only in operational environments. Results of Study 3 indicate that collective efficacy does not distinguish such performance in a training environment. Is there a theoretical reason why this relationship exists? Are experience and ability moderators of the collective efficacy - performance relationship?

A potential answer to that question and also an area of both theoretical and empirical work is the notion of "naive" or "inflated" collective efficacy. It seems in Study 3, that extremely high collective efficacy levels simply failed to predict performance in a meaningful way. But can there be a down-side to inflated collective efficacy? Could it actually be harmful for collective efficacy to be naively high? Perhaps it could result in reduced vigilance or attention to detail, or perhaps decreased motivation. It should be clearly stated that I am making no inference that such dangers exist based on the results of this study. Rather, I offer the question as a logical extension to the high collective efficacy scores encountered during Study 3.

On a related manner, is there a way to construct questions in order to more accurately gauge collective efficacy? In accordance with the efficacy literature (c.f. Gist and Mitchell, 1992) several of these efficacy questions were stated so as to ask about specific performance attainments. It could be that unless individuals have the necessary experience with the task or heed instruction from experts on performance levels, they may possess improper performance level expectations. Mitchell, Hopper, Daniels, George-Falvy, and James (1994) provided empirical evidence in their work on self efficacy that predicted performance is harder to judge accurately with less experience. If indeed performance levels were difficult to judge at Altus, the young crewmembers might have easily fallen prey to inflated collective efficacy. This could especially be the case in an environment like the Air Force with high efficacy "demand characteristics". In this environment, three of Bandura's (1982) four categories of experience are utilized to develop efficacy: 1. Vicarious experience (modeling of other military flyers), 2) verbal persuasion (Aim High / Air Force), and 3) physiological arousal (e.g. anxiety about the checkride).

Each of the three studies on transactive memory offer a unique perspective on the construct. Additionally, we have learned much about collective efficacy and performance. The next section offers an overall discussion and critique of the three studies, and follows with implications for practice, empirical work, and theory.

Chapter 7

Conclusion

Overall Discussion

Taken together, the results of the lab and two field studies inform our understanding of both transactive memory and performance. The field studies clearly demonstrated that turnover or failure to train as integral teams can have an adverse effect on transactive memory. All three studies showed a positive association between higher transactive memory and some elements of performance. The lab study hints at why transactive memory may translate into better performance. It seems that crewmembers have higher situational awareness of events outside of the cockpit under high transactive memory conditions. Perhaps the improved coordination as a result of knowing who knows what, who does what, and who is good at what, allows crewmembers to focus on important aspects of the situation versus having to closely monitor the actions of fellow crewmembers. Furthermore, increased transactive memory may lessen the need for verbal coordination of actions. Additionally, crewmembers may feel they are not performing as well in a low transactive memory situation and are less confident they can gauge other crewmember's situational states under these conditions.

The relationship of transactive memory and performance seems consistent across studies. Transactive memory may help maintain less experienced crewmembers at the "average level" of performance and keep them from falling to unsatisfactory levels. However, there are indications that under conditions of high experience and ability that collective efficacy may mediate the relationship between transactive memory and performance. The field studies also indicate that transactive memory is an important input into the crew's collective efficacy. The knowledge of one another add to the team's confidence.

Collective efficacy was shown to be a powerful predictor of performance under conditions of realistic collective efficacy judgments. Specifically, collective efficacy tends to separate the exceptional crew from the average crew under conditions of experience and well-developed ability. However this relationship can be "short-circuited" when team member's possess a naively inflated sense of collective efficacy. Under these conditions transactive memory's effect on collective efficacy can be reduced as well.

Experience and ability were demonstrated to be related to both performance and collective efficacy, but not as strongly as one might think. For example, in Study 2 collective efficacy added a great deal of explanatory power about performance even after experience and ability were controlled for. Experience and ability do influence collective efficacy, but this relationship is also reduced under inflated efficacy ratings.

Practical Implications

Is it important to keep aircrews together during training and in operational squadrons? That is one of the key questions that ignited this research. The answer seems to be yes. If squadron leadership desires the highest level of performance, they should strive to keep crews together. By keeping crews together, transactive memory is increased. This increased knowledge of who knows what, who does what, and who is good at what may allow the crews to have higher situational awareness outside of the cockpit and greater ability to judge one another inside the cockpit. This will positively influence performance. Additionally, a crew higher in transactive memory will be a more confident crew. Results were clear that crews that know each others strengths well, tend to be the crews that are more confident. This finding assumes greater importance, when one considers the role of experience and ability. Experience and ability were both positively related to collective efficacy in an operational environment. These commodities are quickly being lost as the Air Force (and other services) endure retention problems. Experience and ability are not quickly or easily replaced. Therefore, other measures need to be utilized to keep crew efficacy high. Maintaining high transactive memory through crew integrity is one tool that squadron leadership can control and influence.

Is the price one pays to keep the crew together, worth the additional performance earned? Only the leadership can ultimately answer that question, but there are several factors to consider. Keeping crews together can be a burden from a scheduling perspective; it is sometimes hard to keep individuals together to fly in this high tempo operations environment. A potential compromise is offered: For routine missions, crew integrity may not be as important. Crews should be able to perform proficiently under these circumstances. However, whenever crews will be entering high stress, unknown environments, I recommend that a "hard crew" be employed. Perhaps it would be wise to have a few "hard crews" on hand for short notice deployments such as the Kosovo crisis. However, the squadron leadership may not have such a luxury. As an alternative, crews are sometimes deployed with short notice to a forward location where they then wait for hostilities to begin. During this waiting period crews could have several training flights together to increase both transactive memory and collective efficacy. This should translate into a performance advantage that could make a difference in a crisis situation.

Anecdotally, the Air Force employed such a strategy during the Desert Shield portion of the Southwest Asia campaign. When hostilities began under Desert Storm, crews were already configured as "hard crews". Their success rate speaks for itself. These studies provide empirical evidence for this practice.

It should also be mentioned that these results indicate that flying crewmembers together from the same squadron, even without many flights together, has advantages. These crews tend to have higher transactive memory from simply interacting in a flying atmosphere. On bases housing several squadrons, or at forward locations with personnel from several squadrons, I recommend that personnel from the same squadron fly together as much as possible.

Limitations

The most salient of the limitations of this research has already been acknowledged. All three studies suffered from a small sample size. Working with natural teams whether in the lab or the field, can be expensive and time consuming. It is difficult to assemble the various team members for research, especially when they are busily engaged in their normal tasks and duties. Study 2 is a classic example of these rigors. There were 73 individual respondents, which translates into 19 crews for the study. There were twenty crews evaluated during the Fairchild inspection. Evaluators had hoped to have at least 25, however, the Kosovo crisis erupted three days before the inspection. Nine crews were deployed, resulting in a fewer number available for the inspection. Evaluators were delighted to even have 20 crews at that point.

Keep in mind that there are several positives associated with the samples. The lab study enjoyed one of the highest quality subject pools found anywhere in the literature. All the aviators had over 1500 hours in military aircraft. Seven of the eight were combat veterans of either the Vietnam or Persian Gulf War. One of the participants has been the subject of an aviation book chapter with his three "MIG kills" in Vietnam. These flight subjects offer greater validity in simulator studies than using college students. The same can be said for the field studies where actual KC-135 crews, evaluators, and aircraft were employed in the study. A study cannot be conducted in a more realistic environment. It seems that the sample size trade-off for the quality of the sample itself, was a good one.

Additionally, strong findings emerged using small sample sizes, indicating robust phenomenon at work. For example, the collective efficacy relationships in Study 2 enjoyed correlations in the .60 range. Crew hardness translated into transactive memory with correlations around .50 in both field studies. Transactive memory exhibited an extremely consistent relationship with performance across field studies. The sample sizes employed make these results even more impressive.

There are several potential confounds associated with Study 3. With no investigator on site, data collection relied on volunteer efforts of several instructors at Altus AFB. While their efforts are to be commended, they could not be expected to exert that same effort as the Investigator in ensuring a high survey return rate. It is not clear how many surveys were issued to crews. Surveys were issued from a central dispatch desk. On occasions some crews did not pick up their survey package. On other occasions crews picked up more than one package and then discarded the extra package. Therefore, it is not known how many crews chose to respond. However, threats of response bias are greatly reduced by considering that I am sampling from a very homogenous population. All the trainees at the base encountered similar selection procedures into the military in general and into aviation in particular. Training at initial training flying assignments prior to Altus AFB is extremely standardized by the Higher Headquarters Standardization and Evaluation Division. Headquarters prides itself on consistency in training. It is very likely that each training crew at Altus was more similar than different from the next aircrew.

The evaluator forms were kept simple. The investigator briefed many of the evaluators on the specifics of the forms in late February, just prior to data collection. On a separate occasion, the chief evaluators were briefed in additional detail and in turn covered the form with their personnel at the weekly evaluator's meeting. Results of the survey indicate that evaluators avoided "firewalling" their ratings. Firewalling is the tendency to give all evaluatees the highest marks possible. The other major rating area, Critiques per Crew, was to be translated from the official Air Force evaluation form onto the researcher's evaluation form. The only confusion in this case, may be whether a critique item could be attributable to crew coordination. In most cases, this would be a clear call, however, if several critiques are related, it may be difficult to separate the exact number attributable to crew coordination as discussed earlier.

Another draw back to Study 3 was the mixing of instructors with trainees during the evaluation. For example, Crew 1 is having a checkride on the AC and the Co. The Boom Operator and Navigator are both students earlier in the training sequence. Meanwhile, Crew 2 is having a checkride on the AC and Co also. There are no other trainees available, so permanent party instructors acting as regular crewmembers fill the Boom Operator and Navigator positions. It can be argued that Crew 2 has it much easier as they will not be required to monitor the other crewmembers as closely as they would have to do with fellow trainees. They can then focus more on the task at hand. Additionally, it could be more difficult for the Evaluator to distinguish if the good crew coordination exhibited by Crew 2 is due to their crew coordination skills or due to the presence of instructors on board. On the other hand, Crew 1 may possess excellent crew coordination skills that are severely challenged by newer trainees on board.

While the possibility of these factors influencing the finding of Study 3 cannot be ruled out, the fact that the "non-checked" positions on a checkride are randomly assigned helps to rule out this possibility. Through randomization it is hoped that no systematic performance bias crept into the results. Randomization certainly makes such systematic bias less probable. It was desired to have each checkride to be conducted with an evaluatee in every position. However, real world constraints make this an impossibility. The effects of instructors mixing with trainees is a condition that all skill and ability levels have an equal chance of encountering. The same can be said for the issue of mixing crewmembers flying their checkride with trainees flying at a much earlier phase in their program.

Due to a shortage of qualified subjects, Study 1 was unable to obtain a fully counterbalanced design. The counterbalance is employed to protect against any difference in advisors. Sixteen subjects would have been required for such a counterbalance. Therefore, the first half of the counterbalance design was employed. T-tests indicated no difference between subjects during the first week of testing, when any adverse effects due to a difference in the advisors would emerge. Additionally, only Week 2 results were used in the transactive memory analysis; further eliminating the possibility of performance differences due to experimental design.

The final limitation to the field studies concerns the issue of the measurement of transactive memory. It was hoped that the three components would exhibit the high correlations found by Moreland and colleagues (1998), but they did not. Following their suggestion, the individual components of transactive memory were used to investigate relationships with the criterion variables and Strength was found to be a reliable predictor across studies. The primary measurement issue deals with the Accuracy component. It seems clear from the literature that Accuracy is an important component of transactive memory. In both Study 2 and 3, the Accuracy component exhibited very low mean levels. As a result, the predictive validity of transactive memory may have suffered. When Accuracy was removed from the transactive memory measurement, the construct showed vastly improved criterion related validity. Further studies should endeavor to more precisely measure this component.

Future Research Directions

The studies described here point to several avenues of future research. Studies 2 and 3 revealed two different patterns for the collective efficacy - performance relationship. Under normal levels of experience, collective efficacy was a potent predictor of performance. However, under training conditions, collective efficacy demonstrated practically no relationship with performance. Is experience and ability a moderator of the Collective efficacy - Performance relationship? Further studies, especially in the field, should be conducted by manipulating experience levels and examining if and how the collective efficacy - performance link is changed. Alternatively, collective efficacy and performance could be contrasted for two conditions, a training condition and a "mature condition" where the task has been well learned. Another interesting test would be a longitudinal approach where collective efficacy - performance relationships are tested for crews during training and then later once the skills have been mastered.

On a related matter, collective efficacy was found to be highly positively related to transactive memory in Study 2, but much less so in Study 3. Again, does the relationship of collective efficacy and transactive memory change depending on the training or experience level of the participant? This issue could be addressed in similar manner to the collective efficacy - performance issue described above.

The differences in the transactive memory, collective efficacy, and performance relationships listed above could be attributed to a false or naive sense of collective efficacy. Study 3 revealed inflated collective efficacy levels on the part of new trainees. Can these inflated collective efficacy levels be harmful to performance? Is it possible that such inflated levels may lead to a false sense of security and a resulting lack of vigilance. Alternatively, the inflated levels could lead the subject to attempt a performance level that could be harmful to the subject or others involved. It could also be that these inflated levels simply "wash-out" any predictive power on the part of collective efficacy and are otherwise not harmful. Testing of these hypotheses would be difficult. Manipulating collective efficacy levels in order to test for dangerous results would be unethical.

The inflated collective efficacy issue also highlights the importance of researchers thinking through the efficacy questions posed to respondents in order to protect against such inflated ratings. However, as was seen in these examples, the questions had a very specific nature and were still inflated for the training groups. Researchers cannot control all of the subject's bents.

Future research should also focus on how to best measure the Accuracy component of transactive memory. Moreland and colleagues (1998) were able to consistently measure Accuracy when it came to a very concrete result such as wiring a radio. They found Accuracy levels comparable to both Agreement and Strength levels. Study 2 and 3 required a more abstract Accuracy estimate, that of defining another person's strength from a forced choice list. This task proved to be too difficult. The lesson learned seems to be that the more abstract the Accuracy estimation, the greater the chance of poor results. If the task is comprised of mostly abstract components, it may be difficult to find a suitable accuracy measurement. There is a balance to be reached in finding an accuracy component that is not too simple (resulting in ceiling effects) or as was in this case, an accuracy component that is too difficult (resulting in floor effects). Perhaps accuracy questions should focus more on specific situations that a team may encounter. So given a certain situation, how would you expect Team Member A to react? This may be a more fruitful approach to accuracy and merits further research.

Researchers should also continue to pursue transactive memory work mindful of the consistent results of the field studies using only the Strength component. It may be that the ease of measurement, coupled with the predictive power of the Strength component, may render the more difficult collection of the other two components of transactive memory of little value. This could potentially provide practitioners with a simple tool to gauge team transactive memory levels quickly and efficiently. Alternatively, research should be pursued on inaccurate Strength estimation. A false or inaccurate sense of Strength could be detrimental. This situation could arise when members are in the same organization for a long period of time, but not on the same team. Once put on a team, the members may feel they know each other's strengths well, when in fact they do not. Would this be harmful? Members in this situation may actually take longer to figure out the actual strengths of fellow crewmembers.

Future research should be focused on potential boundary conditions of transactive memory. There may be context contingencies that increase the value of transactive memory. Perhaps transactive memory would be most value in novel, unusual, or crisis situations requiring flexibility on the part of the team. These non-routine situations may present several possible solutions and the team high in transactive memory may better gauge the best solution for that team. The team high in transactive memory may be better able to compensate for weaknesses in team by simply being aware that they exist. This would allow even a team with poor skill levels to perform better than they may otherwise. Furthermore, team high in transactive memory may be better able to select replacements for departing team members as they are acutely aware of the knowledge that has departed with the outgoing member. Certain tasks may moderate the importance of transactive memory to the team. In a similar vein the presence or absence of assigned team role may increase the importance of transactive memory. It is possible that teams without assigned roles would receive greater benefits from transactive memory as they may be more flexible in how they operate.

Finally, research is needed on how to develop or build up transactive memory in a team. Time on the team is clearly one avenue to develop transactive memory. However, there may be further steps that a team can take. Rulke and Rau (1997) found that early in the team's life cycle, teams employ a strategy of small spiral encoding cycles of question-expertise-coordination. It seems that this practice would be beneficial for mature teams as well. Other group techniques such as Pulling apart Success exercises could also facilitate the development of transactive memory. Groups do not tend to re-assess themselves naturally, so this practice may require a team intervention. Another important time for such exercises may be when outgoing team members are replaced by newcomers. With arrivals and departures, this spiral encoding cycle could be beneficial. In groups with an assigned leader, the leader could institute such practices to increase the team's transactive memory and performance.

Summary

Taken as a whole, the results of these studies point to the positive relationship between transactive memory and performance. Though these effects were not exceptionally strong, a leader desiring to improve performance should take heed. Training teams in groups and keeping them together as much as possible during normal real world functioning, should result in improved performance. Additionally, the powerful role of collective efficacy in determining performance in mature teams was highlighted. This performance difference may be what separates the "exceptional" from the "average". Means to improve collective efficacy include increasing experience, ability, and the team's transactive memory.

REFERENCES

Bandura, A. (1977). Self-Efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215

Bandura, A. (1982). Self efficacy: Mechanism in human agency. *American Psychologist, 37*, 122-147.

Bandura, A. (1986). *Social foundations of thought and action: A social-cognitive view.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning: *Educational Psychologist, 28*, 117-148.

Bandura, A. (1997). *Self-Efficacy: The exercise of control.* New York, NY: W.H. Freeman.

Brannick, Roach, Salas (1993). Understanding team performance: A multimethod study. *Human Performance, 6*, 287-308.

Cannon-Bowers, J.A., Converse, S. (1993). Shared mental models in expert team decision making. In *Individual and Group Decision Making, Current Issues*, (Ed.) Castellan, N.J., 221-246.

Cascio, W.F., Zedeck, S. (1983). Open a new window in rational research planning: Adjust alpha to maximize statistical power. *Personnel Psychology, 36*, 517-526.

Church, A. H. (1996). From both sides now: The power of teamwork-fact or fiction? *TIP, 34,* 85-96.

Endsley M.R. (1995). Measurement of situation awareness in dynamic systems. Special Issue: Situation awareness. *Human Factors*, *37(1),* 65-84.

Fuller, S.R., Aldag, R.J. (1998). Organizational tonypandy: Lessons from a quarter century of the groupthink phenomenon. *Organizational Behavior & Human Decision Processes, 73*, 163-185.

Gibson, C.B. (1996) They do what they believe they can? Group-efficacy beliefs and group performance across tasks and cultures. *Paper presented at the 1996 Academy of Management Meetings, Cincinnati OH*

Gist, M.E., Mitchell, T.R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review*, *17*, 183-211.

Goodman, P.S., Ravlin, E., Schminke, M. (1987). Understanding groups in organizations. In L. Cummings, & B. Staw (Eds.), *Research in Organizational Behavior, 9*, 121-173. JAI Press.

Gruenfeld, D.H., Hollingshead, A.B. (1993). Sociocognition in Work Groups The evolution of group integrative complexity and its relation to task performance. *Small Group Research, 24*, 383-405.

Guzzo, R.A., Yost, P.R., Campbell, R.J., & Shea, G.P. (1993). Potency in groups. *British Journal of Social Psychology, 32*, 87-105.

Hackman, J.R. (1983, November). A normative model of work team effectiveness (Teach. Rep. No 2.). New Haven, CT: Yale School of Organization and Management, Research Program on Groups Effectiveness. In Goodman et al. 1987

Hollingshead, A.B. (1998). The impact of knowing who knows what. *Paper presented at the Academy of Management meetings, San Diego, CA 1998.*

Hegel, P.T. (1807/1910). *The phenomenology of mind* (Transl.). London: Allen and Unwin.

Hodges, L., Carron, A.V. (1992). Collective efficacy and group performance. *International Journal of Sport Psychology, 23*, 48-59.

Janis, I.L. (1972). *Victims of Groupthink.* Boston: Houghton-Mifflin.

Klimoski, R., Mohammed, S. (1994). Team mental model: Construct or Metaphor? *Journal of Management, 20*, 403-437.

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*, 384-393.

Lindsley, D.H., Mathieu, J.E., Heffner, T.S., Brass, D.J. (1994). Team efficacy, potency, and performance: A longitudinal examination of reciprocal processes. *Paper presented at the 9th Annual SIOP Conference, Nashville, TN.*

Little, B.L., Madigan, R.M. (1994). Motivation in work teams: A test of the construct of collective efficacy. *Paper presented at the Academy of Management meetings, Houston, TX 1994.*

Mischel, L.J., Northcraft, G. (1997). "I think *we* can, I think *we* can...". The role of efficacy beliefs in group and team task effectiveness. in *Markovsky, (Ed.) Advances in Group Processes.*

Mitchell, T.R., Hopper, H., Daniels, D., George-Falvy, J., James, L. (1994). Predicting self-efficacy and performance during skill acquisition. *Journal of Applied Psychology, 79*, 506-517.

Moreland, R.L., Argote, L., Krishnan, R. (1996) Socially shared cognition at work. Transactive memory and group performance. in *Nye & Brower (Eds) What's Social about Social Cognition? Research in Socially Shared Cognition in Small Groups. Sage*, 57-84.

Moreland, R.L., Argote, L., Krishnan, R. (1998) Training people to work in groups. in *Tindale, Edwards, & Posvac (Eds) Applications of Theory and Research on Groups to Social Issues: Plenum.*

Nelson, R.R., Winter, S.G. (1982). *An evolutionary theory of economic change.* Cambridge, Mass. Belknap Press of Harvard University Press.

Osterman, P. (1994). How common is workplace transformation and who adopts it? *Industrial and Labor Relations Review, 47*, 173-188.

Peterson, E., Mitchell, T., Thompson, L., Burr, R. (1996). Group Efficacy and shared cognition as predictors of group process and performance. *Paper presented at the Academy of Management meetings, Cincinnati, OH 1996.*

Polanyi, M. (1959). *The study of man.* Chicago: University of Chicago Press.

Polanyi, M. (1962). *Personal knowledge ; towards a post-critical philosophy.* New York : Harper Torch Books.

Prussia, G.E., Kinicki, A.J. (1996). A motivational investigation of group effectiveness using social cognitive theory. *Journal of Applied Psychology, 81*, 187-198.

Rousseau, J.J. (1767). *A treatise on the social contract.* London: Beckett and DeHondt.

Rulke, D.L., Rau, D. (1997). Examining the encoding process of transactive memory in group training. Paper presented Academy of Management Meeting Aug 1997 Boston Best paper award from Organizational Behavior Division.

Sundstrom, E., DeMeuse, K.P., Futrell, D. (1990). Work teams: Applications and effectiveness. *American Psychologist, 45*, 120-133.

Wegner, D.M. (1987). Transactive memory: A contemporary analysis of the group mind. in *Theories of Group Behavior*, Mullen & Goethals (Eds.). pp 185-208.

Wegner, D.M., Erber, R., Raymond, P. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology, 61*, 923-929.

Wegner, D.M., Giuliano T., Hertel, P. (1985). Cognitive interdependence in close relationships. In W.J. Ickes (Ed.), *Compatible and incompatible relationships.* 253-276. New York: Springer-Verlag.

Weick, K.E., Roberts, K.H. (1993). Collective Mind In Organizations: Heedful Interrelating On Flight Decks. *Administrative Science Quarterly, 38*, 357-381.

Wood, R.E., Bandura, A. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. *Journal of Personality and Social Psychology, 56*, 407-415.

Appendix A
Training checklist for the pilot

F-22 Simulator Training Checklist

1. Conduct training on Fly Training: Dogfighting, Flight Model: Realistic , # enemy aircraft 3, Enemy skill level: Harmless.

2. Review stick device
   -left/right turn, climb/descent
   -throttle operation
   -afterburner (* key)
   -trigger function
   -missile launch button
   -buttons on the joy stick

2. Allow subject time to practice basic flight maneuvers
   left/right turn
   climb/decent

3. Cover practice/evasive maneuvers.
   split S (requires 6,000 ft.) to accomplish. Used when missile is (1 mi < missile < 2 mi) and coming from the 4 to 8 o'clock position. Or turn hard into the missile. Turn into missile coming from the front.

4. Go over the Head's Up Display (HUD).
   Begin at the Heading indicator and go <u>counterclockwise</u>
   heading indicator
   airspeed is in Knots Indicated Airspeed or KIAS often referred to as *indicated airspeed*
   M = mach
   G = g's on the pilot. Grey/blackout can occur at greater than 6 sustained g's. -1.5 red-out
   A = AOA or Angle of Attack. Not important to this simulation
   F = Fuel (Not important)

   Weapon selected and number of weapons remaining
   AIM 120C good for 5-15 nautical miles Triangle sight
   AIM 9X good for within 5 nautical miles Circle sight
   G480 guns dashed circle sight (small circle = where bullets would hit 2,250' away)
   NOTE each of these weapons brings up a different sight. Weapons will be further demonstrated later in the training period.

   Bottom center small triangle is the turn and slip indicator N/A for this simulation

   Bottom right hand corner: #/type of waypoint, relative degrees to next waypoint and distance to waypoint in nautical miles

Appendix A (Continued)

Estimated Time of Arrival (ETA) to next waypoint under current heading and airspeed.

Target List
Type of aircraft and position in the shootlist (in parentheses)
altitude of enemy aircraft
Heading and distance to enemy aircraft
Bogies airspeed
VC = closure speed of bogey

Altimeter in MSL (Mean Sea Level) also capable of giving Radar Altimeter by pushing Shift A key. MSL is fine for this simulation

Missile range bar. Describe and emphasis having the dot in range before firing
left bar: weapons effectiveness range, dot = bogey
thin t bar: max possible range of weapon
right bar: bogey's best weapon max range, arrow = your aircraft

Also cover Shoot cue
Shoot comes up in circle, dot indicates where the bogey is according to the missile's view.
Heading bug

5. Cover important keyboard keys
   * = afterburner
   B = Speed brake
   Below is important only if not using the throttle buttons
   Enter = toggles through the weapons
   S = brings up shootlist / changing weapons deletes shoot list
   T = toggles through the shoot list
   Tab = warp speed

6. Practice a few air to air engagements.
   NOTE: ENSURE PILOT'S HEADS DOWN DISPLAY IS COVERED
   emphasize proper weapon selection
   emphasize how to bring up a shoot list
   emphasize the missile in range bar on the HUD
   emphasize the different sights (shapes), shoot cues, and target vector bar
   emphasize use of the HUD as the Head's Down Display will not be available.
   emphasize evasive maneuvers from incoming missiles and SAMs
      -must listen to advisor
      -split S and turns into the missile
      -what to do in case of black out - lessen the stick pressure or release depending on position

Appendix A (Continued)

7. Fill out a practice SA (situational awareness) form
    -wingmen are on the wing, unless sent elsewhere.
    -airspeed, altitude, and aircraft attitude.

8. Review how to handle wingmen Hit U, ?, ?
    - wingmen, two on the right wing and one on the left unless sent elsewhere

9. Begin the criterion test
    NOTE: ENSURE PILOT'S HEADS DOWN DISPLAY IS COVERED
    twice shooting down all three aircraft without being shot down
    at least twice during the criterion test have them stop and *verbally* state what they would put on the SA form
    NOTE: recommend 25,000 ft on altitude (good for maneuvering) and
    hit E and V for auto chaff, flare, and ECM.

BEGINNING THE MISSION

1. What buttons for the mission "fly mission" ?

2. Hit E, V for auto chaff, flare, and ECM

3. Hit A for autopilot, TAB x 3 to go on warp speed times three.

4. Once near the border, hit ???? TAB to come out of warp speed

5. Go on Pause (P button), fill out sim time on the form.

Appendix B
Situational Awareness Measures

Time of test ____________ (from F22 clock)

NOTE All of the questions refer to the time since the last test

errors

Wingmen
- Twinkletoes 2
  - Distance__________ ____ .... ____ (± 20%)
  - Direction__________ ____ .... ____ (± 30 deg)
- Twinkletoes 3
  - Distance__________ ____ .... ____ (± 20%)
  - Direction__________ ____ .... ____ (± 30 deg)
- Twinkletoes 4
  - Distance__________ ____ .... ____ (± 20%)
  - Direction__________ ____ .... ____ (± 30 deg)

Bogies? (y/n)_____ ____ .... ____
- Bogey 1, type________ ____ .... ____
  - Distance__________ ____ .... ____ (± 20%)
  - Direction__________ ____ .... ____ (± 30 deg)
- Bogey 2, type________ ____ .... ____
  - Distance__________ ____ .... ____ (± 20%)
  - Direction__________ ____ .... ____ (± 30 deg)
- Bogey 3, type________ ____ .... ____
  - Distance__________ ____ .... ____ (± 20%)
  - Direction__________ ____ .... ____ (± 30 deg)
- Bogey 4, type________ ____ .... ____
  - Distance__________ ____ .... ____ (± 20%)
  - Direction__________ ____ .... ____ (± 30 deg)
- Bogey 5, type________ ____ .... ____
  - Distance__________ ____ .... ____ (± 20%)
  - Direction__________ ____ .... ____ (± 30 deg)

Aircraft
- Altitude ________ ____ .... ____ (± 20%)
- Attitude________ ____ .... ____
- Speed_________ ____ .... ____ (± 20%)
- Wpn selected_______ ____ .... ____

Weapons shot? (y/n)________ ____ .... ____
- number__________ ____ .... ____
- type______________ ____ .... ____
- result_____________ ____ .... ____

Total questions_____
Total errors_______

Please turn the page

Appendix C
Value of Communication Perception Measures

(test 1)

The information you provide about your partner will be confidential

1	None	All the time	Amount of communication
2	None could have done it without	Essential Could not have done it without	Value of communication
3	Pilot 100%	Advisor 100%	Instigator of communication
4	Very poor	Very good	Advisor's performance
5	Very poor	Very good	Pilot's performance
6	Very low	Very high	Your partner's SA
7	Not confident	Very confident	Your confidence in accuaracy of your response to question 6

Please turn the page
in preparation for the
next test

Appendix D
Performance Measures

DATA SHEET ViP2

PILOT

Date________________ Time (real)________________

Trial number ________________

Experimenter__________________

Subject 1 (pilot)_______________

Subject 2 (advisor)______________

1) Trial start time (from F22 clock)_____________

Mission description

RESULTS

Result____________
BDA_____________
Effectiveness_______
Enemy a/c_________

2) Mission end time ___________
(from F22 clock)

Mission duration (2-1)_________

How ended (crash, shot down etc)

Appendix E
Videotape Coding Form

GROUP CODE

Please rate each team on the following items based upon your observation of the videotapes.

1. Please rate the degree to which members *remembered different* aspects of the task (e.g. location of bogies, location of wingmen, weapons selected, effectiveness of weapons, etc.)

1	2	3	4	5	6	7
both remembering the *same* things						each remembering *different* things

2. Please rate the degree to which *individual* team members were responsible for (or focused on) different tasks within the mission. (e.g. avoiding ground threats, engaging the enemy, specific maneuvers, weapons selected, navigation, threat identification, etc.)

1	2	3	4	5	6	7
responsibility for tasks had little *overlap*						responsibility for tasks had a great deal of *overlap*

3. Please rate the degree of *task* coordination (working smoothly) between team members.

1	2	3	4	5	6	7
team members do not coordinate with each other at all						team members coordinate with each other a great deal

4. Please count the incidences of confusion (lack of understanding) in communication between team members
[Note: clarification is not confusion]

____

Appendix E (Continue)

5. Please rate the degree to which team members sequenced their communication effectively and appropriately (focus on communication, verbal sequencing). [Note: stepping on each other]

1	2	3	4	5	6	7
very ineffective sequencing						Very effective sequencing

6. Please rate the degree of cooperation between team members

1	2	3	4	5	6	7
very little cooperation						a great deal of cooperation

7. Please rate the degree of trust (in judgment and ability) between team members

N/A will get from the self-reports

-3	-2	-1	0	1	2	3
do not trust each other at all						trust each other a great deal

8. Please rate the degree the amount of criticism (negative communication, can be tone of voice) between team members

1	2	3	4	5	6	7
Nonexistent						occurred frequently

9. Please rate the degree to which team members behaved (including non-verbals, sarcasm) as if they were frustrated with one another.

1	2	3	4	5	6	7
Nonexistent						occurred frequently

10. Please rate the degree to which team members accepted procedural suggestions from one another.

1	2	3	4	5	6	7
Non-acceptance or heavy resistance						Open acceptance no resistance

Appendix F
Sample Crewmember Form Fairchild

COPILOT'S QUESTIONNAIRE

PLEASE COMPLETE *INDIVIDUALLY!*

***Date of Flight* _____ *Call Sign*_____**

Please circle the most accurate response or fill in the blank

Background

1.	My crew position is	AC	Co	Nav	Boom
2.	What is your approximate total flight time in the Air Force?	_____hrs			
3.	How many hours do you have in the KC-135 (best estimate)	_____hrs			
4.	Were you a Distinguished Graduate (DG) at UPT ?	Y / N			
5.	Were you a Distinguished Graduate (DG) at Altus/Castle ?	Y / N			
6.	How long have you been mission qualified in this crew position?	_____yrs	_____ mos		

Appendix F (Continued)

Please check the most appropriate box

Crew Responsibilities

	Boom	Nav	Co	AC
11. Which crewmember generally takes responsibility for maintaining and updating the weather information?	☐	☐	☐	☐
12. Which crewmember generally fills out the majority of the 781?	☐	☐	☐	☐
13. Which crewmember generally keeps track of mission paperwork (compiling and organizing it) on the ground?	☐	☐	☐	☐
14. Which crewmember generally communicates with Command Post?	☐	☐	☐	☐
15. Which crewmember generally assumes responsibilities for the Custom Forms on an overseas flight?	☐	☐	☐	☐
16. If the crew required an HF phone patch in flight, who would generally take care of it?	☐	☐	☐	☐
17. Which crewmember does the majority of the communication on the Comm 1 UHF radio?	☐	☐	☐	☐
18. Which crewmember does the majority of the communication on the VHF radio?	☐	☐	☐	☐

Please check the most appropriate box

Crew Skills

	Boom	Nav	Co	AC
19. On this crew, the person most "expert" in his or her aircraft *systems knowledge* is the	☐	☐	☐	☐
20. Then, who has the next highest level of *systems knowledge* in their respective *systems* ?	☐	☐	☐	☐
21. Who, in your opinion, has the *most proficiency* in their respective crew position?	☐	☐	☐	☐
22. Then, who has the next highest level of *proficiency* in their respective crew position?	☐	☐	☐	☐

Appendix F (Continued)

Please check the most appropriate box

Crew Strengths	Very limited knowledge	Some Knowledge	Moderate Knowledge	Substantial Knowledge	As well as you can know them
23. How well do you know the strengths of this AC in his/her crew position?	☐	☐	☐	☐	☐
24. How well do you know the strengths of this Nav in his/her crew position?	☐	☐	☐	☐	☐
25. How well do you know the strengths of this Boom in his/her crew position?	☐	☐	☐	☐	☐

Please check the most appropriate box

Crew Strengths cont.	Crew Coordination	Technical Proficiency (e.g. landing, navigating, refueling)	Systems and EP knowledge(of systems they are responsible for)
26. If I had to pick one I would say that this AC's greatest strength as a crewmember is	☐	☐	☐
27. If I had to pick one I would say that this Nav's greatest strength as a crewmember is:	☐	☐	☐
28. If I had to pick one I would say that this Boom's greatest strength as a crewmember is:	☐	☐	☐
29. At the present time, If I had to pick just one I would say that MY greatest strength as a crewmember is:	☐	☐	☐

Appendix F (Continued)

Please check the most appropriate box

Performance	Strongly Agree	Agree	Slightly Agree	Neutral	Slightly Disagree	Disagree	Strongly Disgree
44. I think this crew has the necessary *aviation skills and knowledge* to successfully complete this checkride.	☐	☐	☐	☐	☐	☐	☐

Percent confident	0	10	20	30	40	50	60	70	80	90	100
47. How confident are you that at least one member of your crew will receive an "Outstanding Performance OP" or "Exceptionally Qualified EQ" rating on this checkride?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
49. How confident are you that no member of your crew will receive more than three downgrades on this checkride?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Below Average	Slightly Below Avg	Average	Slightly Above Avg.	Above Average
50. Compared to the average crew, I think this *crew's checkride performance* will be:	☐	☐	☐	☐	☐

Thank you for participating!

Appendix G
Sample Evaluator Form

CHECK NAVIGATOR'S QUESTIONNAIRE

Date of Flight _____ ***Call Sign***_____

Please check the most appropriate box

Crew Evaluation

1.	The following crewmember received checks on this flight.(circle all that apply)	AC	Co	Nav	Boom	
2.	I was giving checks to the (circle all that apply):	AC	Co	Nav	Boom	
3.	*Taken as a crew*, overall this *crew's* checkride performance was: NOTE: it is vital that you be accurate in your assessment. Avoid rating inflation and DO NOT firewall ratings.	Below Average ☐	Slightly Below Avg ☐	Average ☐	Slightly Above Avg. ☐	Above Average ☐

Please check the most appropriate box

Nav Evaluation

4.	*Taken as a individual*, overall this *individual's* checkride performance was: NOTE: it is vital that you be accurate in your assessment. Avoid rating inflation and DO NOT firewall ratings.	Below Average ☐	Slightly Below Avg ☐	Average ☐	Slightly Above Avg. ☐	Above Average ☐
5.	From my observations of the flight and evaluations on the ground, at the present time I rate this evaluatee's greatest strength as	☐ Crew Coordination	☐ Technical Proficiency (e.g. landing, navigating, refueling)	☐ Systems and EP knowledge (of systems they are responsible for)		

Overall Checkride Result: EQ Q Q2 Q3

How many critique items (if any) could have been avoided with better crew coordination? _____

Appendix H
Cover letter

Air Force Institute of Technology (AFIT)
University of Washington
Seattle, Washington

Dear Copilot

The AMC DO has directed me to examine what makes an effective Air Force crew. I am an active duty Air Force Major and KC-135 pilot, currently working on a doctoral degree at the University of Washington. My dissertation focuses on aircrew effectiveness.

I am specifically interested in aircraft with multiple crew positions. The KC-135 is an ideal aircraft for this study. The most efficient way to address this issue is through the use of a short questionnaire. All answers to the survey will be strictly confidential. The Air Force will never receive individual answers. Only group level information will be presented to the Air Force and the University of Washington. Call signs are requested in order to combine crew answers.

Your honest and voluntary answers will help complete my aircrew effectiveness research and will also help the Air Force learn more about aircrew composition issues.

Your cooperation is greatly appreciated.

Sincerely

//signed//
Daryl R. Smith, Major, USAF
School of Business Administration
University of Washington

INSTRUCTIONS

1. Complete the individual survey *for your crew position.*
2. Please fill out the survey *individually*, please do not discuss answers with one another.
3. Place the individual survey back into the large envelope.
4. Once all crewmembers have completed the survey please seal the envelope.
5. *Turn over* the sheet to begin the survey.

Appendix I
Sample Crewmember Form Altus

COPILOT'S QUESTIONNAIRE

PLEASE COMPLETE *INDIVIDUALLY!*

Date of Flight_____Call Sign_____

Please circle the most accurate response or fill in the blank

Background

1.	Reminder: This is the Copilot's questionnaire					
2.	What is your approximate total flight time in the Air Force?	______hrs				
3.	How many hours do you have in the KC-135 (best estimate)	______hrs				
4.	Were you a Distinguished Graduate (DG) at UPT ?	Y / N				
5.	What is your total flight time (military + civilian) ?	______hrs				
6.	The following members of the crew on this checkride are going through qualification training	AC	Co	Nav	Boom	
6a.	Circle the members of the crew receiving checks on this flight.	AC	Co	Nav	Boom	

Please fill in the blank

Experience with the Crew

7.	How many previous times have you flown with the Aircraft Commander assigned to this sortie	______
8.	How many previous times have you flown with the Navigator assigned to this sortie	______
9.	How many previous times have you flown with the Boom Operator assigned to this sortie	______
10.	How many flights has this crew had as an integral crew (i.e. where the AC/CO/Nav/Boom have all been the same as on this flight) [You may consult others on this question]	______

Appendix I (Continued)

Please check the most appropriate box

Crew Responsibilities	Boom	Nav	Co	AC
11. Which crewmember generally takes responsibility for maintaining and updating the weather information?	☐	☐	☐	☐
12.. Which crewmember generally fills out the majority of the 781?	☐	☐	☐	☐
13. Which crewmember generally keeps track of mission paperwork (compiling and organizing it) on the ground?	☐	☐	☐	☐
14. Which crewmember generally communicates with Command Post?	☐	☐	☐	☐
15. Skip	☐	☐	☐	☐
16. If the crew required an HF phone patch in flight, who would generally take care of it?	☐	☐	☐	☐
17. Which crewmember does the majority of the communication on the Comm 1 UHF radio?	☐	☐	☐	☐
18. Which crewmember does the majority of the communication on the VHF radio?	☐	☐	☐	☐

Please check the most appropriate box

Crew Skills	Boom	Nav	Co	AC
19. On this crew, the person most "expert" in his or her aircraft *systems knowledge* is the	☐	☐	☐	☐
20. Then, who has the next highest level of *systems knowledge* in their respective *systems* ?	☐	☐	☐	☐
21. Who, in your opinion, has the *most proficiency* in their respective crew position?	☐	☐	☐	☐
22. Then, who has the next highest level of *proficiency* in their respective crew position?	☐	☐	☐	☐

Appendix I (Continued)

Please check the most appropriate box **Crew Strengths**	Very limited knowledge	Some Knowledge	Moderate Knowledge	Substantial Knowledge	As well as you can know them
23. How well do you know the strengths of this AC in his/her crew position?	☐	☐	☐	☐	☐
24. How well do you know the strengths of this Nav in his/her crew position?	☐	☐	☐	☐	☐
25. How well do you know the strengths of this Boom in his/her crew position?	☐	☐	☐	☐	☐

Please check the most appropriate box **Crew Strengths cont.**	Crew Coordination	Technical Proficiency (e.g. landing, navigating, refueling)	Systems and EP knowledge(of systems they are responsible for)
26. If I had to pick one I would say that this AC's greatest strength as a crewmember is	☐	☐	☐
27. If I had to pick one I would say that this Nav's greatest strength as a crewmember is:	☐	☐	☐
28. If I had to pick one I would say that this Boom's greatest strength as a crewmember is:	☐	☐	☐
29. At the present time, If I had to pick just one I would say that MY greatest strength as a crewmember is:	☐	☐	☐

Please check the most appropriate box **Interpersonal Relations.**	Casual Acquaintance	Some Knowledge	Moderate Knowledge	Substantial Knowledge	Know them very well
30. How well do you know the AC on a personal basis (off the job)?	☐	☐	☐	☐	☐
31. How well do you know the Boom on a personal basis (off the job)?	☐	☐	☐	☐	☐
32. How well do you know the Nav on a personal basis (off the job)?	☐	☐	☐	☐	☐

Appendix I (Continued)

Please check the most appropriate box **Process Assessment**	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
36. Members of this team know each other so well, that there is little need for detailed and specific planning for each phase of flight	☐	☐	☐	☐	☐
37. On this crew there tends to be little confusion between crewmembers while carrying out crew duties	☐	☐	☐	☐	☐
38. As a crew, we effectively sequence our communication (e.g. avoid "stepping on" each other, give the required communication in a timely manner, etc.).	☐	☐	☐	☐	☐
39. I sometimes feel frustrated because I don't know what to expect from one or more member's of this crew	☐	☐	☐	☐	☐
40. Each crewmember on this crew has established his or her expertise and it is rarely challenged by other crewmembers.	☐	☐	☐	☐	☐
41. Compared with the "average" crew (as far as smoothly and efficiently and cooperatively working together to accomplish the mission) this crew is more like the	Best Crew ☐	☐	Avg Crew ☐	☐	Worst Crew ☐
42. Compared to the average crew, the members of this crew on the whole are more or less open to suggestions from other members of the crew.	Totally Open ☐	Slightly Open ☐	Neutral ☐	Slightly Closed ☐	Totally Closed ☐
43. I feel other members of this crew are critical of how I execute my flight duties	Never ☐	Rarely ☐	Sometimes ☐	Frequently ☐	Always ☐

Appendix I (Continued)

Please check the most appropriate box

Performance	Strongly Agree	Agree	Slightly Agree	Neutral	Slightly Disagree	Disagree	Strongly Disgree
44. I think this crew has the necessary *aviation skills* to successfully complete this checkride.	☐	☐	☐	☐	☐	☐	☐
45. I think this crew has the *knowledge* needed to successfully complete this checkride.	☐	☐	☐	☐	☐	☐	☐
46. I think this crew has the *teamwork skills* needed to successfully complete this checkride.	☐	☐	☐	☐	☐	☐	☐

Percent confident	**0**	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**
47. How confident are you that at least one member of your crew will receive an "Outstanding Performance OP" or "Exceptionally Qualified EQ" rating on this checkride?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
48. How confident are you that each member of your crew will receive a QI rating on this checkride?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
49. How confident are you that no member of your crew will receive more than three downgrades on this checkride?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Below Average	Slightly Below Avg	Average	Slightly Above Avg.	Above Average
50. Compared to the average crew, I think this *crew's checkride performance* will be:	☐	☐	☐	☐	☐

VITA

Daryl Raymond Smith

University of Washington
Graduate School of Business
Dept. Of Management and Organization
Box 353200
Seattle, WA 98195-3200

EDUCATION

Ph.D. University of Washington 1999
Major: Human Resource Management and Organizational Behavior
Minor: Research Methods
Minor: Psychology

M.S. Wright State University 1994
Human Factors Engineering

B.S. United States Air Force Academy 1984
Human Factors Engineering