Winter 2013

The Effects of Group Size on Student Learning, Student Contributions, \Mental Effort, and Group Outcomes for Middle-Aged Adults Working in an Ill-Structured Problem-Solving Environment

Gary Lee Roemmich
Old Dominion University

THE EFFECTS OF GROUP SIZE ON STUDENT LEARNING, STUDENT CONTRIBUTIONS, MENTAL EFFORT, AND GROUP OUTCOMES FOR MIDDLE-AGED ADULTS WORKING IN AN ILL-STRUCTURED PROBLEM-SOLVING ENVIRONMENT

By

Gary Lee Roemmich
B.S. June 1976, University of Washington
M.M.A. May 1984, University of Rhode Island

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY
(EDUCATION)

INSTRUCTIONAL DESIGN AND TECHNOLOGY

OLD DOMINION UNIVERSITY
December 2013

Approved by:

Ginger S. Watson (Director)

Gary R. Morrison (Member)

Linda S. Bol (Member)

ABSTRACT

THE EFFECTS OF GROUP SIZE ON STUDENT LEARNING, STUDENT CONTRIBUTIONS, MENTAL EFFORT, AND GROUP OUTCOMES FOR MIDDLE-AGED ADULTS WORKING IN AN ILL-STRUCTURED PROBLEM-SOLVING ENVIRONMENT

Gary Lee Roemmich
Old Dominion University, 2013
Director: Dr. Ginger S. Watson

Group work has become increasingly important within adult education as educators strive to present students with problems and processes that they encounter in their professional lives. In many work environments, individuals are expected to function as a part of a team to solve complex problems. Consequently, there has been a shift towards teaching students how to solve problems as part of a group rather than individually. An important question becomes "What size group maximizes students learning?" This study compared student learning, student participation levels, and mental effort for middle-aged, professional students in large (six students) and small groups (three students) while working in a collaborative, ill-structured problem solving environment to determine if group size impacted student performance. This study found that there was no significant difference in learning, participation, and mental effort between large and small groups. It also confirmed earlier research demonstrating that group product scores, even when adjusted for student participation, did not predict individual student learning. A multiple regression was used to determine if group size, participation, mental effort or group scores could be used to predict individual student learning. The study showed that for middle-aged professional students, group size, mental effort, participation, or group quality were not effective predictors of student learning.

This dissertation is dedicated to my wife, Raeanne Roemmich whose unending support, not only through this process, but my entire professional career has been the linchpin to my success. Without her support, understanding, and inspiration, I would not be where I am today.

ACKNOWLEDGEMENT

I would like to acknowledge all the people who supported me as I have pursued my doctorate. It starts with the faculty at Old Dominion University, who have taught me so much over the past six years: their guidance, support, and mentorship have been instrumental to my arriving at this point today. The faculty members at the Joint Forces Staff College have also been important in my pursuit of knowledge. From the daily discussions and debates in hallways, offices, and classrooms, all have helped form my understanding and knowledge. And of course, their support and understanding in allowing me to conduct research with their student body goes without saying.

Table of Contents

LIST OF TABLES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

If the goal of education is to prepare students to function as members of a community, then knowledge transfer becomes as important as knowledge retention, creating a serious challenge to educators as they prepare students for the uncertain world the future will present them (Bruffee, 1999). Wertheimer (1959) noted that students excel when asked to solve problems of a type they have already learned to solve, but are unable to transfer that knowledge to solve new problems unlike any they have solved in the past. Their ability to transfer what they have learned in the classroom to "real-world" problems affects the degree of success students will enjoy later in life. It can be argued that schools do poorly in preparing students to meet real-world problems and as such, should better teach students how to solve problems in preparation for real-life (Jonassen, 2010). The problems they are taught to solve in traditional educational settings are normally focused within a specific domain and require learners to apply a few rules, heuristics, or principles along a preferred solution path that leads to answers that are clearly identifiable as "right" or "wrong" (Wood, 1983). These problems are often presented as shallow, independent stories that most students are unable to relate to (Jonassen, 2010).

Unfortunately, life beyond school is comprised of problems structured far differently from those encountered in the classroom. This presents a formidable challenge for students attempting to transfer their skills/knowledge learned in the classrooms to real-life problems (Jonassen, 2010). Real world problems typically reside on the opposite end of the problem continuum when compared to those problems learners are taught to solve in school. Real world problems usually require multiple knowledge domains to

successfully resolve. They usually have no simple "correct answer" but involve difficult choices where potential solutions are "better or worse" than alternative solutions with the distinct possibility that in fact there may not be a solution (Kitchener, 1983). These problems present a dilemma that must be addressed if our educational systems are to better prepare students for life after school.

Any study that attempts to examine learning in the context of problem-solving environments must first understand the nature of problems themselves and their associated characteristics. A problem is simply defined as the "fulfillment of a need and the uncertainty as to how to immediately get it" (Newell & Simon, 1972, p. 72). Each problem possesses an area called the problem space, which by definition is the space in which all problem solving activities take place. It is a mental representation of the problem consisting of symbolic structures (states) and a set of operators over the space (Newell & Simon). Problems themselves span a continuum ranging from very simple and straightforward, to highly complex and convoluted. In an attempt to provide more fidelity to this problem continuum, Jonassen (2000) categorized problems into eleven different types with many of the simpler problem types embedded into the more complex types. As an example, decision-making is a mid-range category problem, but is a sub-skill for most of the more complex problem types (strategic performance, policy-analysis, design, and dilemmas). In fact, in solving more complex problems, the problem solver is likely to exercise multiple decision-making cycles (Jonassen, 2010). Each type of problem is defined by a set of characteristics based on external factors. These factors identify the problem's structure, complexity, context, dynamicity, and domain specificity (Jonassen, 2000). This distinction in problem types becomes important because different problem

types require different skill sets to solve (Greeno, 1980; Allaire & Marsiske, 2002; Kitchener, 1983; Hong, Jonassen, & McGee, 2003).

The most important characteristic differentiating problems lies with the continuum of problem structure that ranges from well-structured to ill-structured (Arlin, 1989; Jonassen, 1997, 2000; Newell & Simon, 1972; Voss & Post, 1988; Wood, 1983). At one end of the continuum are well-structured problems represented by problem spaces that are well-understood. They have sufficient states and identified operators within the problem representation resulting in easier problem resolution. Such problems normally require the application of a limited number of rules or principles within a specific domain, in a prescribed manner leading to the correct solution (Wood, 1983). Well-structured problems are what we typically see in education where the problem representation is embedded in an abstract story that has little relevance to the learner and is of questionable value in preparing students to solve problems in the post-educational environment (Jonassen, 2010).

At the opposite end of the problem continuum are ill-structured problems. Ill-structured problems are characterized by having unknowns in the problem space and as the number of unknowns increase, the problem becomes more ill-structured (Frensch & Funke, 1995; Spering, Wagener, & Funke, 2005). Some unknowns may emerge as the problem space is analyzed and solutions are developed but others will remain hidden. Ill-structured problems cross multiple domains and while they have multiple solutions available through multiple solution paths (Kitchner, 1983). There is no simple "right" or "wrong" answer, but when comparing possible solutions, they can be characterized as being better or worse than other solutions. The criteria used to compare potential

solutions are often based on the values and biases of the decision-maker, which in turn influences the solution paths leading to problem resolution (Jonassen, 2010). Unlike well-structured problems, ill-structured problems are embedded in everyday life and as such are susceptible to the problem solvers' belief systems which are directly influenced by external social, cultural, and organizational drivers (Jonassen, 2000; Meacham & Emont, 1989, Smith, 1991).

Complexity is also used to characterize problems and is defined by the size of the problem space, which consists of a number of nodes with normally a high degree of connectivity (Funke, 1991). Changes at a single node often change the status of numerous other nodes making it difficult to anticipate potential changes for a given situation. Problem complexity is determined in large part by the size of the problem space. As the size of the domain and general knowledge required to solve a problem increases, the size of the problem's space increases correspondingly, with a subsequent increase in problem complexity (Kotovsky, Hayes, & Simon, 1985). As the number of activities required for problem resolution increases, or the number of relationships between those activities increases, problem complexity increases as well (Wood, 1985). As problem complexity grows, it becomes more difficult for students to identify the best solution path (Jacobs, Dolmans, Wolfhagen, & Scherpbier, 2003) and it places an increased load on the cognitive resources available to solve the problem (Kirschner, Paas, & Kirschner, 2008). While complexity is related to problem structure, it is sufficiently independent to warrant individual consideration (Jonassen, 2010). Many complex problems are thought to be ill-structured, but it is possible to have highly complex, well-structured problems as well (chess is an example of a highly complex, well-structured problem).

As problems become less structured, problem context becomes more important for students conducting the cognitive activities required to solve the problem (Lave, 1988: Rogoff & Lave, 1984). While context is secondary in well-structured problems, it is so essential to ill-structured problems that it becomes a part of the problem itself, which then becomes part of the desired solution (Wood, 1983). Very ill-structured problems have a tendency to become so context dependent that once removed from that context, the problem itself loses all meaning (Jonassen, 2010), making them similar to the well-structured problems used in many of our schools. Understanding the problem context is critical to success as context helps to determine what skills to apply and when best to apply them. This process requires students to monitor and control their cognitive processes (Harris & Pressley, 1991). The best way to attain this level of expertise is by having students learn problem-solving within the context of realistic problem-solving situations (Mayer, 1998). A number of methodologies including anchored instruction (Bransford, Sherwood, Hasselbring, Kinzer, & Williams, 1990), guided design (Wales & Stager, 1977), and problem-based learning (Barrow, 1983) have been developed to teach problem solving in real-world contexts.

While schools focus on teaching students to solve problems individually, the real world often expects them to solve problems collaboratively as part of a team. When students are presented with simple problem solving and learning tasks such as recalling items, research has shown students learn more when using individual instructional methodologies rather than group methodologies (Andersson & Rönnberg, 1995; Meudell, Hitch, & Kirby, 1992; Weldon & Bellinger, 1997). However, research also indicates that for students working on relatively complex problem solving tasks in group environments,

their individual schema acquisition exceeds that of individuals (Kirschner, Paas, & Kirschner, 2009a; Laughlin, Bonner, & Miner, 2002; Laughlin, Hatch, Silver, & Boh, 2006). Collaborative teams develop better, deeper solutions that address issues from multiple perspectives (Schwartz 1995). They outperform individuals in problem solving (Laughlin, Carey, & Kerr, 2008) while achieving greater educational efficiency by having higher levels of learning with lower mental effort among individual group members (Kirschner et al., 2008; Kirschner et al., 2009a; Kirschner, Paas, Kirschner, & Janssen, 2011).

There is agreement in the literature regarding the success of collaborative groups when solving complex, ill-structured problems. However, is there a point where the group becomes too large and loses its effectiveness in promoting schema acquisition? The focus of this study will be to determine if the size of a group influences student learning in solving complex, ill-structured problems.

Problem Solving Skill Sets

Problem characteristics are an important factor in successful problem-solving as different skill sets are required to solve different problem types (Kitchener, 1983; Schraw, Dunkle, & Bendixen, 1995; Hong, Jonassen, & McGee, 2003; Allaire and Mariske, 2002). Kitchener (1983) developed a model that identifies the three basic skills required in problem solving: inferential rules and strategies normally associated with the solving of problems (level "1" skills); meta-cognitive skills (level "2") were used to select and monitor level "1" skills; and the epistemic skills problem solvers use to develop assumptions concerning the limits and certainty of knowledge (level "3"). Kitchener theorized that level "1" & "2" skills are used for well-structured problem

solving whereas successful, ill-structured problem solving required the use all three levels. Schraw et al. (1995) expanded upon Kitchener's model and found that performance on well-structured and ill-structured problems is independent; success in solving well-structured problems had no impact on a person's ability to solve ill-structured problems. Epistemic beliefs, which are essential to solving ill-structured problems, were not correlated with well-structured problem solving. Allaire and Marsiske's (2002) research supported Kitchner's earlier model as they found measures that predict well-structured problem solving in elderly people could not predict the quality of their solutions to ill-structured problems. Hong, Jonassen, and McGee (2003) found ill-structured problems in an astronomy simulation called on different skills than well-structured problems, including meta-cognition and argumentation. Thus, it is not surprising that studies using well-defined problem-solving rarely showed transfer between domains (Chi, Glaser, & Farr, 1988) while studies using ill-defined problems frequently do (Kitchener & King, 1981; Kuhn & Frank, 1991; Perkins, Farady, & Bushey, 1991). Cho and Jonassen (2002) found that students solving ill-structured economics problems also produced more extensive arguments than did students solving well-structured problems. They theorized this was the result of the importance of generating and supporting alternative solutions when solving ill-structured problems. This reliance on argumentation produced communication patterns within teams that differed for teams solving well-structured and ill-structured problems (Jonassen & Kwon, 2001). The generation of alternative solutions and subsequent analysis of potential solutions is best approached from a group problem-solving environment.

Group Problem Solving

Research shows that groups outperform individuals for complex problems solving tasks (Laughlin, Carey, & Kerr, 2008 Laughlin et al., 2006; Ohtsubo, 2005) and, they significantly outperform the best students working individually (Heller, Keith, & Anderson, 1992). Individuals who had previous group experience consistently outperformed individuals with no previous group experience. A single group experience was alone sufficient for the group-individual transfer effect to occur, but some of the benefits of prior group experience may take time to fully emerge. Groups outperformed those individuals who had previous group experience (Laughlin et al., 2008).

The resources required to solve complex problems usually require more resources than an individual can provide, forcing the necessity to use groups. Functioning as a group allows for prompt evaluation, explanation, or reflection; supports the planning and execution of complex tasks; and contributes new perspectives, novel information, or skills that other students may not possess (Wiley & Jensen, 2006). Group members may detect errors and provide immediate feedback from perspectives that individuals may not possess (Schoenfeld, 1989). Working as a group also allows for the development of more reflective, explicit, or abstract mental models than lone individuals can develop (Moreland & Levine, 1992; Schwartz, 1995).

Group size impacts on learning.

Webb and Palincsar (1996) in their review of earlier work on cooperative learning found that most researchers recommended groups of three to four students to optimize learning. They compared dyads and triads and found that triads outperformed dyads, as well as individuals, to include even the highest ability individuals in a nominal group

analysis. Triads solved problems in fewer trials, were more effective in their equation selection, were better at monitoring the redundancy of their solution attempts, and were better able to take advantage of group heterogeneity on this task (Wiley & Jensen, 2006). They observed that for dyads, both members had to have math skills for the group to be more effective. In triads, as long as one member had math skills, the group was more effective. This finding was similar to those of Laughlin and Adamopoulos (1980) who found that in groups of six, there must be at least two proficient students for group success. Larger groups provide a larger variety of viewpoints, which may challenge superficial viewpoints. Considerable individual learning during group interactions and collaborative decision making may have had a major influence on subsequent individual responses. The disparity between groups and individuals in recognizing correct answers, rejecting erroneous answers, and implementing effective collective information processing strategies grows as the task becomes more intellectual.

Some research found that pairs performed differently than did larger groups (Wiley & Jensen, 2006). Pairs were preferred because such small groupings allowed students to maximize their engagement with the material, increasing their opportunity to participate in group activities and discussions (Lohman & Finkelstein, 2000). With fewer group members, there were fewer distractions, allowing students to remain focused on the content (Dugosh, Paulus, Roland, & Yang, 2000). Pairs greatly reduced the likelihood that students would "free ride" or "socially loaf", with students more likely to feel highly invested in the product or activity (Stroebe & Diehl, 1994).

Apedoe, Ellefson, and Schunn (2011) in their study on design-based instruction for high school students in science classrooms, found there was a difference in group size

effects between advanced classes and mainstream classes. In advanced classes, students appeared to have difficulty sharing intellectual workload in large groups but were successful in pairs. Apedoe et al. theorized advanced students might have preferred to work alone rather than attempting to split intellectual workload among members of a large group. This preference may have resulted in lower student motivation which was displayed through social loafing tendencies (Karau & Williams, 1993). Such behaviors would eventually result in poor group performance. Mainstream classes performed better in larger groups and Apedoe et al. theorized perhaps they were better able to distribute the intellectual workload more effectively, benefitting from the greater diversity of viewpoints.

Laughlin et al. (2006) also found evidence contradicting the earlier work on pairs. They compared three, four, and five person groups with pairs. Their study showed the larger groups required fewer trials to reach a solution and proposed more complex equations than either pairs or individuals. They found no significant difference between the groups of three, four, or five. Their study found no significant difference between the pairs and the best individual students from groupings of five students working independently. The pairs did require fewer trials to solution than did the second-best individuals from the independent grouping of five. This finding supported Laughlin, Zander, Knievel, and Tan's (2003) earlier work where they compared three-person groups to individuals. Their work found that groups of three students outperformed the best of three independent individuals making the use of three-person groups a more efficient use of human and logistic resources.

Lou, et al. (1996) conducted a meta-analysis of cooperative learning studies comparing grouped and ungrouped students and found that students in small groups achieved significantly more than did students working in ungrouped classrooms. They also found that group size was significantly related to learning outcomes. Paired students learned significantly more than did ungrouped students with groups of three to four students being the optimal size for learning. They also reported that when the size of the groups ranged between five and ten students, those students in groups did not learn more than did students in ungrouped classes.

Lyons, McIntosh, and Kysilka (2003) argued that for more complex problems, groups of four to seven students were ideal as it allowed sufficient viewpoints to stimulate active discussions, allowing students the opportunity to think beyond a simple affirmation of their beliefs. However, when groups numbers exceeded seven students, introverted students were found to be less likely to engage in the group discussions. Student participation in group discussions is a critical factor contributing to student achievement when facing complex problems. Group interactions require that students defend their problem-solving strategies while projecting potential impacts. This facilitates student development of more effective solution strategies that persist with students for subsequent problem-solving iterations. Students who had previously adopted ineffective solutions, shifted to more effective solutions following group discussions (Cooper, Cox, Nammouz, & Case, 2008).

Group member composition impacts on learning.

While group size appears to influence student learning, it is not the only group characteristic that does so. The composition of the group members themselves also plays

a role in affecting the effectiveness of the group. A major contributor to overall group success is the composition of the group, particularly with regards to ability of group members. Students with a specific ability level may interact differently in groups with varying student compositions (Webb, 1991). In studies where groups were heterogeneous with students of high, moderate, and low ability, researchers found the high and low-ability students formed teacher-student relationships with significant interactions while moderate ability students interacted at much reduced levels (Peterson, Janicki, & Swing, 1981; Webb, 1980, 1982; Webb & Kenderski, 1984). When moderate ability students performed in groups consisting of only moderate ability students, they provided (Webb, 1980, Webb & Kenderski, 1984) and received (Webb, 1980, 1982) more explanations, demonstrating a higher level of achievement. In most empirical studies, high-ability students have performed equally well on achievement tests after working in heterogeneous and homogeneous groups (Azmitia, 1988; Hooper & Hannafin, 1988; Hooper, Ward, Hannafin, & Clark, 1989; Skon, Johnson, & Johnson, 1981).

When the heterogeneous groups consisted of only high and moderate or moderate and low ability students, researchers observed very different behaviors. Those groups saw significant interactions between students of both ability groups with no grouping being left out of the interactions. Every student in these groups interacted with other group participants at higher levels than did those students in mixed-ability groups consisting of high, moderate, and low-ability students (Webb, 1982, 1984; Webb & Cullian, 1983). Moderate-ability students in these groupings provided and received more explanations and demonstrated higher achievement levels than did comparable moderate-ability

students in heterogeneous groups consisting of students at all three ability levels (Webb & Kenderski, 1984).

A common complaint used against grouping is that high-ability students may be "held back" by being forced into groups with students of lesser ability (Oakes, 1990). However, research does not appear to support these suppositions as high-ability students may appear to learn more when grouped with students of lesser abilities than when grouped only with students of high ability (Webb, 1980). When grouped with students of lesser ability, high-ability students often assumed a teaching role and explained the material to other group members, typically the low-ability students (Bereiter & Scardamalia, 1989; Webb, 1991). When high-ability students were grouped only with other high-ability students, they assumed (often incorrectly) that everyone knew how to solve the problems and therefore, felt no need to provide explanations to other group members as they believed that everyone was competent enough to master the material without assistance. Lou et al. (1996) found that high-ability students performed well regardless of the group composition. Collins and Onwuegbuzie (2000) also found that high-ability groups performed at higher levels and produced a higher quality group product than did lower ability level groups in a college level educational research methodology course. They also found that heterogeneous groups performed better than the more homogeneous groups.

Not surprisingly, low-ability students working on novel mathematics problems received more explanations when grouped with students of greater ability than when grouped only with low-ability students (Webb, 1980). In homogeneous groups, low-ability students tended not to exchange correct explanations. This was most likely due to

their lack of the specific skills/knowledge as well as their perceived lack of confidence. Azmitia (1988) reported novice children paired with expert children learned how to build Lego models better than novices paired with other novices. His research supported earlier research where novice children paired with expert children also received more explanations and demonstrations than novices paired with other novices (Hooper & Hannafin, 1988; Webb, 1980). Tudge's (1989) study of children working with a mathematical balance beam found that when lower-ability students were paired with higher ability students, the low-ability students were likely to be exposed to reasoning at higher levels than they could generate on their own. However, Hooper et al. (1989) found that there were no significant differences in the achievement of low-ability students when grouped with students of higher ability and when grouped only with low-ability students

Early research by Mugny and Doise (1978) found different interaction patterns for heterogeneous groups of all ability levels. Their study showed low-ability students making significant gains when they worked with moderate-ability students, but not when they worked with high-ability students. Moderate-ability students allowed low-ability students to verbalize their problem-solving strategies, voice their doubts and perceived difficulties by allowing the low-ability students to participate in all aspects of the collaboration. The high-ability students appeared to dominate the group work and rarely took time to explain their solutions to the low-ability students. The low-ability students were unable to learn from their interactions with the high-ability students. When high-ability students did provide explanations, they usually were far beyond the ability of most low-ability students to understand. This did not appear to be the case when receiving explanations from the moderate-ability students. Mugny and Doise's (1978) research

findings show that for low-ability students, their participation in the collaborative effort is critical for success and that while membership within heterogeneous groups provides opportunities for assistance and shared collaboration, it does not guarantee them.

Learning in groups.

Multiple studies have demonstrated that learning is enhanced when teachers place students into group environments using cooperative and collaborative learning methodologies (Bossert (1988); Johnson, Maruyama, Johnson, Nelson, & Skon, 1981; Johnson, Johnson, Stanne, & Garibaldi, 1990; Slavin, 1995). However, the type of problem plays a large role in determining if group methodologies will be effectives. For relatively complex problem solving tasks, research indicates group learning is superior to individual learning (Kirschner et al., 2009a; Laughlin, Bonner, & Miner, 2002; Laughlin, Hatch, Silver, & Boh, 2006), while individual learning is superior to group learning for relatively simple recall tasks (e.g. Andersson & Rönnberg, 1995; Meudell, Hitch, & Kirby, 1992; Weldon & Bellinger, 1997).

Early research on learning in groups provided mixed results. Some studies showed no significant difference between individuals and those who participated in groups (Cohen, 1994). A closer examination showed that for many of those earlier studies, researchers failed to understand the significance that problem type plays on group dynamics. Most of the studies used well-structured problems that individual students were able to solve problems without the assistance of group members. The assigned tasks did not require the resources of a group to solve (Cohen & Archevala-Vargas, (1987) and thus, it was easier for individuals to solve the problems without the additional requirement for interactions with group members. The key element that arose

from those early studies was that not all tasks are appropriate for group work. When the working memory of an individual is sufficient to process all elements of a problem, the additional mental effort required for group interactions detracts from schema construction (Kirschner et al., 2008). It is only when the resources required to solve the problem exceeded that available to individual students that the additional cost of group interactions contribute rather than detract from student learning.

Considerable individual learning occurs during group interactions as well which indicates that cooperative decision-making groups may have major influence on subsequent individual responses (Laughlin & Adamopoulos, 1980). Group influence changes the way individuals make judgments rather than merely changing their opinion to adhere to the group norm. The group influence also allows the student to generalize the information beyond the specific issues and items discussed by the group (Stasson & Hawkes, 1995.

Transactive memory systems.

A key element of group dynamics is the collective knowledge or memory system of the group. This collective memory system is referred to as the transactive memory system and is defined simply as the set of individual memory systems within the group combined with the communications that takes place between individuals (Wegner, 1978). The use of the transactive memory system is reliant on the communication pathways within the group as well as the groups' combined domain knowledge. For new groups, an initial period is often required for the development of the group's transactive memory system. Hollingshead (1998a, 1998b, 1998c) developed a series of propositions that describe the development of the transactive memory system (Table 1). These propositions focus on

how the group determines what each person knows (propositions 1-4), their level of expertise (proposition 5), how the group will search for information that the group collectively lacks (propositions 6-8), and how group members communicate knowledge with non-verbal and paralinguistic cues (proposition 9). How quickly the group develops their transactive memory system impact overall group performance as time spent in memory system development usually results in time where the group operates at lower levels of effectiveness. This impact can be mitigated by having the group participate in a team-building exercise prior to working the group task. This should ensure that groups have already conducted the initial group formulation stages and are prepared to operate at higher efficiency levels when ultimately presented with the planning task (Liang, Moreland, & Argote, 1995). The transactive memory system allows for distribution of mental effort throughout the group resulting in reduced mental effort for individuals (Kirschener et al. 2009b).

Table 1

Propositions for the Development of Transactive Memory Systems (Hollingsworth, 1998a)

Number	Proposition
1	Groups begin their discussions by comparing answers or preferences.
2	If all group members retrieve the same answers, it is likely the group will assume the answer is accurate with little or no discussion regarding whether it is actually accurate.
3	If group members retrieve different answers, or if only one member or a subset of group members retrieve an answer, the group will need to assess its accuracy. The group must first establish and recognize expertise.
4	If no group members are able to retrieve the answer, the group will develop and execute a shared information search strategy. The group must establish and recognize expertise before executing the shared strategy.
5	If relative expertise is established and the recognized expert has an answer, the group will adopt the answer. Otherwise, the group will implement an information search.

6	When group members are unable to remember shared knowledge, they will try to cue one another by engaging in transactive information search.
7	If the group has not reached consensus regarding the correct answer, members will focus their attention on cueing the recognized experts in an information search.
8	If the group agrees that the correct answer is outside the collective knowledge of all members, the group will generate a guess or will look to information sources outside the group.
9	Nonverbal and paralinguistic cues are important in the retrieval and communication of knowledge in transactive memory systems.

Assessing group products.

For students working in collaborative groups, completion of the group product itself is the driver for forcing interactions among group members during collaboration and is the instrument most often used to measure the overall productivity of the group. The group product provides an indicator of the quality of the effort in terms of subject material and the relative effectiveness of the collaborative efforts of group members. Quite often, the same group grade is assigned to every group member, regardless of the contributions made to the overall group effort by that individual (Race, 2001; Zhang & Ohland, 2009). However, assigning all group members the same score, regardless of how much each member has contributed, seems unfair and thoughtless, does little to motivate most students (Kagan, 1995), and provides opportunities for those students displaying social loafing and free-riding behaviors to be rewarded for the efforts of their teammates. For this reason, many good students abhor group projects, as they feel that their grades are solely reliant upon their own efforts with few additional contributions from other group members (Oakley, Felder, Brent, & Elhajj, 2004; Pfaff & Huddleston, 2003). This lack of accountability for individual contributions has led some to argue against the use of group grades in schools as the group product score by itself is a poor indicator of overall student effort (Kagan, 1995).

Those scores from group product assessment by themselves may also not be a valid indicator of the individual competency for many students (Webb, 1993). Although a student may receive a high group product score, that score by itself may not reflect actual student competency without an additional assessment of their efforts during group collaboration as well (Webb). At higher cognitive levels, an assessment of their group interactions in addition to performance is required to effectively evaluate a student's ability (Mathematical Sciences Education Board, National Research Council, 1993). How to measure that effort and reflect such efforts in the overall grade has become the major concern when evaluating group work (Gillies & Ashman 2003). It is only by combining the group product score with the assessment of student interactions that an honest assessment of a student's true knowledge and skills can be developed (Webb, 1997).

Measuring student contributions to the group effort.

Teachers usually only have the final group product, which represents the collective effort of all group members, and their own observations of student efforts which are incomplete (Zhang, Johnston, & Kilic, 2008). As a result, various methods have been adopted to awarding marks which reflect not only the outcome of the group project but include considerations for the contributions of the individual group members (Conway, Kember, Sivan, & Wu, 1993). The challenge is in how teachers account for the differences in the contributions of the individual group members (Gilles & Ashman, 2003).

Most methods for measurement of student contributions to group work are reliant on peer and self-rating of student contributions (Conway et al., 1993; Lejk & Wyvill, 1996, 2001, 2002; Johnston & Miles, 2004). While teachers may not be fully aware of the

individual student contributions, students understand how much each individual has contributed or detracted from the overall success or failure of their group. These self- and peer ratings alone readily provide the vital source of information available to effectively gauge individual contributions (Zhang et al., 2008). While many have questioned whether students can accurately assess other students, studies in higher education have demonstrated a high degree of agreement between student and teacher ratings in various disciplines (Falchikov & Boud, 1989; Falchikov & Goldfinch, 2000).

It is difficult for the teacher to accurately assess the contributions of individual students involved in a team project as they usually have multiple groups working simultaneously with significant work occurring outside of scheduled classroom times (Ohland et al., 2012). Students, as inside members of a group, are better able to observe and evaluate members' contributions than are teachers, who are outsiders to the group (Millis, Cottell, & American Council on Education, 1997).To address this issue, group members are often asked to provide self and peer evaluations as a means of assessing an individual's contribution to the overall group effort (Johnston & Miles, 2004).

Peer-assessment has been shown to promote independent, reflective, critical learning (Somervell, 1993), enhances motivation for participation amongst students (Michaelsen, 1992), and encourages students to take responsibility for their learning (Rafiq & Fullerton, 1996). Self-appraisals are often used with peer evaluations because ratees want to have input in their evaluations while the information they provide can facilitate a discussion about their performance (Inderrieden, Allen, & Keaveny, 2004). Self- and peer evaluations may also be used to provide feedback to improve students' team skills and develop reflective and self-management skills that enable students to

become lifelong learners (Chen, Donahue, & Klimoski, 2004; Dochy, Segers, & Sluijsmans, 1999; Felder & Brent, 2007; Young & Henquinet, 2000). Peer-assessment feedback also allows students to assess their own communication skills in group settings. Student understanding that their individual contributions to the group project were to be assessed by their peers, decreased the probability that students would demonstrate free-riding and social loafing behaviors with a subsequent increase in student participation for the group-based learning projects (Johnston & Miles, 2004).

Research also shows a high degree of agreement between student and teacher ratings in various disciplines of higher education (Falchikov & Boud, 1989; Falchikov & Goldfinch, 2000). However, most research pertaining to the quality of peer and self-rating has been devoted to the agreement between student and teacher ratings and not on the consistency of ratings among students themselves (Zhang et al., 2008). This is due in large part because each student only receives ratings from other members of their group. Such small sample sizes make it difficult to calculate reliability using traditional means (Zhang et al., 2008).

While fairly effective, peer evaluation systems are not without their own set of problems. Stronger students have a tendency to under-represent their contributions while weaker students tend to over-represent their contributions (Zhang & Ohland, 2009). Often, average and below-average performers display a tendency to not differentiate their ratings of team members because they worry that providing accurate ratings may damage social relations in the team (Saavedra & Kwun, 1993). When conducted in secrecy, students are "more discriminating" but the use of the peer survey as a formative instrument to assess student contributions is greatly reduced. Group size was also found

to impact self-ratings. Students were found to rate themselves lower if they were in groups of 5 rather than groups of 3-4 (Johnston & Miles, 2004).

Past research has found that both students and employees in organizations are reluctant to evaluate their peers, particularly for administrative purposes, such as for adjusting grades or determining merit-based pay (Bettenhausen & Fedor, 1997; Sheppard et al., 2004). Many students feel peer evaluations are biased by friendships, popularity, jealousy, or revenge (Ohland et al., 2012) and recent research suggests that these concerns may be well-founded (Taggar & Brown, 2006). Peer ratings have been found to impact how students feel about other group members. Students who received positive peer ratings felt better about their teammates and rated them higher on subsequent peer evaluations; whereas students who received negative peer ratings grew to dislike their teammates and subsequently gave them lower ratings on future peer evaluations. This occurred even though the students only saw aggregated feedback from multiple raters and not individual feedback.

Such evaluations normally adopt one of two methods: a point distribution system or a standard system. The point distribution system is based on students dividing up a set number of points for group members. These systems are common because they are simple and yield a score that can easily be used for grading. Unfortunately such systems do not focus on the important teamwork behaviors that are most important, and fails to provide specific feedback for individual improvement (Ohland et al., 2012). The standard system is focused on desired behaviors and provides criteria for scores (a rubric). Unlike the point distribution system, this system treats the contributions of each group member individually against the standards rather than against other group members. Research has

shown that individual accountability (task visibility) (George, 1992; Harkins & Szymanski, 1989; Williams, Harkins, & Latane, 1981), using reliable performance evaluation instruments that are based on clear standards (Harkins & Szymanski, 1989; Szymanski & Harkins, 1987), is a major factor in reducing social loafing and free-riders. The standard system is focused on holding individuals accountable towards established standards.

Cognitive Load

Cognitive load theory (Kirschner, 2002; Paas, Renkl, & Sweller, 2003; Sweller, 1988; Sweller, Van Merriënboer, & Paas, 1998) explains how learning is impacted by our understanding of the interaction between long-term memory and working memory that is limited in capacity (Miller, 1956; Cowan, 2001) and duration (Peterson & Peterson, 1959). The theory asserts that learning is hampered when working memory capacity is exceeded in a learning task. Early cognitive load theory identified three different types of contributions to total cognitive load. Intrinsic cognitive load related to inherent characteristics of the content to be learned, extraneous cognitive load is the load that is caused by the instructional material used to present the content, and finally, germane cognitive load referred to the load imposed by learning processes (Chandler & Sweller, 1991; Sweller, 1988; Sweller & Chandler, 1991). However, recent literature now no longer considers germane cognitive load as a category; dropping the categories to just intrinsic and extraneous. Mental resources must be applied against both categories of mental load. The resources applied against the intrinsic load are considered to be germane to learning and are referred to as germane resources (Sweller, Ayres, & Kalyuga, 2011).

Cognitive load theory has been focused on individuals but recent research has examined how collaborative learning can be used to allow for group resources to compensate for individual working memory limitations (Kirschner, Paas, Kirschner, & Jannsen, 2011). Individuals collaborating on complex learning tasks can divide the task and its associated interactive elements (intrinsic cognitive load) among group members with the resultant lowering of the cognitive load of individual group members (Kirschner, Paas, & Kirschner, 2009b; Ohtsubo, 2005; Stasser, Stewart, & Wittenbaum, 1995). This process is referred to as the distribution advantage (Kirschner et al., 2011). It is countered by the transactional costs for groups. Transactional activities take into consideration the cognitive load cost of communications and interactions between group members (Ciborra & Olson, 1988; Kirschner et al., 2009a; Yamane, 1996). The combined effects of the distribution advantage and transactional activities allows group members to make use of the group's transactive memory system by sharing the cognitive load imposed by a task, allowing members to process information elements more deeply while constructing higher quality schemas than learners working individually (Kirschner et al., 2009a).

Research has shown that for fairly simple recall tasks, groups performed worse than individuals, indicating that collaboration was detrimental to learning (Meudell et al., 1992; Stephenson, Clark, & Wade, 1986; Weldon & Bellinger, 1997). For those low-complexity tasks, individual learners had sufficient working memory capacity to carry out the tasks alone. Dividing the information among the group members forced students to communicate information and coordinate their actions, which for low-complexity tasks imposes a relatively high load (in relation to the benefits that will be accrued) thereby negating the distribution advantage (Kirschner et al., 2009a). Transaction costs drove the

efficiency of the learning for low complexity learning tasks. If transaction costs could be minimized then the learning process for groups would be as efficient as that for individuals but if the transaction costs are high, then any distribution effect is negated and the groups can be expected to learn less efficiently than individuals (Kirschner et al., 2011). For tasks of medium complexity, research has found that group interactions became necessary to successfully resolve the problem resulting in groups that performed significantly better than individuals (Kirschner et al., 2009a; Laughlin & Adamopoulos, 1980; Laughlin, Hatch, Silver, & Boh, 2006; Laughlin, Zander, Knievel, & Tan, 2003; Laughlin, Bonner, & Miner, 2002).

Measuring cognitive load.

Measuring cognitive load is somewhat challenging. It has been argued there are no standard, reliable, and valid measures for the main constructs of cognitive load (Moreno, 2010). The method used most frequently is a self-reported, one-item scale, developed by Paas (1992), in which learners indicate their "perceived level of mental effort" on a 9-point rating scale. His scale asked students to rate their mental effort ranging from very, very low to very, very high. This scale is based on the people being capable of providing a numerical indication of their perceived mental burden (Gopher & Braune, 1984). Such scales are sensitive to small differences in cognitive load and are viewed as being valid, reliable, and not intrusive (Paas, Van Merrienboer, & Adam, 1994). Paas (1992) originally described mental effort as representing an index of cognitive load but later redefined it being that part of cognitive load allocated to address the requirements of the task and as such indicates the actual cognitive load (Paas, Tuovinen, Tabbers, & van Gerven, 2003).

While the concept of a single-item scale has proven to be popular in research, there has been significant variation as to how many points and just what the students were specifically rating on that single item (de Jong, 2010). Studies varied from as few as five points (Camp, Paas, Rikers, & Van Merriënboer, 2001; Huk, 2006; Salden, Paas, Broers, & Van Merriënboer, 2004) to as many as one hundred (Gerjets, Scheiter, & Catrambone, 2006). Studies also varied on when and how often students were asked to complete the survey. Some surveys asked the same question once (Ayres, 2006; Kalyuga, Chandler, & Sweller, 1999; Pociask & Morrison, 2008) while others queried students multiple times during and after the learning process (Kester, Kirschner, & Van Merriënboer, 2006; Paas, Van Gerven, & Wouters, 2007; Stark, Mandl, Gruber, & Renkl, 2002; Tabbers, Martens, & Van Merriënboer, 2004; Van Gog & Paas, 2008; Van Merriënboer, Schuurman, De Croock, & Paas, 2002). Those researchers that took measurements of cognitive load throughout their studies argued that this provided a more accurate measurement as it accounted for variations of the cognitive load experienced at different stages of the learning process (Paas et al. 2003). While most studies examined cognitive load (Ayres, 2006) other studies combined difficulty and understanding into a single question (Pollock, Chandler, & Sweller, 2002). Some researchers queried effort and difficulty separately but then combined the results into a single metric (Moreno & Mayer, 2007; Zheng, McAlack, Wilmes, Kohler-Evans, & Williamson, 2009).

Problems with cognitive load.

Cognitive load measures are relative in that the measures do not show overload; just a relative load. Overload is an absolute measure and thus one cannot predict when overload will or will not occur (de Jong, 2010). Cognitive load ratings provide little help

in interpreting results in terms of cognitive load. Studies that measure only one overall concept of cognitive load do not do justice to its multidimensional character (Ayres, 2006). When measured as a single concept, researchers are unable to differentiate between intrinsic and extraneous cognitive load (de Jong, 2010). The cognitive load measure used does not differentiate between the cognitive load due to perceived difficulty of the subject matter, presentation of the instructional material, or being engaged in relevant learning activities. (Wouters, Paas, & Van Merriënboer, 2009). The most commonly used measures are not sensitive to variations over time and rarely do cognitive load studies take into account the time spent on task when rating cognitive load (Paas et al., 2003).

Purpose of the Study

Research has clearly shown that groups outperform individuals when presented with complex, ill-structured problems (Laughlin & Adamopoulos, 1980; Laughlin et al., 2006). They develop deeper schemata (Kirschner et al., 2009a; Laughlin et al., 2002; Laughlin et al., 2006; Webb & Palincsar, 1996) while requiring less mental effort, making the instruction more efficient (Kirschner et al., 2009b; Ohtsubo, 2005; Stasser et al., 1995). What is not understood, is how group size affects the learning and mental efforts of individuals within the group. Is there a specific size where groups lose their effectiveness; where the dispersal of efforts becomes too diluted and a reduction of learning is experienced when compared to smaller groups? The purpose of this experimental study was to determine if the number of group members influences learning in complex, collaborative, ill-structured problem solving learning environments for mid-grade military officers. The independent variable was defined as the size of the group

itself. Outcome variables were individual student learning, peer assessment of student contributions, individual mental effort for group members and the normalized group product scores. Five research questions guided this study.

Research Questions

1. Is there a difference in higher order learning outcomes for students solving ill-structured problems in small and large groups as measured by a student post-test?
2. Is there a difference in the level of student contributions for students in large and small groups as measured by peer assessments of participation?
3. Is there a difference in the cognitive load of students collaboratively solving ill-structured problems in large and small groups as measured by the mental effort instrument?
4. How well do normalized group scores predict individual learning?
5. Which combination of factors (i.e., group size, student contributions, individual cognitive load, and normalized group product scores) best predict individual student learning in collaborative ill-structured problem-solving environments?

CHAPTER 2

METHOD

Participants

There were 107 students from six seminars of the Joint and Combined Warfighting School (JCWS) at the Joint Forces Staff College (JFSC) that comprised the 26 groups in the study. Of those, 84 students consented to participate and completed associated consent forms (see Appendix A). The JCWS uses a seminar format where students are permanently assigned to a seminar upon arrival. At the time of the study, students had been together for five weeks. Each military student was a mid-grade officer in the United States armed forces and had between 10 to 18 years of military service. There were two government agency civilians among the study participants. Neither civilian student had any previous military experience. All students enrolled in this course were college graduates with approximately 80% of the graduates possessing a master's degree and about 3% holding doctoral degrees.

Research Design

This quasi-experimental mixed-methods study compared the effects of large (6 students) and small (3 students) groups on student learning, student contributions, and cognitive load in an ill-structured problem-solving environment. For the study, each student group completed an analysis of a complex problem. The specific problem analyzed was an operational level, fictional military planning problem based on difficulties within a geographic region. The analysis required each group to determine the current conditions, identify the desired end states, and develop a rudimentary design framework for problem solutions they would address in a later course (outside the scope

of the study). Groups needed to evaluate their work (current conditions, future end states, and design framework) to ensure that their work was inter-connected and that it addressed the specific tasks that initiated the problem analysis. As future staff officers, they must be able to evaluate the work of others as well as their own. The study session covered 12 hours spread over a three day span. Qualitative analysis using a complementary methodology (Greene, Caracelli, & Graham, 1989) was conducted to elaborate on student perceptions as to their functioning as a group.

Treatment.

The two treatments compared in this study were small (3 students) versus large (6 students) groups. Table 2 documents how students were distributed among the groups within each seminar. While 107 students were involved in the group activities, only 84 students opted to participate in the study. Assignment of students to groups was random within branch of service, ensuring that each group had representation from the Army, Air Force, and Naval Services that facilitated later attainment of the overall curriculum objective of "joint" acculturation.

Table 2

Distribution of Students among Seminars

	Large Groups			Small Groups		
Seminars	Groups	Total Students	Participating Students	Groups	Total/Part. Students	Participating Students
1	2	14*	12	1	3	1
2	2	24	20	2	12	10
3	1	18	12	4	36	29
Total	9	56	44	17	51	40

* denotes groups of seven students due to reduced number of students within the seminar.

Dependent Variables

Individual learning.

A 20-item multiple choice post-lesson test was developed by the researcher to measure individual learning outcomes for the mission analysis and design processes (see Appendix B). Lesson objectives and doctrine were used to develop a Table of Specifications (Appendix C). Based on a group product from a previous class (Appendix D), the Table of Specifications was used to develop post-test items that measured the students' ability to evaluate the completeness, quality, and effectiveness of the analysis conducted by previous students. Faculty from the Contingency Planning Working Group conducted an expert review of the post-test to assess content-related validity and usability (Vogt, 2007). They recommended numerous minor changes to item and response wording. The initial post-test was pilot tested by 8 students in an earlier JFSC Class with possible scores ranging from 0 indicating minimal knowledge to a perfect score of 100. The descriptive statistics for the pilot test were $M = 38.13$, $SD = 6.51$. As a result of the pilot test, response options for five questions were modified to ensure one best correct answer, five different questions were modified for clarity, and two questions were dropped and replaced with alternatives.

The current amended instrument was used for the study. The descriptive statistics from the study were $M = 60.29$, $SD = 12.58$, which demonstrated significant improvement over those recorded for the pilot test. Scores ranged from a low of 29.41 to high of 88.24. A Kuder-Richardson Formula 20 Reliability Coefficient (Thorndike, 2005) of .27 was calculated for the post-test, which is well below the desired level of .70. Study

results for the post-test should be interpreted with caution due to low reliability and validity of these scores.

Group learning.

At the end of the lesson, each group prepared and presented a briefing to the other groups within the seminar communicating their understanding of the environment, the current and desired future conditions, and their rudimentary operational design. (Topic areas and sample slide are contained in Appendix E.) Following the study, a team of three experienced faculty members used a rubric (Appendix F) to score the briefing developed by each student groups. The rubric was based on existing military doctrine and matched the values weighting contained within the Table of Specifications (Appendix C). Six subject-matter experts in the joint operation planning process at the Joint Forces Staff College provided validity evidence for the rubric content and associated scores. Their comments were used to modify the initial rubric. Those same individuals were then asked to provide comments on the modified rubric. Following the second expert review for content validity, the rubric was pilot tested to evaluate group products from an earlier seminar class. No changes were proposed to the rubric as a result of the pilot test. For the study, three faculty raters used the rubric to score group products with results ranging from 12 to 83 with M =48.97 and SD = 17.09. The raters received a short training session that familiarized them with the rubric and its use. The raters were all experts on the problem and process used by the students. Inter-rater reliability for the faculty raters was determined by using intra-class correlation coefficients (Shrout & Fleiss, 1979). The intra-class correlation for single raters was .86 with a Cronbach's α = .95, which is above the acceptable level of .8 (Girden, 2001).

Student contributions.

The Student Contribution Survey (Appendix G) was used to measure student perceptions of their contributions and those of their peers during completion of the group project. The Student Contribution Survey was based on the instrument developed by Li (2001) and contained two distinct two parts. Part one had seven topic areas and an overall rating that used a five-point Likert scale to rate the contributions of each student for each indicator area as outstanding (5), better above average (4), average (3), willing but not very successful contributions (2), and did not contribute in this way (1). The topic areas reflected the specific tasks that student groups completed during their analysis. For students who rated themselves far differently from their peer ratings, their self-ratings were excluded from their scoring, using only the marks provided by their peers. This process ensured students did not artificially inflate their own contributions. For each student, the mean of the scores provided by their group peers and the self-rating were calculated for each area and overall and then compared for large and small groups. Where students failed to provide an overall rating, an average of the scores from the previous seven topic areas was used to calculate the overall rating. While individual scores ranged from a low of 1.00 to a high of 5.03, mean scores ranged from the lowest (M = 3.28, SD = .83) for force analysis to the highest (M = 3.63, SD .66) for the development of the presentation. Reliability was not calculated to assess the consistency of scores for each student because the number of ratings for each student was too small to calculate reliability using traditional means (Zhang et al., 2008).

Part two of the Student Contribution Survey, which was not part of the Li survey instrument, consisted of four open-ended questions soliciting comments on the dynamics

within their group. These questions provided a glimpse into the group interactions and helped determine where contributions by members were constructive or detrimental to the overall group effort. Specifically, responses provided indications for where conflict spurred open discourse within the group or where the level of conflict detracted from group efforts.

Cognitive load.

Students were asked to rate their own mental effort using a single item, 9-point symmetric Likert scale developed by Bratfisch and Borg (1972) and modified by Paas (1992) (Appendix H). While there is considerable debate regarding the use of the single item scale for measuring mental effort (de Jong, 2010; Moreno, 2010), it remains the standard subjective measure for determining overall cognitive load (Paas et al., 2003). The instrument required students to rate mental effort following eight different topic areas, as well as upon completion of the mission analysis survey. Students were also asked to provide an overall mental effort score for the entire study. The nine-point Likert scale assumed that students can be introspective with regards to their own cognitive processes and report the amount of effort required to analyze the problem (Gopher & Braune, 1984). Scores on the scale range from one, representing very, very low mental effort transitioning to nine, which represents very, very high mental effort. The scale was provided and explained to the students prior to the study and during the lessons. Actual scores for the study ranged from a low of 1 to high of 9 for various topic areas with force analysis rated the lowest (M = 4.92, SD = 1.76) and objective development rated the highest (M = 6.00, SD = 1.49). Paas (1992) found that the Cronbach's alpha reliability coefficient (α) was .90 for studies of short duration. This study asked students to rate their

cognitive load for each topic area, the exam, and their cumulative effort for the entire project, and examined each topic area separately to assess how changing task complexity impacted mental load in a group setting.

Normalized group learning.

Individual normalized group learning scores were calculated using the process developed by Li (2000) (see Appendix I). This process was used to calculate a normalized group product score for each student based on their self-reported and peer-assessed contribution scores. The normalization process included the group product scores that were previously developed to measure group learning. Normalized group learning scores ranged from 14.47 to 84.13 with $M = 54.64$, $SD = 15.46$. Inter-rater reliability analysis using intra-class correlation coefficients was performed to determine consistency among raters. The inter-rater reliability for raters was .79 indicating a good level of agreement.

Data Analysis

Table 3 below displays the sample size, dependent variables, instruments and methods used to analyze each research question.

Table 3

Research Questions and Dependent Variables

Research Questions	Sample Size (*N*)	Dependent Variables	Instruments	Analysis Method
1. How does group size effect learning?	64	Individual Learning	Post test	Independent-samples *t*-test
2. How does group size effect student contributions?	84	Participation	Participation Survey	Two-way ANOVA
3. How does group size effect cognitive load?	70	Individual cognitive load	Mental Effort Survey	Two-way ANOVA
4. Do Group scores predict individual learning?	57	Learning scores, Group Learning Scores, Participation	Learning Assessment, Group Product Rubric, Participation Survey	Pearson correlation coefficient
5. Which factors best predict student learning?	54	Group size, participation, individual cognitive load, group learning, and normalized group product	Learning, Participation Survey, Mental Load Surveys	Multiple regression

Procedures

Pre-instruction.

Prior to the start of the study, participating faculty members received a briefing on the goals of the study, how groups were to be formed, and how the lesson was expected to flow. Faculty members were shown the data instruments and received explanations as to how and when they were to be completed.

Three weeks prior to the study, students received a presentation discussing the overall intent of the study and expectations for participants. They were asked to sign an informed consent document (see Appendix A) indicating that they understood the intent of the study, expectations of them as participants, and the degree of risk involved as participants.

A student briefing was provided for each seminar prior to the study to establish a common baseline for understanding the operational environment, i.e. describing current and desired conditions, identification of who the key players were, what those key players intentions were, and how they interacted with each other. This information was essential in enabling students to develop group mental models of the environment in the short time span available in the classroom. The study covered only a block of instruction that incrementally guided students through a mission analysis and development of an initial, complex operational design.

Instruction.

The study session began with each seminars faculty leading a review of the activities involved in the mission analysis step (see Appendix E), contextualized within the Joint Operational Planning Process (JOPP). This activity had been introduced earlier

during the earlier Theater Campaign Planning Course. The faculty followed this up with a guided discussion of how the step differs when done within the context of contingency planning. The two sessions required 40 minutes to complete. The JOPP mission analysis step is similar to the ill-defined problem-solving process steps 1 (articulation the problem space and contextual constraints) and 2 (identification and clarification of alternative opinions, positions, and perspectives of the stakeholders) described by Jonnasen (1997). This review focused on the differences between steady-state and deliberate planning with an emphasis on preparing students to participate in the exercise portion of the lesson. Faculty members assigned a student leader for each seminar. The student leader participated as a group member while assisting faculty in keeping students on schedule and leading the effort to consolidate the multiple briefs into a single seminar product that was to be used in later lessons.

The study itself took place over a single lesson consisting of 13 hours of instruction that incrementally guided students through development of an initial operational design and mission analysis for a specific ill-structured or wicked military-style problem. The 13 hours for the study were divided into 3 separate sessions as follows: day 1, introduction and two hours class work; day 2, two hours of class work, and day 3, eight hours of class work. Students attended electives classes on day 2 that were not associated with the study. Following the brief review, students were informed of their group assignments and commenced with their analysis and design efforts. Each group was provided a briefing template (see Appendix I) that acted as a scaffold to assist groups in structuring their efforts. Groups were free to organize their efforts as they felt necessary to accomplish the task. They had to gain consensus on how best to accomplish

each step, what information they needed to accomplish the step, and how best to utilize that information. Their background, experiences, as well as doctrinal guidance all played key roles in their processing of information. Although each group used the JOPP framework (Joint Staff, 2011), the process and factors each group used to solve problems, the mental models of the problem space and the actual problem defined as requiring resolution was somewhat different for each group. The JOPP process was simply a vehicle for to students think critically about analyzing problems. Upon completion of each topic area, students were asked to rate their mental effort and the participation level of themselves and each group member for that particular topic area. Upon completion of the overall analysis, each group prepared a formal briefing that explained their analysis and how that led to their proposed operational design. Groups were asked to question and evaluate the group presenting the briefing to better understand their perspectives, beliefs, and the process used to derive their analysis. Group briefings were later evaluated for quality and completeness by a team of raters.

Post Instruction.

On the day following the final class session, each student individually completed the post-test which was used to measure their academic achievement. Retrospectively, they rated their own overall mental effort for the entire study period. They also completed the Student Contribution Survey (Appendix G) that incorporated the self and peer-evaluation instruments.

CHAPTER 3

RESULTS

Individual Learning

Post-test item analysis from the final study indicated three questions had correct response rates below 7 percent. Those questions were excluded from scoring and the remaining seventeen questions were used to calculate post-test scores. Based on only seventeen questions, the relative value of each question in scoring was raised to 5.88 or 1/17 of the potential total score of 100. An independent-samples *t*-test was conducted on the post-test scores to determine if student learning was greater for students working in small groups (3 students) or large groups (6-7 students) while collaboratively solving complex, ill-structured, real-world type problems. The test ($N = 64$) was not significant $t(62) = -.42$, $p = .68$. There was no significant difference in the level of learning between students in small ($M = 59.01$, $SD = 11.43$) and large groups ($M = 60.29$, $SD = 13.28$). Table 4 shows the learning score statistics for each group.

Table 4

Individual Learning Scores

Group Size	Sample Size	Mean	Standard Deviation
Small Groups	32	59.01	11.43
Large Groups	32	60.29	13.28

Student Contributions

An 8 X 2, two-way ANOVA was conducted to evaluate the effects of group size on student contributions across eight topic areas of the mission analysis process used for

collaborative problem-solving of ill-structured problems. The means and standard deviations for student contributions as a function of the two factors are presented in Table 5. Results indicate no significant main effects for group size, $F(1,656) = .18$, $p = .67$, partial $\eta^2 = .00$, a significant main effects for sub-task, $F(7,656) = 3.40$, $p = .001$, partial $\eta^2 = .04$, and no significant interaction between the sub-tasks and group size $F(7, 66) = .66$, $p = .71$, partial $\eta^2 = .01$. A primary purpose of the study was to determine if group size impacts group member contributions to the overall group effort for those in a collaborative problem-solving environment. The findings suggest group size had no impact on the level of student contributions for students in a collaborative problem-solving environment.

Table 5

Student Contributions Based on Group Size and Sub-Task

	Small Groups			Large Groups			Cumulative Total		
Sub-Task *(N = 84)*	Rank	Mean	SD	Rank	Mean	SD	Mean	SD	Rank
Fact identification	4	3.52	.69	6	3.47	.51	3.49	.60	6
Mission development*	2	3.59	.73	3	3.60	.44	3.60	.60	3
Problem identification	5	3.45	.94	2	3.62	.45	3.54	.73	4
COG analysis	6	3.44	.75	1	3.62	.46	3.53	.63	5
Force analysis	8	3.27	.99	8	3.29	.63	3.28	.83	8
Risk analysis	7	3.29	.93	7	3.30	.65	3.29	.80	7
Presentation development*$	1	3.73	.72	5	3.53	.60	3.63	.66	1
Overall*	3	3.59	.61	4	3.61	.36	3.60	.50	2

* Denotes significant difference from Force Analysis at .05 confidence level

$ Denotes significant difference from Risk Analysis at .05 confidence level

The main effect for sub-task indicated that some of the sub-tasks of the mission analysis process had higher levels of student contributions than others. Follow-up analyses of the main effect for topic areas consisted of all pairwise comparisons among the eight sub-tasks of the mission analysis process. The Tukey HSD procedure was used to control for Type I error across the pairwise comparisons. The results of this analysis indicated that the contribution levels for students working the force analysis sub-task was significantly less than for students developing the group presentation, or the overall contribution level of students during the entire study period. There was also a significant difference in the contribution level of students developing the brief and for the period during the risk analysis sub-task. There were no other significant differences among student contributions to the group effort for any of the other mission analysis sub-tasks.

Student Contribution Narrative Results

The Student Contribution Survey instrument used four open-ended questions to elaborate the results of the quantitative findings of the study. They addressed group decisions and behaviors that occurred within each group over the course of the study. Comments were received from 56 students on at least one of the four questions. The comments received represented inputs from eight of nine large groups and 13 of 17 small groups. Content analysis (Patton, 2002) was conducted on the comments with multiple reviews conducted. The coding was based on inductive reasoning to determine categories of group behaviors. Items and categories that were the product of a lone student were dropped or removed. The initial number of categories was consolidated and reduced to a total of nineteen classifications spread over the four questions. Two raters received training on coding and then coded the comments into the new categories. Inter-rater

reliability was calculated using the Miles and Huberman (1994) methodology to compare the number of agreements divided by the number of agreements plus the number of disagreements. The inter-rater reliability for the two raters was .92. Table 6 shows the results of the analysis grouped by question and major topic.

Table 6

Open-ended responses by group

Issue	Large Groups: Groups reporting Issue	Large Groups: Total # Groups	Large Groups: % Groups	Small Groups: Groups reporting Issue	Small Groups: Total Groups	Small Groups: % Groups
How well group functioned and any issues?						
Functioned well with no issues	8	8	100	10	13	77
Shortage of time	3	8	38	0	13	0
Social loafing demonstrated	1	8	13	1	13	8
Task organized	0	8	0	3	13	23
Lack of SME's	1	8	13	1	13	8
Compromise difficulty	4	8	50	1	13	8
Dysfunctional group	0	8	0	1	13	8
Workload distribution						
Task organized	4	8	50	3	13	23
Worked as whole group	3	8	38	0	13	0
Even workload	4	8	50	8	13	62
Unequal workload	2	8	25	1	13	15
Time factor	3	8	38	2	13	15
Level of conflict within the group						
Little to no conflict	7	8	88	10	13	77
Major conflict	1	8	13	1	13	8
Great discourse	8	8	100	11	13	85
Social loafing	1	8	50	1	13	8
Other comments						
Equitable split of experience	1	6	17	1	11	9
Effective group size	3	6	50	6	11	55
Leader importance to effort	1	6	17	1	11	9

Students almost universally felt that their groups worked well. All of the large groups and 10 of 13 (77%) small groups indicated that they perceived their groups to be effective over the course of the study. Two small groups did not provide indications of group effectiveness. The final small group was simply dysfunctional with group members having severe personality conflicts that they were unable to overcome, thereby forcing them to work individually. Their group product was completed by one experienced group member. Half of all of the large groups (3 of 6) had students report that they felt that six students was the optimal size for group interactions. In an interesting twist, small groups also had the same percentage of groups (6 of 11) with students report that three students was the optimal group size. This study did not allow students to cross groups and as such students did not experience working groups of a different size. These statements regarding optimal size are more likely a testimonial to the student's perceptions with regards to the overall effectiveness of their groups. One small group captured their thoughts on the effectiveness of small groups with; "This was probably the most effective group I've ever worked with. All players contributed equally - first time I've ever seen it."

An interesting and unexpected outcome was that half of the large groups (4 of 8) and a quarter (3 of 13) of the small groups organized themselves into smaller components to work the topic areas. Although the numbers are too small to make a determination, indications were that there was little difference between those students in large groups that worked as a single entity ($M = 61.27$; $SD = 10.18$) and those students from large groups who broke into smaller entities ($M = 59.13$: $SD = 11.88$). Thus their scores were not much different from students that were in small groups. In the case of three large groups, they explained that time was a significant factor and felt that they could not

accomplish their analysis in the allotted time if they worked every topic area as a single integrated unit. The time issue became apparent after students assessed their progress following the initial two-hour class session. One large group expressed, "Initially, we tried to handle each task with all participants but changed on day two where we broke into sub-groups to address separate tasks and back brief the other sub-groups on our product." This change allowed the group to reach consensus faster resulting in an overall acceleration of the process. The sub-groups were not formally assigned and were different within each group. Three large groups indicated that they worked the entire study as a single entity and did not assign topic areas to individuals or sub-groups. One large group provided no indications as to whether they did or did not re-organize themselves into sub-groups to work topic areas.

While three small groups task organized into individual effort, they did not indicate that time was the factor that drove that decision. Only one small group indicated time was an issue but that group did not break down into smaller groups or individuals. None of the remaining small groups specifically reported that they worked the exercise as a single unit but researcher observations indicated that most small groups did not re-organize by topic area. The small group that was dysfunctional did not indicate if they had issues with time. For students in the three small groups that did task organize, the post-test scores for students appear to be much lower ($M = 54.12$; $SD = 12.75$) than the sample mean. While such a small sample size makes it difficult to conclude much based on these scores, the evidence appears to support the significant body of earlier research whereby students in groups outperform individuals.

Half of the large groups (4 of 8) and two thirds (8 of 13) of the small groups indicated that they felt the student workload was equally distributed among group members. Only two large groups and one small group reported that students felt the workload was uneven. While two large groups reported an uneven workload, one additional group reported the lone case of social loafing with two students who principally observed the process rather than participate. That group indicated that it was easier and more productive for the overall group to let those few members remain disengaged rather than pressure them for involvement. The one small group reported an uneven workload had one member who spent an inordinate amount of time on his cell phone.

Students indicated that there was a great deal of discourse with little conflict between group members (100% of the large groups and 85% of the small groups). Where conflict was noted, its character can be best described by this statement "conflict was relevant and productive towards objectives." Students indicated that the discourse was positive and moved the group forwards towards resolution of their group project. Only in the dysfunctional small groups did conflict negatively impact the progression towards group effectiveness.

A final issue noted in the comments was related to the experience levels of the group members. The overall student population had widely divergent, formal planning experience. Some had a great deal of experience while others had very little. One large and small group indicated that they felt their group had no group members with experience and as such, they struggled with completion of the group project. Although the sample sizes were very small, the large group reporting a lack of experience had post

test scores that appeared to be above the overall mean (M = 63.24, SD = 2.94) whereas the small group post-test scores appeared to be well below the overall group mean (M = 52.94, SD = 10.19).The small group that reported this issue was also one of the two small groups that reported having issues with time. This is the same group that reported they also did not have enough group members to "task-organize." This group's overall group product score was the lowest score recorded for any of the small groups (45.67). It appears that this group was less effective than other groups which may have been the result of having no planning experience within the group to help guide group members through the process.

Cognitive Load

An 8 X 2, two-way ANOVA was conducted to evaluate the effects of group size on cognitive load across eight topic areas of the mission analysis process used for collaborative problem-solving of ill-structured problems. The means and standard deviations for cognitive load as a function of the two factors are presented in Table 7. The ANOVA indicated no significant main effects for group size, $F(1,544) = 1.42$, $p = .23$, partial $\eta^2 = .00$, significant main effects for sub-task, $F(7,544) = 3.96$, $p = .00$, partial $\eta^2 = .05$, and no significant interaction between the sub-tasks and group size $F(7, 54) = .97$, $p = .45$, partial $\eta^2 = .01$. A primary purpose of the study was to determine if group size impacts cognitive load for those in a collaborative problem-solving environment. The main effects for group size were not significant suggesting that group size had no impact on the cognitive load experienced by students.

Table 7

Cognitive Load by Group Size and Sub-Task

	Small Groups		Large Groups		Cumulative Total		
Sub-Task (N = 70)	Mean	SD	Mean	SD	Mean	SD	Ranking
Fact identification	5.17	1.79	5.23	1.56	5.20	1.67	7
Assumption development	5.69	1.59	5.29	1.82	5.49	1.71	5
Termination criteria	5.29	1.86	6.17	1.58	5.73	1.77	4
COG analysis*	5.94	1.31	5.94	1.64	5.94	1.47	2
Objective development*	5.86	1.63	6.14	1.44	6.00	1.53	1
Force analysis	4.94	1.83	4.89	1.81	4.91	1.81	8
Risk analysis	5.26	2.10	5.71	1.73	5.49	1.92	5
Overall*	5.89	1.02	5.97	1.04	5.93	1.03	3

* Denotes significant difference from Force Analysis at .05 confidence level

The topic area main effect indicated that some of the sub-tasks of the mission analysis process required greater mental effort than others when collaboratively solving complex, ill-structured problems. Follow-up analyses of the main effect for sub-tasks examined this issue. The follow-up tests consisted of all pairwise comparisons among the eight sub-tasks of the mission analysis process. The Tukey HSD procedure was used to control for Type I error across the pairwise comparisons. The results of this analysis indicate that the cognitive load experienced while conducting force analysis was less than that experienced by students developing termination criteria, developing objectives, and significantly less than the overall cognitive load expended during the study. This finding was not surprising as participation results indicated significantly lower scores for force analysis as well, indicating that less thought and overall effort was expended by the group

on the force analysis sub-task. There were no significant differences among the cognitive load for any of the other mission analysis sub-tasks.

Individualized Group Scores

A Pearson Product Moment correlation coefficient was computed for the post-test score and the adjusted group product score. This methodology used peer and self-ratings of student contributions to modify the overall group product score, producing a normalized or adjusted group product score for each individual. Normalized scores were calculated using the methodology articulated in Appendix I (Li, 2001). The Pearson's Correlation coefficient for Adjusted Group Scores and Individual Learning Scores was not statistically significant, $r(55) = .21$, $p = .12$. The results indicate that the group scores, even when adjusted for a group member's participation, do not predict the level of student learning for individual students.

Predicting Individual Student Learning

A multiple regression was conducted to evaluate how well student observable behaviors within a group predicted individual student learning. The predictors were group size, individual cognitive load, individual participation levels, and group product quality. The linear combination of student observable behaviors within groups was not significantly related to individual student learning $F(4,52) = .71$, $p = .59$. The multiple correlation coefficient was .23, indicating that approximately 5% of the variance in student learning was accounted for by the linear combination of observable student behaviors within groups.

Table 8 presents bivariate and partial correlations of the individual predictors. All bivariate correlations between the observable behaviors and the individual learning index

were positive, with none of the four indices being statistically significant ($p = .59$). The partial correlations between the student observable behaviors were not significant. Combined, they account for only 5% of the variance of the individual student learning index and as such, are relatively unimportant in predicting student learning for students participating as members of collaborative problems-solving groups in learning environments.

Table 8

Bivariate and Partial Correlations of the Predictors with the Individual Student Learning Index

Predictors	Correlation between each predictor & post-test	Correlation between each predictor & post-test controlling for all other predictors
Team size	.15	.10
Overall cognitive load	.13	.13
Overall participation	.10	.03
Group product quality	.14	.12

CHAPTER 4

DISCUSSION

This study examined middle-aged, professional students, working in a collaborative, ill-structured problem solving environment, to determine if there were differences in student learning, individual student contributions to the group effort, or mental efforts between students in large (six students) and small groups (three students). It also looked at how group product scores, modified for student contributions predicted student learning. The final study area examined whether contributions, mental effort, modified group product scores predicted student learning.

Individual Student Learning

This study found that there were no significant differences between large (6 students) and small (3 students) groups. The course of instruction that students were enrolled in required them to develop a set of solutions to complex, international relations problems. They were to address the use of all of the instruments of national power (diplomacy, information, military, and economic) to develop those solutions. The study period covered only a small portion of that course (12 of an overall 56 hours). The lessons used for this study required students to analyze the problem, the specifics of what they were directed to accomplish, develop a mental model of the operational environment, and then develop an operational framework. Their efforts from the study period were used later in the course to develop solutions and then evaluate their solutions for adequacy and feasibility. The complexity of the problem was of a type that students would expect to encounter in real-world situations and could be asked to solve in collaborative problem-solving environments. As such, many of these students had

varying levels of previous experience in solving such problems in collaborative environments.

Group dynamics.

While the study showed no significant differences between large and small groups, the open-ended questions showed major differences in the dynamics of the groups. Time or specifically the lack of it became a major driving factor in how groups constructed their approaches to accomplish their tasks. Allowing everyone to voice opinions/thoughts not surprisingly took longer for the larger groups and subsequently, they reported more instances of time-related issues. They also indicated they had more difficulty in reaching consensus on group positions before moving on to the next topic area. Some groups broke down into smaller groups, worked topic areas, and then compiled their efforts into their consolidated briefing. Other groups worked every topic area as a single entity. Some groups reported that having no students with prior experience presented difficulties for the group as they struggled to accomplish topic-related tasks.

One large and one small group specifically indicated that a lack of subject matter experts adversely impacted their overall group effort. While the large group indicated that "some areas caused hang-ups" due to the lack of SME's, the impacts for the small groups were reported to be more significant as they felt that the group members "weren't smart enough on the subject." The small group post-test scores appeared to be influenced by one score that was just above the mean. The scores for the other two individuals from that group were well below the overall mean. This is the same group that reported they also did not have enough group members to "task-organize". This group's overall group

product score was also the lowest score recorded for any of the small groups (45.67). It would appear that this group was less effective than other groups which may have been the result of having no planning experience within the group. While the sample was isolated, it does support earlier research that showed a lack of skills resulted in lower group effectiveness (Wiley & Jensen, 2006; Laughlin & Adamopoulis, 1980). The composition of the other study groups with regards to student expertise levels was not reported to be an issue. For the study, groups were compiled randomly with the only requirement being that each group consisted of an appropriate military service mix which prevented an assurance of an expertise spread among the seminars. As the number of groups within a seminar increases, the likelihood that groups would have no previous planning expertise increases. With an increased number of groups to monitor, faculty members had less time available to monitor group activities and may have reached the point whereby they cannot appropriately monitor the number of student groups they are using. This was particularly important upon exercise start as student confusion and uncertainty are highest at that point. Student expertise within the process became a bridge to "scaffold" other students allowing faculty to cover more seminars. Initial problems and misunderstandings have the potential to cascade as students' progress thru the process, impacting group effectiveness and subsequently student learning.

Task-organized groups.

One surprise the open-ended questions uncovered was the number of large seminars that decided to task-organize in to smaller entities. This division of labor can be a common problem when using collaborative groups (Akella, 2012) with a subsequent risk of the groups becoming competitive, isolated, and apathetic. Half of the large

seminars opted to divide the sub-tasks among members and collate their inputs later. As a comparison, only three of the small groups opted to task organize. Time or the apparent lack of it appears to the key determinant in this decision for the large groups but not the small groups. The small groups that reported time was an issue were not the groups that task organized. While time may have been a factor in their decision, it was not reported as such. Students in the large groups felt that there was insufficient time available for them to accomplish their learning tasks based on their initial impressions of the time required for discourse among six group members. Students commented on how long normal discourse took when all six group members provided their perspective. One large group indicated that they attempted to conduct everything as a single entity until they finished the first two-hour class session. They realized their lack of progress jeopardized their efforts and decided the only way to accomplish the task in the time remaining was to break the large group into smaller entities. All four groups that task organized also indicated that they had difficulty reaching consensus. Because large groups only reported this, it would appear that for larger groups, normal discourse and subsequent attainment of group consensus took longer with more group members.

Since half of the large groups functioned as small groups, the question arose as to if there was a difference in the learning between these two sets of groups that could be attributed to their behavior. However, we saw that there was little difference in the post-test scores between small groups, task-organized large groups, and integrated large groups. This on the surface appears to be surprising as students in the integrated groups would be expected to engage all of the material while those in the task-organized groups would only engage the sub-tasks they worked on.

However, in the task–organized groups, students returned to the larger group to combine their material into a single brief. Using Blooms taxonomy (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1984), students that worked on a particular topic area were involved at the synthesis level of learning. Thus students in the single entity groups were involved at the synthesis level for all of the material. In the task-organized groups, when students combined their material for the final brief, students were forced to evaluate their product as well as that from everyone else to ensure that all of the parts were consistent with parts developed by others and supported the overall effort. Thus, while task-organized students may only have synthesized a small number of group products, they were engaged in the analysis and evaluation of all of the groups' products and as such remained at very high learning levels.

Optimal group size.

The earlier research of Laughlin et al. (2006) found that there were no differences in the performances between groups of three, four, and five participants when involved in solving highly intellective problems. These groups required fewer runs to solution and proposed more complex solutions. Groups of five were the largest groups involved in the study and as such, it is possible that larger groups may have been as effective. Lyons et al. (2003) proposed that for more complex problems, groups of four to seven students were ideal as larger groups allowed for sufficient viewpoints to stimulate active discussion. They did not note that time was an issue or limiter in the development of student discussions. They argued that for groups larger than seven, introverted students were more likely to not become engaged in discussions. Groups of three were covered in Laughlin but not Lyons. Groups of six were outside the scope of Laughlin's study but

included in Lyons. This study found that for middle-aged students solving complex problems, there was no significant difference in the learning between students in groups of three and groups of six. These results seem to indicate that perhaps there is overlap in optimal group sizes between what Laughlin et al. and Lyons et al. proposed and perhaps optimal sized groups in ill-structured problem-solving environments range from three to seven individuals. These results would appear to run counter to the meta-analysis conducted by Lou et al. (1996). However, their meta-analysis examined students in cooperative learning environments which involved students solving very different types of problems. Students in this study, as they were in Laughlin's study were working in a collaborative environment with very different types of problem sets. And as indicated earlier, insufficient time played a major role in group decision making and forced students into the decision to task-organize.

The post-test was based on a student sample of mission analysis. As expected, it had a high level of complexity and required significant effort to digest. It was used only for the study and played no part in the assignment of student grades. As such, many students did not expend the requisite time necessary to properly analyze the student sample to enable sufficient understanding to answer all of the survey questions. A number of students indicated to the researcher that after answering the first few questions, they understood they had not committed sufficient time to analyze the product and were unwilling or unable to provide the additional time necessary to complete the learning survey instrument. This likely lead directly to the lower than expected scores, the lower number of students completing the learning survey instrument, and contributed to the low reliability score for the instrument itself..

Student Contributions to Group Effort

In this study, there was no significant difference in the cognitive load experienced by students in large and small groups. While there were no significant differences between large and small groups, significant differences were found between the topic area Force Analysis and the topic areas for Center of Gravity (COG) Analysis, Objective Development and the overall lesson. This finding indicated that the greater transactional memory system of the larger group did not play a role in students' ability to accomplish the educational task or impact individual student learning. Since learning, cognitive load, and participation, displayed no significant differences between group sizes, this would seem to indicate a couple of very important items. In this study, students did not appear to exceed their cognitive load capabilities. If they had done so, one would have expected to have seen a lower cognitive load from the members of the large group as they would have been able to use the group's transactional memory system to spread the cognitive load among the group members. This would have resulted in a subsequent lowering of individual cognitive load for larger groups. However, there were no indications that this happened for the larger groups.

Cognitive load was also examined to determine if it remained the same across all of the steps of the process. The study showed the student cognitive load experienced during force structure analysis step was less than that experienced during the development of termination criteria, objectives, or the overall cognitive load experienced for the entire study period. This may be the result of two independent factors. One factor was that force structure analysis can be fairly straight forward in that one looks at the potential foe and uses algorithms to determine if they have been provided sufficient force.

The other factor was that four small groups and one large group failed to conduct force structure analysis. The low score was likely a reflection of that failure to conduct that topic area. The narrative does not indicate why but the assumption was students elected either not to conduct force structure analysis because of the critical time issue or spent less time developing it as it was one of the final topic areas to be completed. Most students understand that while force structure analysis may start during mission analysis, it becomes much more focused as the planning effort moves through course of action development. These groups may have determined that they would complete this step later in the planning process, but during a period of time outside the scope of this study.

Cognitive Load

There were no significant differences in contributions made by students in large and small groups. Early work showed that small groups (3-4) were the optimal size for learning (Webb & Palincsar, 1996). Smaller groups maximized student engagement of the material while increasing the opportunity to participate in group activities (Lohman & Finkelstein, 2000). Smaller groups were less likely to have students to display "free riding" or "social loafing" behaviors (Stroebe & Diehl, 1994). Yet, this study showed that there was no significant difference in participation levels between students in large and small groups. The data was supported by narrative comments as only two large groups indicated that there was an unequal workload. They indicate that the majority of the work was done largely by only a couple of students. They did not indicate the reason for such but there were no indications of students not wanting to participate. This may be a case of students who have experience stepping in and "taking over" to drive the group to completion. This population has no shortage of people willing to "take charge" and lead

an effort. If so, then this could be a case where previous experience had a negative impact as those with experience did a disproportionate amount of work to the detriment of the remainder of the group. While two groups commented on unequal workload, all of the large groups indicated that they felt the groups were effective. The lone small group that reported an unequal workload was also the group that was dysfunctional. What was interesting was that the two large groups that reported unequal workload were not those that reported social loafing. In both cases, the comments would lead they researcher to believe that this was more a case of "type A" personalities stepping forward rather than others backing away from participating. The experience level of those who did more was unknown but it is possible that those with experience simply took over and lead the effort.

Regardless of group size, students were highly engaged with only a few cases of "social loafing" behavior reported. As previously discussed, the sample consisted on mid-grade military officers with significant time in their professional field. As professionals, when working in their professional field, their motivation and desire to perform exceed what would be expected from a truly random sample of the overall populace. The military aspect of the sample likely also plays a large part in the determination of results. As such, they are all experienced with working in collaborative environments and as military, they are usually task-focused ensuring that they modify their processes and efforts to ensure that the deliverable becomes the highest priority. Instances of social loafing and free-riding are greatly reduced because the military members are usually team-focused and as such have a strong desire to not fail their fellow group members. While some students reported in their comments that their groups task-organized, the

individual participation scores within those groups did not reflect such. Students were rated fairly consistently evenly by the group across the seven topic areas as well as for the overall effort. Rarely were students rated highly for some topic areas (the parts they worked on) and then low on parts where they only provided comments during the brief compilation. For groups where they task-organized, wider variation in participation between the seven distinct segments was expected but the data does not support such. This finding indicated students likely completed the participation instrument after the study rather than complete each section as it was accomplished allowing perceptions with regards to overall participation to be formulated during the times when the large group was together rather than when divided into smaller entities. With groups that break-up into smaller groups, they had to recombine and consolidate their results prior to the briefing. To make the briefing coherent and consistent, group members had to be critical in their evaluations of each component, allowing students to contribute in areas that they did not necessarily work on.

An interesting perspective was that for small groups, the activity where the participation level was highest was for the development of the presentation. That activity was only the fifth highest activity for large groups. This may reflect the fact that half of the large groups task-organized and those group members had to interface with the whole group only during the time when the brief was actually compiled and not for the activities each individual actually worked on. For much of their effort, these four large groups worked cooperatively to develop the briefing. For the smaller groups, every member worked collaboratively on all of the parts and as such all participated in the development of the briefing. So while there may appear to be a trend showing differences, in actuality,

there were no significant differences in the participation levels for group members in the development of the briefing. A number of groups commented on the importance of the person who collated the material into the brief. This person often became the most important person as they played instrumental roles in keeping everyone "on track" and not allowing the groups to pursue discussions that detracted from the mission analysis effort. This study did not capture how much of a group member's contributions were administrative (collating material into the brief) and how much was cognitive (development of the material to put into the brief). If their role was the former, then the expectation would be that their learning scores would be significantly lower as they would not have actually conducted any of the sub-tasks.

Normalized Group Scores

Normalized group scores, which are group product scores adjusted using a methodology from Li (2000) were found not to be a valid predictor of student learning. A number of studies have been conducted to determine if there was a better method of assigned individual grades to students for work on group projects (Goldfinch & Raeside, 1990; Conway, et al., 1993; Goldfinch, 1994; Li 2001). The most common methods involve the use of peer and self-ratings with regards to each student's contributions to the project. While Webb (1992) found that the grades assigned to group products did not reflect individual student learning, this study attempted to determine if group project scores, weighted based a peer assessments of student contributions, would be useful in predicting individual student learning. Groups products often are the result of only some of the members with the inherent difficulty in ascertaining individual contributions to the overall group product. As a result, free-riding and social loafing behaviors often have the

potential to occur in group projects. Although combining the individual and group contributions may portray a more honest assessment of a student's true knowledge and skills (Webb, 1997), it still does not correlate with actual learning scores. This result supports Webb's (1992) earlier work in that the quality of the group product, even when adjusted for a group member's level of participation, does not predict how much that individual student learned.

Predicting Individual Student Learning

The final part of the study looked to see if group size, individual cognitive load, participation or group product quality predicted group learning. Most of these are items the teacher can observe (mental effort to some degree) so they have value in assisting the teacher in evaluating how well students are learning. Earlier parts of this study informed us that the learning score was not going to be impacted by the difference in group size and the quality of the group product and thus the lack of correlation between these variables and student learning was expected. Finding that student contributions did not correlate well with student learning was an unexpected outcome. The expectation was that as student contributions increased, their involvement with the learning material would increase which would result in a subsequent increase in learning. While there were outlying students that had higher and lower levels of student contributions, across the student sample, student contributions were fairly consistent. The findings seem to indicate that as students had correspondingly higher and lower levels of contributions to the group effort, the individual learning scores did not reflect correspondingly higher or lower scores. This study used peers to evaluate student participation and contributions to the overall group effort. That allows students to acknowledge actual contributions to the

overall student effort and not reward behaviors that detract or confuse the group efforts. This finding supported the earlier research where measurements of student interactions did not correlate with student grades (Davies & Graff, 2005: Morris, Finnegan, & Sz-Shyan, 2005).

For students attempting to bound social problems, their limiting factor became the definition with which the group can develop a model of the problem space. The assumption is that the better the understanding of the problem space, the better the solution that the problem solver can develop. It seems logical, that using case based reasoning (Kolodner, 1992), and larger groups would have a greater number of cases to draw from and thus would be able to develop mental models with greater definition (Moreland & Levine, 1992; Schwartz, 1995). The implication that may be made from this study is that the mental models between large and small groups are of similar size but that may be an incorrect assumption. This class session asks students to synthesize various mission analysis products. The focus of the lesson is on students understanding the process and how the various elements fit together to best apply the elements of the Operational Art of Warfare to solve operational-level military problems. While mental models drive solution sets, this lesson and subsequently this study did not evaluate the mental models that students constructed. As a result, their relative size and complexity when compared against other student efforts is unknown. It is possible that larger groups constructed more complex models with more actors with additional relationships. These additional factors to consider may not have been evident as larger groups had more cognitive resources with which to address these resources making it appear that the mental effort required for the large group was the same as that for the smaller group. And

then again, the mental models may have been the same between large and small groups and the material was sufficiently complex enough that the mental effort for all students was the same. If the mental models were of different complexity, there may become a particular point at which the group size becomes too large but that break point was greater than the group sizes used for this study.

Future Research

This study focused on the performance of groups and the individuals within each group; it did not examine the group dynamics or behaviors exhibited by the groups. While there were no differences in learning or group product quality, how the groups arrived at such point was significantly different. Earlier studies regarding group learning were of short duration or used problems that were not necessarily suitable to groups. The military students in this study showed initiative in modifying group behaviors to ensure that the groups' outcomes were accomplished. Very little research has been done to examine the behavior and dynamics of the groups. Understanding what conditions elicit specific behaviors from groups is important in the design of instructional environments. Future research on student group behaviors could inform the effectiveness of group learning efforts.

As students develop their understanding of the environment, they develop a mental model based on the collective understanding of operators and associated relationships. The complexity of the developed metal model may be one area where large groups possess a major advantage over smaller groups. That mental model developed by each group is based on the background and experience of the group members (Koldoner, 1996). As the number of group members increases, there is more experience from which

group members are able to draw. This increase in group size should result in a more complex, understanding of the problem space, a better, more detailed understanding of the problem, which should eventually result in a better problem solution (Moreland & Levine, 1992; Schwartz, 1995). This study did not examine the difference in understandings of the problem space and if that eventually leads to better solutions. This study stopped at the analysis step of problem-solving. While the study determined that there were no differences between groups of three and groups of six, it remains to be determined if even larger groups remain as effective or if there is some point where groups become too large to become effective, where students are insufficiently engaged with the material to promote student learning. Additional comparisons of even larger groups to smaller groups will help determine where that break point may be.

Narrative comments indicated that experience may play an important role in the effectiveness of a group. A small group that had no one with experience clearly had difficulties and performed below expectations. Their overall learning scores and group product scores were well below the mean for both categories. Experienced group members are able to guide the inexperienced students through the process, particularly in the early steps of the process. Time was critical factor that placed additional pressure on groups. Once a group fell behind, they were unlikely to catch up and their problems cascaded. The instructor could provide the initial guidance and direction early on but the number of groups determines how much time the teacher has to focus on individual groups. This study clearly indicated that experienced students divided among the groups may increase the overall effectiveness of the groups and may play a role in determining how many groups the teacher should use. The instructor should find that as the number of

students with experiences rises, so does the number of groups that can be effectively monitored. This factor of experience and the role it plays is worth additional study.

Conclusions

There were no significant differences for cognitive load, learning and participation levels for large and small groups becomes important for institutions that work in such realms. Little research has been done on how students in this age group learn. The field of andragogy (Knowles, Holton, & Swanson, 2012) is based on the belief that adults learn differently than children. The importance of background and experience are key factors in those describing those differences. As adults age, the background and experiences from which to draw from grows considerably. How they use that base of knowledge to address learning to solve real-world problems as they age is not well understood. Complex real-world problems present unique challenges in that there are no correct answers, simply answers that are better or worse than others (Jonassen, 2010). This study found that when presented with such problems, middle-aged professionals were highly engaged (mental effort as well as student contributions) regardless of the size of the group. Group size was not a limiting factor on individual learning, levels of student contributions to the overall group effort, or mental effort. These finding support and expand upon the earlier research of Laughlin et al., 2006 and Lyons et al. (2003) indicating that the proposed range of group sizes proposed should be expanded to includes groups of three and six. This study shows that for middle-aged professionals, working in complex, ill-structured problem-solving environments, there is no difference in performance for groups ranging in size from three to six group members. That places the decision as to which group size to use back into a classroom management issue

whereby the teacher can opt to use the size that is easiest to manage from their point of view.

REFERENCES

Akella, N. (2012). The Real Deal on Collaborative Learning. *Education, 2*(3), 23-29.doi:10-5923/j.edu.20120203.01

Allaire, J. C., & Marsiske, M. (2002). Well and ill-defined measures of everyday cognition: Relationship to older adults' intellectual ability and functional status. *Psychology and Aging, 17*(1), 101-115. doi:10.1037//0882-7974.17.1.101

Andersson, J., & Rönnberg, J. (1995). Recall suffers from collaboration: Joint recall effects of friendship and task complexity. *Applied Cognitive Psychology, 9*, 199-211. doi:10.1002/acp.2350090303

Apedoe, X. S., Ellefson, M. R., & Schunn, C. D. (2011). Learning together while designing: Does group size make a difference. *Journal of Science, Education, and Technology*, 1-12. doi:10.1007/s10956-011-9284-5

Arlin, P. K. (1989). The problem of the problem. In J.D. Sinnott (Ed.), *Everyday problem solving: Theory and application* (pp. 229-237). Hillsdale, NJ: Praeger.

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology, 20*(3), 287-298. doi:10.1002/acp.1245

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one. *Child development*, 87-96. doi:10.2307/1130391

Barrow, H. S. (1983). Problem-based, self-directed learning. *Journal of the American Medical Association, 250*(22), 3077-3080. doi:10.1001/jama.1983.03340220045031

Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, 361-392.

Bettenhausen, K. L., & Fedor, D. B. 1997. Peer and upward appraisals: A comparison of their benefits and problems. *Group and Organization Management. 22*: 236-263. doi:10.1177/1059601197222006

Bloom, B. S., Engelhart, M. D., Furst, F. J., Hill, W. H., & Krathwohl, D. R. (1984). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain.* New York: Longman.

Bossert, S. T. (1988). Chapter 6: Cooperative activities in the classroom. R*eview of research in education, 15*, 225-250. doi:10.3102/0091732X015001225

Bransford, J. D., Sherwood, R. D., Hasselbring, T. S., Kinzer, C. K., & Williams, S. M. (1990). Anchored Instruction: Why we need it and how technology can help. In D. Nix & R. J. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (Vol. XII, pp. 115-141). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Bratfisch, O., & Borg, G. (1972). *Perceived item-difficulty in three tests of intellectual performance capacity* (No. 16). Stockholm, SW: Stockholm University.

Bruffee, K. A. (1999). *Collaborative learning: Higher education, interdependence, and the authority of knowledge* (2nd ed.). Baltimore, MD: Johns Hopkins University Press.

Camp, G., Paas, F., Rikers, R.m J.P, & Van Merriënboer, J.J G. (2001). Dynamic problem selection in air traffic control training: A comparison between

performance, mental effort and mental efficiency. *Computers in Human Behavior, 17*, 575-595. doi:10.1016/S0747-5632(01)00028-0

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293-332. doi:10.1207/s1532690xci0804_2

Chen, G., Donahue, L. M., & Klimoski, R. J. 2004. Training undergraduates to work in organizational teams. Academy *of Management Learning and Education.* 3: 27-40. doi:10.5465/AMLE.2004.12436817

Chi, M. T.h, Glaser, R., & Farr, M. J. (1988). *The nature of expertise.* New York, NY: Lawrence Erlbaum Associates.

Cho, K. L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development, 50*(3), 5-22. doi:10.1007/BF02505022

Ciborra, C. C., & Olson, M. H. (1988). Encountering electronic work groups: A transaction costs perspective. *Information Technology & People*, *4*(4), 285-298. doi:10.1108/eb022667Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational research, 64*(1), 1-35. doi:10.2307/1170744

Cohen, E. (1994). *Designing groupwork.* Teachers College Press New York.

Cohen, B.P., & Arechevala-Vargas. (1987). *Interdependence, interaction and productivity* (Working Paper No. 87-3). Stanford: Center for Sociological Research.

Collins, K. M., & Onwuegbuzie, A. J. (2000). Relationship between Critical Thinking and Performance in Research Methodology Courses.

Conway, R., Kember, D., Sivan, A., & Wu, M. (1993). Peer assessment of an individual's contribution to a group project. *Assessment & Evaluation in Higher Education, 18*(1), 45-56. doi:10.1080/0260293930180104

Cooper, M. M., Cox, C. T., Nammouz, M., & Case, E. (2008). An assessment of the effect of collaborative groups on students' problem-solving strategies and abilities. *Journal of Chemical Education, 85*(6), 866-872. doi:10.1021/ed085p866

Cowan, N. (2001). Magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-114. doi:10.1017/S0140525X01003922

Davies, J., & Graff, M. (2005). Performance in e-learning: online participation and student grades. *British Journal of Educational Technology, 36*(4), 657-663. doi:10.1111/j.1467-8535.2005.00542.x

DeJong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science, 38*, 105-134. doi:10.1007/s11251-009-9110-0

Dochy, F., Segers, M., & Sluijsmans, D. 1999. The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education.* 24: 331-350. doi:10.1080/03075079912331379935

Dugosh, K. L., Paulus, P. B., Roland, E. J., & Yang, H. (2000). Cognitive stimulation in brainstorming. *Journal of Personality and Social Psychology, 79*, 722-735. doi:10.1037/0022-3514.79.5.722

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*(4), 395-430. doi:10.2307/1170205

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70*(3), 287-322. doi:10.2307/1170785

Felder, R. M., & Brent, R. 2007. Cooperative learning. In P. A. Mabrouk, ed. Active learning: *Models from the analytical sciences.* ACS Symposium Series, 970: 34-53. Washington, DC: American Chemical Society.

Frensch, P. A., & Funke, J. (1995). Definitions, traditions, and a general framework for understanding complex problem solving. In P.A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 3-26). Hillsdale, NJ: Lawrence Erlbaum Associates.

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Lawrence Erlbaum Associates.

George, Jennifer. (1992). Extrinsic and intrinsic origins of perceived social loafing in organizations. *Academy of Management Journal, 35*(1), 191-202. Doi:10.2307/256478

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations. *Learning and Instruction, 16*, 104-121. doi:10.1016/j.learninstruc.2006.02,007

Gilles, R. M., & Ashman, A. F. (2003). *Co-operative learning: The social and intellectual outcomes of learning in groups.* New York, NY: Routledge.

Girden, E. R. (2001). *Evaluating research articles* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Goldfinch, J. (1994). Further developments in peer assessment of group projects. *Assessment & Evaluation in Higher Education, 19*(1), 29. doi:10.1080/0260293940190103

Goldfinch, J., & Raeside, R. (1990). Development of a peer assessment technique for obtaining individual marks on a group project. *Assessment & Higher Evaluation in Higher Education, 15*(3), 210-231. doi:1080/0260293900150304

Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors: The Journal of the Human Factors and Ergonomics Society, 26*(5), 519-532. doi:10.1177/001872088402600504

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Greeno, J. (1980). Trends in the theory of knowledge for problem solving. In D.T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research* (pp. 9-23). Hillsdale, NJ: Lawrence Erlbaum Associates.

Harkins, Stephen G. & Szymanski, Kate. (1989). Social loafing and group evaluation. *Journal of Personality and Social Psychology, 56*(6), 934-941. Doi:10.1037/0022-3514.56.6.934

Harris, K. R., & Pressley, M. (1991). The nature of cognitive strategy instruction: Interactive strategy construction. *Exceptional Children, 57*(5), 392-404. Retrieved

from http://www.freepatentsonline.com/article/Exceptional-Children/10487825.html

Heller, P., Keith, R., & Anderson, S. (1992). Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Association of Physics Teachers, 60*(7), 627-644. doi:10.1119/1.17117

Hollingshead, A. B. (1998a). Distributed knowledge and transactive processes in decision-making groups. In M.A. Neale, E.A. Mannix, & D.H. Gruenfeld (Eds.), *Research on managing groups and teams* (Vol. 1, pp. 103-123). Greenwich, CN: JAI Press.

Hollingshead, A. B. (1998b). Communication, learning, and retrieval in transactive memory systems. *Journal of Experimental Social Psychology, 34*(5), 423-442. doi:10.1006/jesp.1998.1358

Hollingshead, A. B. (1998c). Retrieval processes in transactive memory systems. *Journal of Experimental Psychology, 34*(5), 423-442. doi:10.1037/0022-3514.74.3.659

Hong, N. S., Jonassen, D. H., & McGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science Teaching, 40*(1), 6-33. doi:10.1002/tea.10058

Hooper, S., & Hannafin, M. J. (1988). Cooperative CBI: The effects of heterogeneous versus homogeneous grouping on the learning of progressively complex concepts. *Journal of Educational Computing Research, 4*(4), 413-424. doi:10.2190/T26C-3FTH-RNYP-TV30

Hooper, S. R., Ward, T. J., Hannafin, M. J., & Clark, H. T. (1989). The effects of aptitude composition on achievement during small group learning. *Journal of Computer-Based Instruction*, 16(3) 102-109.

Huk, T. (2006). Who benefits from learning with 3D models? The case of spatial ability. *Journal of Computer Assisted Learning, 22*, 392-404. doi:10.1111/j.1365-2729.2006.00180.x

Inderrieden, E. J., Allen, R. E., & Keaveny, T. J. 2004. Managerial discretion in the use of self-ratings in an appraisal system: The antecedents and consequences. *Journal of Managerial Issues.* 16: 460-482. Downloaded 26 August, 2013 from http://www.jstor.org/discover/10.2307/40604464?uid=3739936&uid=2&uid=4&uid=3739256&sid=21102584045643

Jacobs, A.E.J.P., Dolmans, D.H.J.M., Wolfhagen, I.H.A.P, & Scherpbier, A.J.J.A (2003). Validation of a short questionnaire to access the degree of complexity and structuredness of PBL problems. *Medical Education, 37*(11), 1001-1007. doi:10.1046/j.1365-2923.2003.01630.x

Johnson, D., Johnson, R., Stanne, M., & Garibaldi, A. (1990). Impact of group processing on achievement in cooperative groups. *Journal of Social Psychology, 130*(4), 507-516. doi:10.1080/00224545.1990.9924613

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin, 89*(1), 47-62. doi:10.1037/0033-2909.89.1.47

Johnston, L., & Miles, L. (2004). Assessing contributions to group assignments. *Assessment & Evaluation in Higher Education, 29*(6), 751-768. doi:10.1080/0260293042000227272

Joint Staff (2011). *JP 5-0 Joint operation planning.* Washington, DC: Government Printing Office.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45*(1), 65-94. doi:10.1007/BF02299613

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*(4), 63-85. doi:10.1007/bf02300500

Jonassen, D. H. (2010). *Learning to solve problems: A handbook for designing problem-solving learning environments.* New York: Routledge.

Jonassen, D. H., & Kwon, H. (2001). Communication patterns in computer mediated versus face-to-face group problems solving. *Educational Technology Research and Development, 49*(1), 35-51. doi:10.1007/BF02504505

Kagan, S. (1995). Group grades miss the mark. *Educational Leadership, 52*, 68-71.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351-371. doi:10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65*(4), 681-706. doi:10.1037/0022-3514.65.4.681

Kester, L., Kirschner, P. A., & Van Merriënboer, J.J G. (2006). Just-in-time information presentation: Improving learning a troubleshooting skill. *Contemporary Educational Psychology, 31*, 167-185. doi:10.1037/0022-3514.65.4.681

Kirschner, F., Paas, F., & Kirschner, P. A. (2008). Individual versus group learning as a function of task complexity: An exploration into the measurement of group cognitive load. In J. Zumbach, N. Schwartz, T. Seufert, & L. Ketser (Eds.), *Beyond knowledge: The legacy of competence* (pp. 21-28). Amsterdam, NE: Springer. doi:10.1007/978-1-4020-8827-8_4

Kirschner, F., Paas, F., & Kirschner, P. A. (2009a). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior, 25*, 306-314. doi:10.1016/j.chb.2008.12.008

Kirschner, F., Paas, F., & Kirschner, P. A. (2009b). Cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review, 21*, 31-42. doi:10.1017/s10648-008-9095-2

Kirschner, F., Paas, F., Kirschner, P. A., & Janssen, J. (2011). Differential effects of problem-solving demands on individual and collaborative learning outcomes. *Learning and Instruction, 21*(4), 587-599. doi:10.1016/j.learninstruc.2011.01.001

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction, 12*, 1-10. doi:20.2026/S0959-4752(01)00014-7

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development, 26*(4), 222-232. doi:10.1159/000272885

Kitchener, K. S., & King, P. M. (1981). Reflective judgment: Concepts of justification and their relationship to age and education. *Journal of Applied Developmental Psychology, 2*(2), 89-116. doi:10.1016/0193-3973(81)90032-0

Knowles, M. S., Holton III, E. F., & Swanson, R. A. (2012). *The adult learner.* Routledge.

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review, 6*(1), 3-34. doi:10.1007/BF00155578

Kotovsky, K., Hayes, J., & Simon, H. (1985). Why are some problems hard? Evidence of Hanoi from tower. *Cognitive Psychology, 294*(2), 248-294. doi:10.1016/0010-0285(85)90009-X

Kuhn, W., & Frank, A. U. (1991). A formalization of metaphors and image-schemas in user interfaces. In D.M. Mark & A.U. Frank (Eds.), *Cognitive and linguistic aspects of geographic space* (Vol. 63, pp. 419-434). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Laughlin, P. R., & Adamopoulos, J. (1980). Social combination processes and individual learning for six-person groups on an intellective task. *Journal of Personality and Social Psychology, 38*, 941-947. doi:10.1037/0022-3514.38.6.941

Laughlin, P. R., Bonner, B. L., & Miner, A. G. (2002). Groups perform better than the best individuals on letters-to-numbers problems. *Organizational Behavior and Human Decision Processes, 88*(2), 605-620. doi:10.1016/S0749-5978(02)00003-1

Laughlin, P. R., Carey, H. R., & Kerr, N. L. (2008). Group-to-individual problem-solving transfer. *Group Processes and Intergroup Relations, 11*(3), 319-330. doi:10.1177/1368430208090645

Laughlin, P. R., Hatch, E. C., Silver, J. S., & Boh, L. (2006). Groups perform better than the best individuals on letters-to-numbers problems: Effects of group size. *Journal of Personality and Social Psychology, 90*, 644-651. doi:10.1037/0022-3514.90.4.644

Laughlin, P. R., Zander, M. L., Knievel, E. M., & Tan, T. K. (2003). Groups perform better than the best individuals on letters-to-numbers problems: Informative equations and effective strategies. *Journal of Personality and Social Psychology, 85*(4), 684-694. doi:10.1037/0022-3514.85.4.684

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life.* Cambridge: Cambridge University Press.

Lejk, M., & Wyvill, M. (1996). A survey of methods of deriving individual grades from group assessments. *Assessment & Evaluation in Higher Education, 21*(3), 267. doi:10.1080/0260293960210306

Lejk, M., & Wyvill, M. (2001). Effect of the inclusion of self-assessment with peer assessment of contributions to a group project: A quantitative study of secret and agreed assessments. *Assessment & Evaluation in Higher Education, 26*(6), 551-561. doi:10.1080/02602930120093887

Lejk, M., & Wyvill, M. (2002). Peer assessment of contributions to a group project: Student attitudes to holistic and category-based approaches. *Assessment & Evaluation in Higher Education, 27*(6), 569-577. doi:10.1080/0260293022000020327

Li, K. (2001). Some refinements on peer assessments of group projects. *Assessment & Evaluation in Higher Education, 26*(1), 5-18. doi:10.1080/0260293002002255

Liang, D. W., Moreland, R., & Argote, L. (1995). Group versus individual training and group performance: the mediating role of transactive memory. *Personality and Social Psychology Bulletin, 21*(4), 384-393. doi:10.1177/0146167295214009

Lohman, M. C., & Finkelstein, M. (2000). Designing groups in problem-based learning to promote problem-solving skill and self-directedness. *Instructional Science, 28*, 291-307. doi:10.1023/A:1003927228005

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of educational research, 66*(4), 423-458. doi:10.3102/00346543066004423

Lyons, R.E., McIntosh, M., & Kysilka, M.L. (2003). *Teaching college in an age of accountability*. Boston: Allyn and Bacon.

Mathematical Sciences Education Board, National Research Council (1993). *Measuring up: Prototypes for mathematical assessment*. Washington, DC: National Academy Press.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science, 26*(1/2), 49-63. doi:10.1023/A:1003088013286

Meacham, J. A., & Emont, N. C. (1989). The interpersonal basis of everyday problem solving. In J.D. Sinnott (Ed.), *Everyday problem solving: Theory and applications* (pp. 7-23). Hillsdale, NJ: Praeger.

Meudell, P. R., Hitch, G. J., & Kirby, P. (1992). Are two heads better than one? Experimental investigations of the social facilitation of memory. *Applied Cognitive Psychology, 6*, 525-543. doi:10.1002/acp.2350060606

Michaelsen, L. K. (1992). Team learning: A comprehensive approach for harnessing the power of small groups in higher education. *To Improve the Academy, 11*, 107-122.

Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: A sourcebook for new methods.* 2nd ed. Newbury Park, CA: Sage.

Miller, G. (1956). Magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97. doi:10.1037/h0043158

Millis, B. J., Cottell, P. G., Jr., & American Council on Education, W. D. C. (1997). *Cooperative Learning for Higher Education Faculty. Series on Higher Education.*

Moreland, R. L., & Levine, J. M. (1992). Problem identification by groups. In S. Worchel, W. Wood, & J.A. Simpson (Eds.), *Group processes and productivity.* Newbury Park, CA: Sage.

Moreno, R. (2010). Cognitive load theory: More food for thought. *Instructional Science, 38*, 135-141. doi:10.1007/s11251-009-9122-9

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(30), 309-326. doi:10.1007/s10648-007-9047-2

Morris, K. V., Finnegan, C., & Sz-Shyan, W. (2005). Tracking student behavior, persistence, and achievement in online courses. *Internet and Higher Education, 8*(3), 221-231. doi:10.1016/j.iheduc.2005.06.009

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology, 8*(2), 181-192. doi:10.1002/ejsp.2420080204

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.

Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science* (Report R-3928-NSF). Santa Monica, CA: Rand Corp. retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED329615&ERICExtSearch_SearchType_0=no&accno=ED329615

Oakley, B., Felder, R. M., Brent, R., & Elhajj, I. 2004. Turning student groups into effective teams. *Journal of Student-Centered Learning, 2*: 9-34.

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., . . . Schmucker, D. G. (2012). The Comprehensive Assessment of Team Member Effectiveness: Development of a Behaviorally Anchored Rating Scale for Self- and Peer Evaluation. *Academy of Management Learning & Education, 11*(4), 609-630. Doi:10.5465/amle.2010.0177

Ohtsubo, Y. (2005). Should information be redundantly distributed among group members? Effective use of group memory in collaborative problem solving. *Applied Cognitive Psychology, 19*, 1219-1233. doi:10.1002/acp.1162

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skills in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429-434. doi:10.1037/0022-0663.84.4.429

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1-4. doi:10.1007/s10648-005-3951-0

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P.W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-72. doi:10.1207/S15326985EP3801_8

Paas, F., Van Gerven, P.W M., & Wouters, P. (2007). Instructional efficiency of animation: Effects of interactivity through mental reconstruction of static key frames. *Applied Cognitive Psychology, 21*(6), 783-793. doi:10.1002/acp.1349

Paas, F., Van Merrienboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*, 419-430.

Patton, M. Q. (2002). *Qualitative research and evaluation methods (3rd ed.).* Thousand Oaks, Ca: Sage.

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 83-106). Hillsdale, NJ: Lawrence Erlbaum Associates.

Peterson, P. L., Janicki, T. C., & Swing, S. R. (1981). Ability x treatment interaction effects on children's learning in large-group and small-group approaches.

American Educational Research Journal, 18, 453-473. doi:10.3102/00028312018004453

Peterson, L., & Peterson, M. (1959). Short-term retention of individual verbal items. *Journal of Educational Psychology, 58*, 193-198. doi:10.1037/h0043158

Pfaff, E., & Huddleston, P. 2003. Does it matter if I hate teamwork? What impacts student attitudes toward teamwork. *Journal of Marketing Education.* 25: 37-45. Doi:10.1177/0273475302250571

Pociask, F. D., & Morrison, G. R. (2008). Controlling split attention and redundancy in physical therapy instruction. *Educational Technology Research and Development, 56*(4), 379-399. doi:10.1007/s11423-007-9062-5

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12*(1), 61-86. doi:10.1016/S0959-4752(01)00016-0

Race, P. (2001). A briefing on self, peer and group assessment. *Assessment series, 9*, 12-12.

Rafiq, Y., & Fullerton, H. (1996). Peer assessment of group projects in civil engineering. *Assessment & Evaluation in Higher Education, 21*, 69-81.

Rogoff, B., & Lave, J. (1984). *Everyday cognition: Its development in social context.* Cambridge, MA: Harvard University Press.

Saavedra, R. & Kwun, S. K. 1993. Peer evaluation in self-managing work groups. *Journal of Applied Psychology.* 78: 450-462. Doi:10.1037//0021-9010.78.3.450

Salden, R.J. C.M., Paas, F., Broers, N. J., & Van Merriënboer, J.J. G. (2004). Mental effort and performance as determinants for the dynamic selection of learning tasks

in air traffic control training. *Instructional Science, 32*(1/2), 153-172. doi: 10.1023/B:TRUC.0000021814.03996.ff

Schoenfeld, A. H. (1989). Ideas in the air: Speculations on small group learning, environmental and cultural influences on cognition and epistemology. *International Journal of Education, 13*(1), 71-88. doi:10.1016/0883-0355(89)90017-7

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well defined and ill defined problem solving. *Applied Cognitive Psychology, 9*(6), 523-538. doi:10.1002/acp.2350090605

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences, 4*, 321-354. doi:10.1207/s15327809jls0403_3

Sheppard, S., Chen, H. L., Schaeffer, E., Steinbeck, R., Neumann, H., & Ko, P. 2004. Peer assessment of student collaborative processes in undergraduate engineering education. Final Report to the National Science Foundation, Award Number 0206820, NSF Program 7431 CCLI-ASA.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428. Doi:10.1037/0033-2909.86.2.420

Skon, L., Johnson, D. W., & Johnson, R. T. (1981). Cooperative peer interaction versus individual competition and individualistic efforts: Effects on the acquisition of cognitive reasoning strategies. *Journal of Educational Psychology, 73*, 83-92. doi:10.1037/0022-0663.73.1.83

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.). Boston, MA: Allyn & Bacon.

Smith, M. (1991). A view from biology. In M.U. Smith (Ed.), *Toward a unified theory of problem solving* (pp. 1-20). Hillsdale, NJ: Lawrence Erlbaum Associates.

Somervell, H. (1993). Issues in assessment, enterprise and higher education: the case for self-, peer and collaborative assessment. *Assessment & Evaluation in Higher Education, 18*, 221-233.

Spering, M., Wagener, D., & Funke, J. (2005). The role of emotions in complex problem solving. *Cognition and Emotion, 19*(8), 1252-1261. doi:10.1080/02699930500304886

Stark, R., Mandl, H., Gruber, H., & Renkl, A. (2002). Conditions and effects of example elaboration. *Learning and Instruction, 12*(1), 39-60. doi:10.1016/S0959-4752(01)00015-9

Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology, 31*, 244-265. doi:10.1006/jesp.1995.1012

Stasson, M. F., & Hawkes, W. G. (1995). Effect of group performance on subsequent individual performance: Does influence generalize beyond the issues discussed by the group. *Psychological Science, 6*(5), 305-307. doi:10.1111/j.1467-9280.1995.tb00515.x

Stephenson, G. M., Clark, N. K., & Wade, G. S. (1986). Meetings make evidence? An experimental study of collaborative and individual recall of a simulated police interrogation. *Journal of Personality and Social Psychology, 50*, 1113-1122.

Stroebe, W., & Diehl, M. (1994). Why groups are less effective than their members: On productivity losses in idea-generating groups. *European Review of Social Psychology, 5*(1), 271-303. doi:10.1080/14792779543000084

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. doi:10.1207/s15516709cog1202_4

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory.* New York, NY: Springer.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*(4), 351-362. doi:10.1207/s1532690xci0804_5

Sweller, J., Van Merrienboer, J. J.g., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-295. doi:10.1023/A:1022193728205

Szymanski, Kate & Harkins, Stephen G. (1987). Social loafing and self-evaluation with a social standard. *Journal of Personality and Social Psychology, 53*(5), 891-897. Doi:10.1037/0022.3514.53.5.891

Tabbers, H. K., Martens, R. L., & Van Merriënboer, J.J G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology, 74*(1), 71-81. doi:10.1348/000709904322848824

Taggar, S., & Brown, T. C. 2006. Interpersonal affect and peer rating bias in teams. *Small Group Research,* 37: 87-111. Doi: 10.1177/1046496405284382

Thorndike, R. M. (2005). *Measurement and Evaluation in Psychology and Education* (7th ed.). Upper Saddle River, NJ: Pearson.

Tudge, J. (1989). When collaboration leads to regression: Some negative consequences of socio-cognitive conflict. *European Journal of Social Psychology, 19,* 123-138. doi:10.1002/ejsp.2420190204

Van Gog, T., & Paas, F. (2008). Revisiting the original construct in educational research. *Educational Psychologist, 43*(1), 16-26. doi:10.1080/00461520701756248

Van Merriënboer, J.J G., Schuurman, J. G., De Croock, M.B. M., & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction, 12*(1), 11-37.

Vogt, W. P. (2007). *Quantitative research methods for professionals.* Boston, MA: Pearson.

Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M.T.H. Chi, R. Glaser, & M.J. Farr (Eds.), *The nature of expertise* (pp. 261-285). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wales, C. E., & Stager, R. A. (1977). *Guided design.* Morgantown, WV: West Virginia Center for Guided Design.

Webb, N. M. (1980). A process-outcome analysis of learning in group and individual settings. *Educational Psychologist, 15,* 69-83. doi: 10.1080/00461528009529217

Webb, N. M. (1982). Peer interaction and learning in cooperative small groups. *Journal of Educational Psychology, 74,* 642-655. doi:10.1037/0022-0663.74.5.642

Webb, N. M. (1984). Stability of small group interaction and achievement over time. *Journal of Educational Psychology, 76*, 211-224. doi:10.1037/0022-0663.76.2.211

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, 366-389. doi:10.2307/749186

Webb, N. M. (1993). Collaborative group versus individual assessment in mathematics: Processes and outcomes. *Educational Assessment, 1*(2), 131-152. doi:10.1207/s15326977ea0102_3

Webb, N. M. (1997). Assessing students in small collaborative groups. *Theory into practice, 36*(4), 205-213. doi:10.1080/00405849709543770

Webb, N. M., & Cullian, L. K. (1983). Group interaction and achievement in small groups: Stability over time. *American Educational Research Journal, 20*(3), 411-423. Doi:10.3102/00028312020003411

Webb, N. M., & Kenderski, C. M. (1984).Student interaction and learning in small group and whole class settings. In P.L. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction. Group organization and group processes* (pp. 153-170).New York: Academic Press.

Webb, N., & Palincsar, A. S. (1996). Group processes in the classroom. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of educational psychology* (pp. 841-873). New York, NY: Simon and Schuster Macmillan.

Wegner, D. (1978). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). New York, NY: Springer-Verlag.

Weldon, M. S., & Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1160-1175.

Wertheimer, M. (1959). *Productive thinking* (1st ed.). Chicago, IL: University of Chicago Press.

Wiley, J., & Jensen, M. S. (2006). When three heads are better than two. In R. Sun (Ed.), *28th Annual Conference of the Cognitive Society* (pp. 2375-2380). Mahway, NJ: Lawrence Erlbaum Associates.

Williams, K., Harkins, S., & Latane, B. (1981). Identifiability as a deterrent to social loafing: Two cheering experiments. *Journal of Personality and Social Psychology, 40*, 303-311. Doi: 10.1037/0022-3514.40.2.303

Wood, P. K. (1983). Inquiring systems and problem structure: Implications for cognitive development. *Human Development, 26*(5), 249-296. doi:10.1159/000272887

Wood, P. K. (1985). The statistical examination of necessary but not sufficient antecedents of problem solving behavior. *Dissertation Abstracts International: Section A. Humanities and Social Sciences.*

Wouters, P., Paas, F., & Van Merriënboer, J.J G. (2009). Observational learning from animated models: Effects of modality and reflection on transfer. *Contemporary Educational Psychology, 34*(1), 1-8. doi:10.1016/j.cedpsych.2008.03.001

Yamane, D. (1996). Collaboration and its discontents: Steps toward overcoming barriers to successful group projects. *Teaching Sociology, 24*, 378-383. Retrieved June 22, 2012, from http://www.jstor.org/stable/1318875

Young, C. B., & Henquinet, J. A. 2000. A conceptual framework for designing group projects. *Journal of Education for* Business, 76: 56-60. Doi: 10.1080/08832320009599051

Zhang, B., Johnston, L., & Kilic, G. (2008). Assessing the reliability of self- and peer rating in student group work. *Assessment & Evaluation in Higher Education, 33*(3), 329-340. doi:10.1080/02602930701293181

Zhang, B., & Ohland, M.W. 2009. How to assign individualized scores on a group project: An empirical evaluation. *Applied Measurement in Education, 22:* 290-308. Doi: 10.1080/08957340902984075

Zheng, R., McAlack, M., Wilmes, B., Kohler-Evans, P., & Williamson, J. (2009). Effects of multimedia on cognitive load, self-efficacy, and multiple rule-based problem solving. *British Journal of Educational Psychology, 40*(5), 790-803. doi:10.1111/j.1467-8535.2008.00859.x

Appendix A: Informed Consent Form

PROJECT TITLE: The Effects of Group Size on Student Learning, Participation, Mental Effort, and Group Outcomes for Middle-Aged Adults Working in an Ill-Structured Problem-Solving Environment

RESEARCHERS

Ginger S. Watson, Ph.D., **Responsible Project Investigator**
Associate Professor
STEM Education and Professional Studies
Darden College of Education, Old Dominion University
gswatson@odu.edu

Gary Roemmich, M.S.
Doctoral Student
STEM Education and Professional Studies
Darden College of Education, Old Dominion University
roemmichg@ndu.edu

PURPOSE

The purposes of this consent form are to provide you information that may affect your decision whether to participate in this research, and to record the consent of those who agree to participate. Several studies have investigated group problem-solving; however, few studies have looked at group size on the learning and performance outcomes of adults solving complex problems. The study will examine how group size impacts problem solving. Approximately 200 individuals will participate in this study.

If you decide to participate, you will be assigned to a group and asked to complete the group assignments (i.e., contingency planning mission analysis) and the end of course test normally required in this course. In addition, at the end of the instruction you will be asked to complete a survey describing your perceived participation and that of other members of your group. You will also be asked to report your mental effort periodically throughout the exercise. Recording your mental effort should take about 5 minutes each time you are asked and completion of the participation survey at the end of the study will take about 10 minutes. No additional time (outside of the regularly scheduled class session) will be required of you for any of the activities associated with this research.

RISKS AND BENEFITS

RISKS: The risk of participation in the study is relatively minor. Participation in the study will have no impact on class standing or grades. All scores/data collected will remain confidential with no feedback to the college or students provided. As with any research, there is some possibility that you may be subject to risks that have not been identified.

BENEFITS: There is no direct benefit to you for participating in this study. However, future students may benefit from being placed into groups that are optimally sized to promote learning.

COSTS AND PAYMENTS

There is no cost to you for participating in this research. Similarly, you will not be paid to participate in this research.

NEW INFORMATION

If the researchers find new information during this study that would reasonably change your decision about participating, then they will give it to you.

CONFIDENTIALITY

The researchers will take reasonable steps to keep private information confidential. The researchers will not share identifiable information collected during this study with anyone outside of the research team. This survey and all study data will be kept confidential and all your data will coded with a unique identifier instead of your name so that your name will not be directly linked to your responses. Encrypted data transmission technology will be used for security purposes. The results of this study may be used in reports, presentations, publications, and a dissertation. However, the researchers will not identify you. The data collected will be compiled into groups and then compared. Individual scores and marks will not be part of the dissertation or resulting publications. Of course, your records may be subpoenaed by court order or inspected by government bodies with oversight authority.

WITHDRAWAL PRIVILEGE

It is OK for you to say NO. Even if you agree to participate now, you are free to walk away or withdraw from the study at any time. Your decision will not affect your relationship with your instructor or university, or otherwise cause a loss of benefits to which you might otherwise be entitled, except extra credit. The researchers reserve the right to withdraw your participation in this study, at any time, if they observe potential problems with your continued participation.

COMPENSATION FOR ILLNESS AND INJURY

If you agree to participate, your consent in this document does not waive any of your legal rights. However, in the event of harm arising from this study, neither Old Dominion University nor the researchers are able to give you any money, insurance coverage, free medical care, or any other compensation for such injury. In the event that you suffer injury as a result of participation in any research project, you may contact Dr. Ginger Watson (gwatson@odu.edu) or 757-683-3246, Dr. George Maihafer the current IRB chair at 757-683 4520 at Old Dominion University, or the Old Dominion University Office of Research at 757-683-3460 who will be glad to review the matter with you.

VOLUNTARY CONSENT

By agreeing to be a part of this study, you are saying several things. You are saying that you have read this form or have had it read to you, that you have understood this form, the research study, and its risks and benefits. Feel free to keep a copy of this form for your records. The researchers should have answered any questions you may have had about the research. If you have any additional questions later on, then you can contact Dr. Ginger Watson (gswatson@odu.edu or 757-683-3246). If at any time you feel pressured to participate, or if you have any questions about your rights or this form, then you should contact Dr. George Maihafer the current IRB chair at 757-683 4520 at Old Dominion University, or the Old Dominion University Office of Research at 757-683-3460 .

If you do not agree to be a part of this study, then simply return this form to the researcher. If you agree to be a part of this study, then put your name, signature and today's date on the line below.

Participant's
name____________________________Signature___________________________
Date________

INVESTIGATOR'S STATEMENT

I certify that I have explained to this subject the nature and purpose of this research, including benefits, risks, costs, and any experimental procedures. I have described the rights and protections afforded to human subjects and have done nothing to pressure, coerce, or falsely entice this subject into participating. I am aware of my obligations under state and federal laws, and promise compliance. I have answered the subject's questions and have encouraged him/her to ask additional questions at any time during the course of this study.

Investigator's Printed Name & Signature	**Date**

Appendix B: Post-Lesson Test

For each question below, select the most appropriate option for each question.

1. Which of the below should not be considered as a fact?

   *a. Tribal influences, particularly Tuareg, will contribute to regional instability.

   b. Each nation will be responsible for the logistical support of its own forces.

   c. The regional balance of air power, prior to U.S. intervention, belongs to the RNA in terms of capability and capacity.

   d. Tunisia has agreed to the US use of airports and seaports.

   e. The US has a bilateral defense agreement with Tunisia.

2. Which of the assumptions listed below would be considered to be invalid?

   a. Current strategic partners will allow the US to conduct operations from sovereign territories.

   b. Tunisia will be able to secure/maintain control of ALOCs and SLOCs into the country prior to hostilities.

   c. Indicators and warnings will allow the U.S./Tunisian militaries 6 days of early warning of an impending RNA attack.

   d. Mali, Mauritania, and Morocco will abide by their bilateral agreements and support U.S. /Tunisia effort against the RNA (provide necessary support)

   *e. U.S. will be granted overflight, landing and basing rights.

3. For the assumption identified in question 2, what makes it invalid?

   a. The assumption is not logical.

   *b. The assumption is not realistic.

   c. The assumption is not essential to the planning effort.

   d. The assumption does not address a gap in knowledge.

   e. The item is a fact and not an assumption.

4. The U.S. – Tunisian Defense Treaty agrees to the introduction of American forces into Tunisia in the event of external aggression or subversion. Which of the below objectives does not support this agreement as it fails to remove the external threat or subversion from Tunisian territory?

a. Tuareg Liberation Front (TLF) and its associated networks are defeated.

b. Tunisian counter-terrorist forces deny extremist organizational safe havens from which to operate.

*c. Peace operations capacity exists to respond to emerging crises in designated states.

c. Tunisia has a viable military capable of self-security.

5. Which strategic end state appears to be unsupported by the military end states?

a. Notional Tunisian territorial integrity is maintained or restored.

b. Notional Tunisian government is capable of performing defense, security and requisite humanitarian operations

c. Aggressor nation/organizations offensive capability is reduced to acceptable self-defense levels

*d. Regional stability is restored to pre-conflict levels.

6. Which of the below is not a valid limitation for the Tunisian Defense Plan?

a. The US cannot initiate combat operations.

b. The plan must use a whole-of-government approach

*c. The US must continue TCP funding for critical partners for the duration of the campaign.

d. The US must reduce the ability of the RNA to conduct offensive operations.

7. Which essential task is not specifically addressed in the mission statement?

*a. Defend Tunisian territorial integrity.

b. Support the effort of other U.S. Departments and Agencies.

c. Protect designated Tunisian critical infrastructure/assets.

d. Information Operations Plan to support Strategic Communication efforts to isolate/deter RNA aggression.

Mission Statement: USEASTCOM, in conjunction with other government agencies, will deter RNA aggression with Tunisia. USEASTCOM Forces will engage hostile forces threatening Tunisia and its territorial waters in order to maintain Tunisian territorial integrity by supporting Inter-agency efforts, protecting designated Tunisian critical infrastructure/assets, and enabling Strategic Communications in order to isolate and deter aggression.

8. The mission statement is incomplete because it is lacking which component(s)?

a. Why

b. What

*c. When

d. Where

e. How

9. Which of the below essential tasks are not supported in the mission statement?

a. Defend Tunisian territorial integrity

b. Support the efforts of other U.S. Departments and Agencies

c. Strategic Communications plan to isolate/deter RNA aggression

d. Protect critical Tunisian infrastructure

*e. All essential tasks are included in the mission statement

10. The CCIR's listed fail to address a key information deficiency. Which of the below explanations addresses that deficiency?

a. Support important information the commander will require for decision-making.

*b. Support resolution of assumptions.

c. Do not link to the strategic end states.

d. Address operational limitations

11. The force structure analysis listed capabilities needed for each phase of the operation. Based on their list, what statement can you make regarding their opinion regarding the availability of forces.

a. There are sufficient forces available to conduct the operation.

b. There are insufficient forces available for the operation.

*c. There are insufficient details available to make a determination.

12. Which enemy capability appears to not be addressed by the initial force analysis?

a. Airborne forces

b. Special Operations forces

*c. Minelaying forces

d. ICBM forces

13. They identified four risks that they determined to be of marginal severity. Of those four, which risk, if unmitigated, did they estimate had the lowest probability of occurring?

a. Loss of access to Tunisia

b. Loss of critical infrastructure

*c. Tunisian military capitulates

d. Humanitarian crisis

14. They identified the risk of a Humanitarian Crisis was likely and attempted to mitigate with force flow adjustments and Information Operations. Their table showed no

effects from the mitigation. Which of the below is most likely to reduce the impact of a humanitarian crisis?

a. Increased ISR.

*b. Support for NGO involvement.

c. Diplomacy.

d. Mil to Mil engagement.

15. The identified enemy operational center of gravity is the RNA East Operational Strategic Command. Which of the below is a critical capability of this center of gravity?

a. African Union.

b. Port Access.

c. RNA Governmental control.

*d. Command and Control.

16. The Tunisian Defense Forces have been identified as a critical capability. Which of the below is not a critical vulnerability of the Tunisian Defense Forces?

a. Corruption of Tunisian Officials.

b. Outdated/inferior equipment.

*c. Strategic partnership support.

d. Insurgency activity within Tunisia.

17. In the initial Operational Approach, which end state appears to have little support with which to achieve the desired conditions?

*a. RNA forces in Tunisian territory are defeated.

b. The Government of Tunisia reassumes performing security and humanitarian operations.

c. Establishment of exclusion zone and international monitoring capability.

d. All'military end states are adequately supported within the initial operational approach framework.

18. The initial operational approach has three end states listed. Each end state has supporting objectives assigned that should allow attainment. Of the three end states, only one appears to have objectives which will support attainment. Which end state appears to have sufficient coverage by objectives to support its attainment?

a. RNA forces in Tunisian territory defeated.

*b. The Government of Tunisia military resumes performing defensive security and humanitarian operations.

c. Establishment of exclusion zone and international monitoring capability in place.

19. Which objective displayed in the Initial Operational Approach is not supported by its associated effects?

a. Movement of ships along regional SLOCs remains unimpeded.

*b. Aggressor force reduced to offensive combat ineffective.

c. Tunisia has a viable military capable of self-security?

d. Tunisian counter-terrorist forces deny extremist organizational safe havens from which to operate.

20. Which of the below listed restraints should not be listed as a restraint?

a. Cannot initiate combat operations.

b. No use of chemical/biological weapons

c. Conduct offensive mining operations.

*d. Land.

* Denotes correct answer

Appendix C: Post-Lesson Test Table of Specifications

Table C1

Table of Specifications

Content Area	Nr Q's	%	Bloom's Taxonomy	Objectives
Known - Facts/ Conditions	**1**	**5**	**Analyze**	**DETERMINE known facts and develop planning assumptions**
Assumptions	**2**	**10**	**Evaluation**	**DISTINGUISH facts, assumptions, limitations and specified tasks in strategic guidance applicable to USEASTCOM's Theater Campaign Plan.**
Objectives	**1**	**5**	**Synthesis**	**PRODUCE Theater Strategic Objectives and related effects that consider ends ways and means to describe the desired conditions for the fictitious USEASTCOM Theater Campaign Plan.**
End State	**1**	**5**	**Analyze**	**ANALYZE policy and strategic guidance to determine national interests, end states, and assigned objectives for the fictitious USEASTCOM AOR.**
Limitations	**1**	**5**	**Analyze**	**DISTINGUISH facts, assumptions, limitations and specified tasks in strategic guidance applicable to USEASTCOM's Theater Campaign Plan.**
Theater Tasks	**1**	**5**	**Evaluation**	**DETERMINE specified, implied, and essential tasks**
Mission Statement	**2**	**10**	**Synthesis**	**DEVELOP a mission statement**
CCIR's	**1**	**5**	**Analyze**	**DETERMINE initial commander's critical information requirements**
Force Structure Analysis	**2**	**10**	**Evaluation**	**IDENTIFY available resources and shortfalls from a comprehensive perspective.**
Strategic Risk Assessment	**2**	**10**	**Evaluation**	**DEMONSTRATE an understanding of strategic risk in an ends ways means risk framework, and the principles of risk assessment**

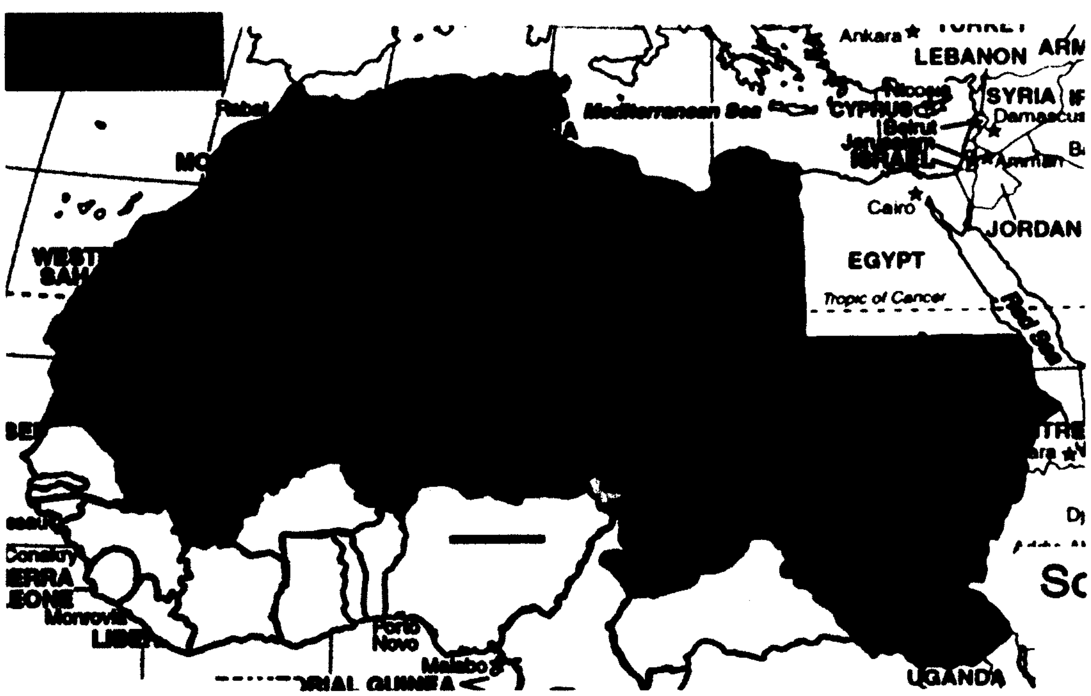
Center of Gravity	**2**	**10**	**Evaluation**	**DETERMINE the termination criteria, and friendly/adversary military end states, strategic military objectives and centers of gravity.**
Initial Operational Approach	**2**	**10**	**Synthesis**	**Create an initial operational approach.**
Refined Operational Approach	**2**	**10**	**Evaluation**	**Refine an initial operational approach.**
Total	**20**	**100**		

Appendix D: Post-Lesson Test Background Information

The information below has been extracted from the draft mission analysis and operational design brief prepared by the Defense of Tunisia Joint Planning Group (JPG) at the USEASTCOM Staff. The JPG Team lead has asked that you provide an outside assessment of their efforts before they brief their Commander. Unfortunately, the JPG was comprised of few individuals with planning experience at the Geographic Combatant Commander (GCC) level and the JPG lead is concerned the group may have lost focus and feels he is too close to the process to make an honest assessment. Review the materials and be prepared to answer a series of questions with regards to the quality and direction of the following material.

The material below is based on a fictional scenario with personalities, intentions, and visions are constructed to present certain planning challenges and do not reflect real people or national aims and as such it is to be used only educational purposes.

1. North Africa Regional Assessment
    a. Background germane to strategy – All countries except for Libya were French territories until the late 1950s/ early 1960s.
    b. Region is predominately stable
    c. Tensions between RNA and Tunisia over oil fields and terrorism
        i. Potential for Tuareg unrest in S. Tunisia; suspected ties to insurgency & RNA support
        ii. Regional Policy: Maintain regional stability, prevent spread of violent extremism and reduce human suffering, (spread of HIV/Aids, eliminate human trafficking, and prevent inhumane treatment of refugees)
    d. US Regional Interests (security, economic, development, etc.) Intensity-vital



e. Factors
    i. External Factors: Chinese interest in oil reserves; terrorist organizations seeking new safe havens;
    ii. Regional Economic and Political Factors: Oil reserves, corruption, refugees, ethnic diversity
    iii. Regional Transnational Threats: Terrorist safe havens
f. Critical Partners
    i. Tunisia
g. Key Supporting partners
    i. Mali
    ii. Mauritania
    iii. Morocco
h. Actors of Concern
    i. RNA
    ii. Western Sahara (ties to RNA)

2. Area of Operations/Area of Interest

3. What are the facts?
    a. Ethnic, demographic, environmental and health problems in N. Africa will continue to worsen and will be exploited by corrupt politicians, VEOs, insurgents and criminals unless there are credible solutions. (GEF, pg 32)
    b. The African Union will become more prominent as the primary continental security organization in Africa; however some member states may be unlikely to subordinate their security policies for the next five years. (GEF, pg 32)
    c. Tribal influences, particularly Tuareg, will contribute to regional instability. (GEF, pg 32)

d. Libya will eventually continue to normalize relations with the United States. (GEF, pg 32)
e. U.S. has a bilateral defense agreement with Tunisia (CDRs thought, pg 3)
f. RNA sees the southern Tunisia situation as critical to their interests in terms of the El Borma Pipeline corridor. (CDRs thought, pg 3)
g. RNA is the only direct antagonist in the defense of Tunisia and it initiates conflict or overt warfare (JIPOE).
h. RNA coastal Navy normally stays within 100nm of the borders. (JIPOE)
    i. Limited Capes to conduct coordinated surface/subsurface attacks within the RNA coastal waters and Tunisian waters west of BON and very limited operations east of CAPE BON (TUNISIA.)
i. Tunisian government has agreed to U.S. use of the airfields and seaports. (DOD Memo)
    i. Need to retain 33% of each port's throughput capacity to maintain essential civil support functions
j. Air Domain: regional balance of power in air capability belongs to the RNA in terms of capability and capacity (without U.S. intervention.)
k. U.S. air capability will be dependent on basing in TUNIS, MORON, ROTA, SIGONELLA, AVIANO as well as the availability of a CSG in the MED. (JIPOE)
l. Each nation will be responsible for the logistical support of its own forces (DOD Memo)

4. JIPOE
    a. Threat of reduced U.S. influence and access into North Africa
    b. RNA Borders secured physically and economically
    c. Counter influence of Islamic Violent Extremist Organizations
    d. Economic dominance of north Africa hydrocarbons industry
    e. RNA leadership in the Tuareg confederations
    f. Champion Tuareg identity and freedom in north Africa and the world
    g. Preclude US from intervening in north Africa, or force its departure before achieving goals
    h. Consolidate southern Tunisia Tuareg population under RNA control
    i. Control EL BORMA OIL FIELDs.
    j. Reduce/eliminate AL-JUCKS control of southern Tunisia

5. Analyze the task and intent; prepare a CONPLAN w/TPFDD for the unilateral defense of Tunisia in case of external aggression and in support of US interests.
    a. <u>Identify the tasks.</u>
        i. Defend Tunisian territorial integrity (GEF/CDR Gdc)
        ii. Maintain a legitimate GOT (GEF)
        iii. Aggressor capability reduced to defensive posture only (GEF)
        iv. Restore regional stability to pre-combat level (GEF)
        v. US credibility in region maintained/increased (CDR Gdc)
        vi. Maintain/strengthen regional partnership (CDR Gdc)
        vii. US/Tunisian forces operate as a combined force (TDM)

viii. US Forces will deploy in Tunisia (TDM)
ix. Develop/Maintain TPFDD (GEF)
x. US Forces will provide their own logistics support (TDM)
xi. Support the effort of other US departments and agencies (JSCP)
xii. Develop options to evacuate or protect US Citizens/allied citizens (JSCP)
xiii. Protect designated Tunisian critical infrastructure/assets (JSCP)
xiv. Support FDO across the DIME instruments of national power (JSCP)
xv. Provide humanitarian assistance (JSCP)
xvi. Conduct phased transition of responsibilities to International Force/Tunisian Security Forces (JSCP/TDM)

b. Implied (Team)
i. Develop logistics support plan
ii. Establish joint mil/civ operations centers @ all PODs ensuring TDM supported
iii. Coordinate open LOC in EUCOM AOR
iv. IO plan to support Strategic Communications to isolate/deter RNA aggression
v. Protect maritime LOCs
vi. Enable/Reinforce diplomatic efforts
vii. Develop interoperability training program with TDF/regional militaries
viii. Establish ID/Coordination CSL in regional centers

c. Determine which tasks are essential.
i. Defend Tunisian territorial integrity
ii. Support the effort of other US departments and agencies
iii. Protect designated Tunisian critical infrastructure/assets
iv. IO plan to support Strategic Communications to isolate/deter RNA aggression

6. What are the operational limitations (constraints/restraints)?

a. What must we do? (Constraints)
i. Whole of gov't approach (GEF/CDR Guidance)
ii. Must continue TCP throughout AOR
iii. TCP Funding for Critical Partners (GEF)
iv. Phase 0 train/equip/advise (GEF)
v. Establish intelligence sharing agreement (GEF)
vi. Provide HA until USAID/GOT assume responsibility (JSCP)
vii. Develop NEO Plan (JSCP)
viii. GOT integration includes LFO, Mil, SDF
ix. Limit Aggressor ability to conduct further offensive ops (GEF)

b. What can't we do? (Restraints)
i. Leahy Amendment (Law)
ii. Force Apportionment (GEF)
iii. International Law of War

   iv. Maritime
   v. Land
   vi. Cannot initiate offensive ops
   vii. No chemical/biological weapons use

7. What are the assumptions necessary for planning? (Develop assumptions for the missing facts.)
   a. Mali, Mauritania, and Morocco will abide by their bilateral agreements and support U.S. /Tunisia effort against the RNA (provide necessary support)
   b. Tunisia will be able to secure/maintain control of ALOCs and SLOCs into the country pre-hostilities.
   c. Indicators and warnings will allow the U.S./Tunisian militaries 6 days of early warning of an impending RNA attack (JIPOE)
   d. Tunisia can successfully execute its defensive plan in the event of a RNA attack (JIPOE)
   e. Diplomatic functions will fail and military action will be required
   f. Key Allies and International & African agencies will support US/Tunisian actions against aggression from the RNA & insurgent organizations
   g. Current strategic partners will allow US to conduct operations from sovereign territories in event of a conflict in Tunisia.
   h. U.S. will be granted overflight, landing and basing rights

8. What are the strategic objective/s?
   a. Mil End State 1: RNA Forces in Tunisian Territory defeated
      i. OBJ #10: Movement of ships along regional SLOCs remains unimpeded
      ii. OBJ #12: Aggressor Force reduced to offensive combat ineffective
   b. Mil End State 2: GOT military resumption of Defense security & Humanitarian operations
      i. OBJ #2: Tunisian forces able to provide populace with safety and basic social goods
      ii. OBJ #3: Tunisian counter-terrorist forces deny extremist organizations safe havens from which to operate
      iii. OBJ #4: Tunisia has viable military capable of self-security
      iv. OBJ #5: Tuareg Liberation Front (TLF) and its associated networks are defeated
      v. OBJ #7: Security forces in Tunisia are professionalized
   c. Mil End State 3: Establishment of exclusion zone and international monitoring capability
      i. OBJ #11: Peace operation capacity exists to respond to emerging crises in designated states

9. Military Objectives/Effects
   a. Movement of ships along regional SLOCs remains unimpeded

i. Tunisian maritime economic exclusion zone is enforced
ii. Tunisian ports remain open and friendly to US

b. Aggressor Force reduced to offensive combat ineffective
c. Tunisian forces able to provide populace with safety and basic social goods
    i. Tunisian security forces are effective at providing citizens safety against violence and crime
    ii. Tunisian security forces and/or the military apply discriminate law enforcement and counter-terrorism responses towards security threats
d. Tunisian counter-terrorist forces deny extremist organizational safe havens from which to operate
    i. Tunisian security forces kill or capture known terrorists within Tunisia
    ii. Tunisian security forces enforce laws and deny safe havens to extremists
e. Tunisia has viable military capable of self-security
    i. GOT invest in the growth of their military
    ii. GOT invest in the training of the military forces
f. Tuareg Liberation Front (TLF) and its associated networks are defeated
    i. TLF influence within Tunisia is negated
    ii. TLF is unable to conduct terrorist activities within Tunisia
g. Security forces in Tunisia are professionalized
    i. Tunisian security forces execute their mission essential tasks
    ii. Tunisian security forces enforce rules of law and recognized international human rights
h. Peace operations capacity exists to respond to emerging crises in designated states
    i. Tunisian security forces enforce the rule of law and recognize international human rights
    ii. GOT supports global peace operations initiative (CPOI)
    iii. GOT supports

10. What is the military end state?
    a. RNA Forces in Tunisian Territory defeated
    b. GOT military reassume of performing Defense security & Humanitarian operations
    c. Establishment of exclusion zone and international monitoring capability

11. What are the termination criteria
    a. ES 1: Tunisian Territory Restored
        i. Aggressor forces removed from Tunisian territory
        ii. Land/Air/Maritime military exclusion zone established/recognized
    b. ES 2: GOT capability of Defense Security/Humanitarian Ops
        i. Command structure restored
        ii. Capability to conduct and maintain border/internal security

        iii. Infrastructure developed & improvement planned and initiated
    c. ES 3: Aggressor reduced to self-defense capability
        i. RNA forces reduce to offensive combat ineffective status
        ii. RNA forces denied access into capable range of GOT borders
    d. ES 4: Regional Stability Returned
        i. Diplomatic monitoring agreements put in place (AU/UN mission)
        ii. North African countries accept/acknowledge shared security agreement
        iii. Economic infrastructure sharing agreement developed

12. Centers of Gravity
    a. Determine centers of gravity (Team)
        i. Critical Capability
        ii. Critical requirement
        iii. Critical vulnerability
    b. Enemy Strategic – RNA-Tuareg Tribal Identity/Ideology
        i. Common Narrative /History
            1. Organizational Scheme
                a. Internal contradiction of RNA narrative with known Tuareg Culture and sub-cultures (that is not our story)
            2. Collective Identity
                a. Contradictions of RNA Tuareg identity with sub-culture identity and narratives (You/they aren't even our people)
            3. Threat to identity
                a. Contradiction in RNA message of threat to identity with RNA actions (You made deals/diplomacy with the devil)
        ii. Ideological Dogma
            1. Islamic religious continuity
                a. Fundamental Islamism narrative
                b. Displaced Arab leadership elite from former Algeria
            2. Monopoly on Tuareg identity "authority"
                a. Competing Tuareg voices of "authority" from other factions
        iii. Use of all available means to advance cause
            1. International and Domestic Shock and revulsion for techniques employed in cause
        iv. Control of RNA Government
            1. Elected "mandate"
                a. International censure of RNA elections
            2. Set Policy and Strategy
            3. Exercise Regional D,I,E
            4. C2 of RNA Military Forces

        a. Centralized C2 structure that lacks agility and ability to exploit opportunity, heavily dependent on C2 communications

  v. <u>Control of RNA Economy</u>
    1. Control of hydrocarbon export
        a. Economic capability based on global oil markets
    2. Control of State Owned Enterprises (SOE)
        a. Culture of patronage/control which leads to mediocrity

c. Enemy Operational – RNA-East Corps/OSC (Operational Strategic Command)
  i. <u>Operational C2 (Command and Control)</u>
    1. Communications Infrastructure
        a. Reliance on unsecure communications platforms
    2. Situational Awareness
        a. Authoritarian/cultural aversion to "bad news"
    3. ISR
        a. Organic ISR capacity beyond HUMINT limited (reliance on int'l market)
  ii. <u>(Mech) Division (RNA) (Move/Maneuver)</u>
    1. Mech/Armor Brigades-Ground Maneuver
        a. Dated systems in most units (T62/72 + BTR60/BMP1)
        b. Force generation timeline produces little combined arms training at Brigade and above level for very long
        c. Morale of conscript army not exclusively Tuareg
        d. Low tactical initiative in junior leaders beyond scripted plans
    2. C2 Infrastructure
        a. Reliance on unsecure FM and Cellular telephones
    3. Situational Awareness/Understanding
        a. Authoritarian/cultural aversion to "bad news"
    4. Operational sustainment
        a. Long lines of communication with limited capacity to secure
        b. Finite fuel refinement and transportation capacity to support mechanized/armored forces
        c. Conscript army challenged to repair/maintain highly technical systems
  iii. <u>Integrated Air/Artillery (Op Fires)</u>
    1. 3.1 Communications Infrastructure
        a. 3.1.1 Reliance on unsecure FM and Cellular telephones
    2. 3.2 Airfields/Support Infrastructure

        a. 3.2.1 Few all-weather air strips to support A/C sorties into Tunisia at high sortie rates
    3. 3.3 Sustainment/Maintenance of systems
        a. 3.3.1 Reliance on foreign technicians/repair parts for systems
        b. 3.3.2 Dated systems in RNA Air Force (3-4 generation)

iv. <u>Integrated Air Defense (Protection)</u>
    1. 4.1 Communications infrastructure
        a. 4.1.1 Reliance on commercial internet/largely unsecure
    2. 4.2 IFF mechanisms
        a. 4.2.1 IFF reliability is low in dynamic environment
    3. 4.3 Sustainment/Maintenance of systems
        a. 4.3.1 Reliance on foreign technicians/repair parts for systems
    4. 4.4 High Altitude Interdiction capability
        a. 4.4.1 No high altitude protection systems

v. <u>Coordinated action with TLF and Criminal Enterprises</u>
    1. 5.1 RNA SOF forces/enduring presence
        a. 5.1.1 Foreign "invader" in southern Tunisia
        b. 5.1.2 Heavy handed techniques of coercion that lead to backlash
    2. 5.2 Financial support
        a. 5.2.1 Criminal enterprises "highest bidder" approach
    3. 5.3 Communications Infrastructure
        a. 5.3.1 Reliance on Commercial Internet/Cellular telephone systems and AM

d. Friendly Strategic – Will of the Tunisian Populace
    i. Critical Capabilities
        1. (1) USEASTCOM
        2. (2) Department of State
        3. (3) African Organizations
        4. (4) United Nations
    ii. Critical Requirements
        1. Information Operations
        2. Military Information Support Ops
        3. Military Exchange
        4. Embassies/Country Teams
        5. USAID
        6. State Support
        7. Refugee Support
        8. Internally Displaced Civilians
        9. Office for Coordination of Human Affairs
    iii. Critical Vulnerabilities

1. Aggressive US Action
2. Inconsistent IO Messages
3. Corrupt Governments
4. Funding
5. Tribal Violence
6. Coordination with NGOs
7. Interagency Support

e. Friendly Operational – Defense of Tunis
   i. Critical Capabilities
      1. (1) Tunisian Defense Forces
      2. (2) U.S. Air Power
      3. (3) Freedom of Maneuver
      4. (4) Sustainment
   ii. Critical Requirements
      1. Maintaining Key Terrain
      2. Communications Infrastructure
      3. Over-flight Rights
      4. ISR
      5. Host Nation Basing (Air & Port)
      6. Security of SPODs & APODs
   iii. Critical Vulnerabilities
      1. Aggressive US Action
      2. Corruption
      3. Phasing of US Force Flow
      4. Outdated/Inferior Equipment
      5. Strategic partnership support
      6. Insurgent activity

13. Decisive Points
   a. Logical Decisive Points
      i. Disrupt/Delay RNA Vanguard
      ii. Host Nation Basing (Air & Port)
      iii. Secure SPODs/APODs
      iv. Establish C2/interoperability w/ Tunisia
      v. Phasing of US Force Flow
      vi. Control/Influence of deliberation in UN
      vii. Control/Influence of criminal organization in Tunisia
      viii. Control of tribal leadership in S Tunisia
      ix. Air Superiority over Tunisia is achieved.
      x. Control of sea lanes entering strategic ports
      xi. Operational Logistics postured to support invasion
      xii. Humanitarian Crisis in contested areas averted
      xiii. Civil Authority in occupied areas established
   b. Terrain Oriented Decisive Points
      i. Maintain key defensive positions of narrow corridors in northern and central parts of the country to delay and disrupt the enemy

ii. Attrite RNA forces vicinity PL Platinum to Gold
iii. Secure and defend Djefna and Bedjez al Bab
iv. Secure and Defend Ben Arous
v. Secure and Defend oil fields

14. Determine CCIR's
   a. PIR's
      i. What are the conditions under which RNA will cease aggression (i.e. UN, AU actions)?
      ii. Where are the RNA logistical support bases?
      iii. What is the disposition and status of RNA strategic forces and East Corps?
      iv. What is the disposition, status and intent of the Tuareg Liberation Front (TLF)?
      v. What is the disposition and status of IDPs in Tunisia?
      vi. What is the status of critical infrastructure in Tunisia and RNA to include the security, medical, and transportation nodes?
      vii. How will Western Sahara support RNA to counter combined operations?
      viii. How will Tuareg tribe support RNA to counter combined operations?
      ix. Will RNA attack/invade Tunisia in the next 6 days? (Assum)
      x. RNA SOF infiltration into Tunisia?
      xi. RNA SOF coordination w/ TLF and criminal elements?
      xii. Is the RNA government making attempts to influence US strategic partners and African organizations (particularly attempts to deny US basing and overflight)? (Assum)
      xiii. Has the RNA activated military reserve forces / reserve recall or initiated an extension of conscription service?
      xiv. Has the RNA deployed forces forward (naval forces east of Algiers, air and ground forces to positions closer to border)?
      xv. Has the RNA positioned SCUD units in attack positions?
      xvi. Are criminal elements in Tunisia limiting port access?
      xvii. Have attacks occurred on symbols of US authority (i.e. US Embassy in Tunisia)?
      xviii. How are regional governments attempting to influence US Strategic Partners?
      xix. Are any other external aggressors threatening the sovereignty of Tunisia?
      xx. Has RNA acquired a WMD capability?
   b. FFIRs
      i. Support or changes in support from a key strategic partner that will impact the combined operation (notably basing and overflight support) (Assum)
      ii. Occurrences of fratricide to combined forces

iii. Incidents that will likely be reported by the media and negatively impacts the components ability to conduct their mission
iv. What is the status of LOCs and PODs (particularly strategic lift, air and sea ports of debarkation degradation that negatively impacts US force flow into theater)?
v. Class III (fuel) or V (ammo) shortage which impacts components ability to conduct their mission
vi. Degradation to component capabilities that impacts their ability to conduct their mission
vii. Has a humanitarian situation occurred/changed that impacts components ability to conduct their mission?
viii. Changes in civilian population support to combined operations that will impact the components ability to conduct their mission
ix. What is the number and status of US citizens in Tunisia?
x. Is Tunisia planning to or has it shut-off pipeline flow from RNA?
xi. What are the types/status/capes of NGOs/IGOs in Tunisia?
xii. Is Tunisia able to execute its defensive plan against an RNA invasion? (Assum)
xiii. Is Tunisia able to secure and maintain control of LOCs and PODs? (Assum)
xiv. What are the transportation/logistical support resources available in Tunisia?
xv. Has the RNA-Tunisian joint oil venture been terminated?

15. Conduct initial force structure analysis.
    a. Structure/Capabilities Needed
        i. Phase I – Deter
            1. Airbase Defense/Sustainment
            2. Air Superiority
            3. Interdiction/Bombers
            4. MEU
            5. CSG
        ii. Phase II – Seize the Initiative
            1. SEAD/Ground Attack
            2. 2-3 IBCT/SBCT's
            3. Logistics
            4. Navy Strike
        iii. Phase III – Dominate
            1. Ground Attack
            2. 1-2 HBCT
            3. MEU
            4. Police/Security Forces
        iv. Phase IV - Stability
            1. Police/Security Forces
            2. Civil Affairs
            3. Engineers

16. Conduct initial risk assessment.

Risk Identification	Mission Risk Assessment			Mitigation	Residual Risk
	Probability*	Severity**	Risk***		
Expansion of regional violence/ instability	O	CR	H/M	Diplomatic, IO TCP	M
Loss of popular support to GOT/ Coup	O	CR	H	Diplomatic, Regional/IGO, Mil to Mil Engagement	H/M
Loss of Access	L	M	M/L	Diplomatic , ISR/ Analytic Focus	L
Loss of Critical Infrastructure	L	M	M/L	ISR responsive targeting, IO, Air Supremacy	L
Tunisian military capitulates	S	M	M/L	IO, Mil to Mil engagement	L
Humanitarian Crisis	L	M	M/L	Force Flow, IO	M/L

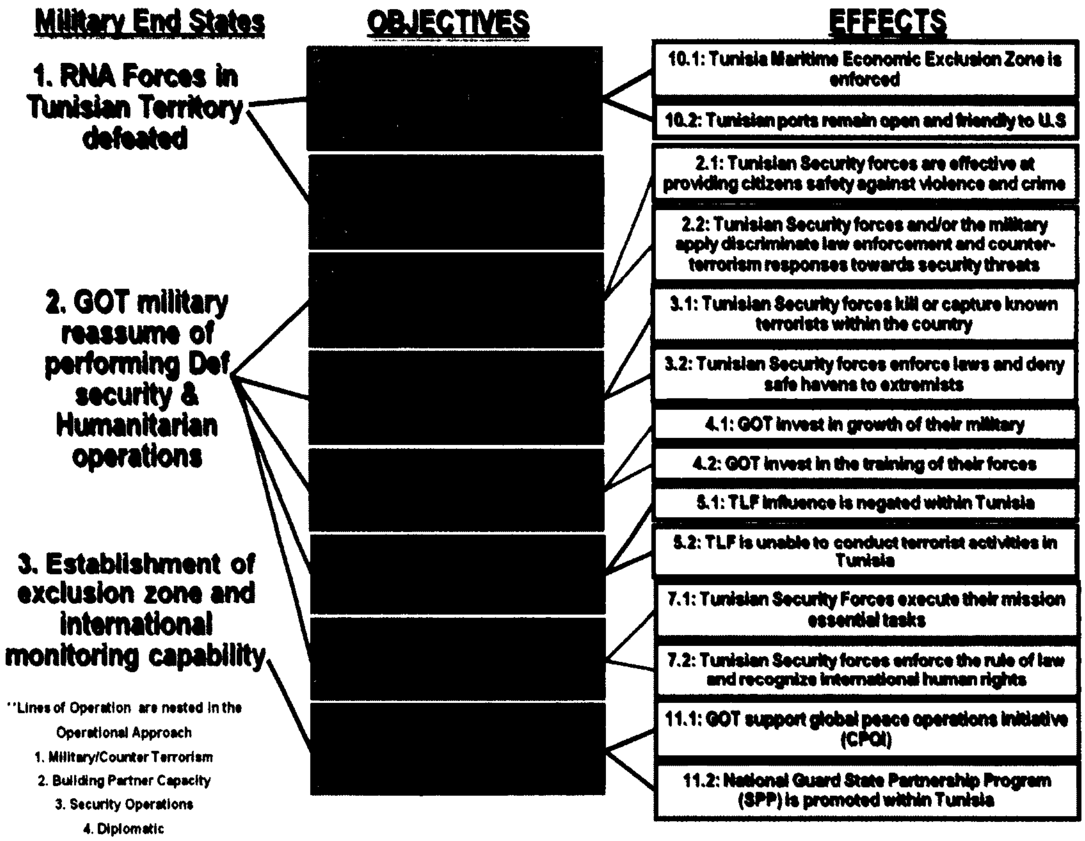
* Frequent, likely, occasional, seldom, unlikely
** Catastrophic, critical, marginal, negligible
*** Extremely high, high, moderate, low

17. Draft mission statement - When directed, USEASTCOM Forces will engage hostile forces threatening Tunisia and its territorial waters in order to maintain Tunisian territorial integrity by supporting Inter-agency efforts, protecting designated Tunisian critical infrastructure/assets, and enabling Strategic Communications in order to isolate and deter aggression.

18. Consolidate mission statements - USEASTCOM, in conjunction with other government agencies, will deter RNA aggression with Tunisia. USEASTCOM Forces will engage hostile forces threatening Tunisia and its territorial waters in order to maintain Tunisian territorial integrity by supporting Inter-agency efforts, protecting designated Tunisian critical infrastructure/assets, and enabling Strategic Communications in order to isolate and deter aggression.

19. Determine strategic themes necessary to accomplish mission.
    a. Promoting regional stability
    b. Supporting Democratic efforts / Rule of Law
    c. Reinforcing the key partnerships within the region
    d. Promoting economic growth

20. Initial Operational Approach



The material above is based on a fictional scenario with personalities, intentions, and visions are constructed to present certain planning challenges and do not reflect real people or national aims and as such it is to be used only educational purposes.

Appendix E: Mission Analysis Activities

Below are the steps of the mission analysis activities (Joint Staff, 2011) that students conducted during the Mission Analysis step of the Joint Operational Planning Process (JOPP). It includes the development of the initial operational approach as part of the process in preparation for later development of a refined operational approach. The briefing template provided to students was simply a title slide with a number of blank slides following that are labeled with headings on them to guide students through the mission analysis step. The slide also includes a doctrinal definition for each heading area. An example is shown below (Figure 1).

1. Analyze higher headquarters planning activities and strategic guidance
2. Review commander's initial planning guidance, including his initial understanding of the operational environment, of the problem, and description of the operational approach
3. Determine known facts and develop planning assumptions
4. Determine and analyze operational limitations
5. Determine specified, implied, and essential tasks
6. Develop mission statement
7. Conduct initial force allocation review
8. Develop risk assessment
9. Develop mission success criteria
10. Develop commander's critical information requirements
11. Prepare staff estimates
12. Develop initial operational approach
13. Prepare and deliver mission analysis brief

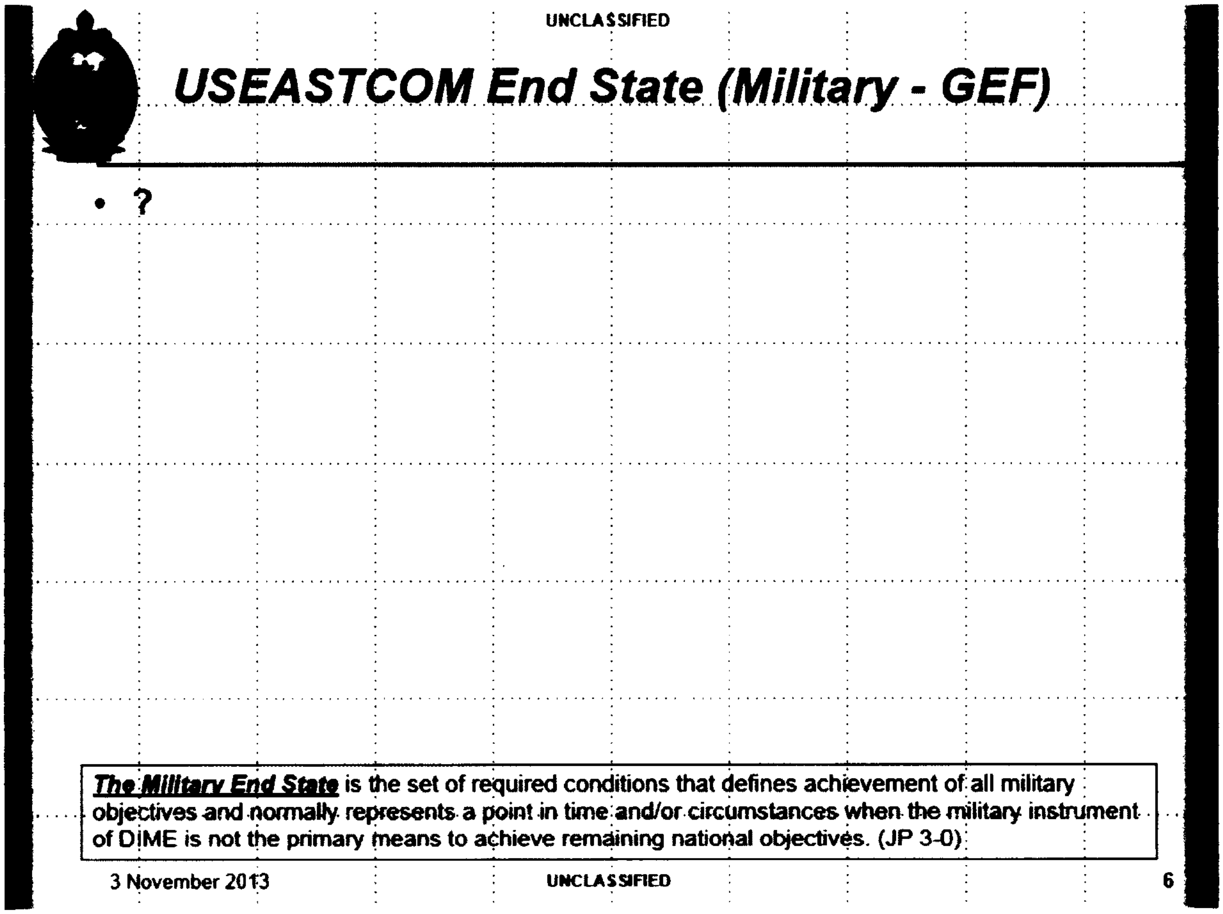


Figure 1 Briefing Template Slide. This slide shows a sample slide from the briefing template provided of student use. It is an example of the level of detail provided for students to complete as part of their analysis efforts.

Appendix F: Group Product Rubric

The lesson used for this research is a collaborative session that requires students to conduct mission analysis for the fictional United States Eastern Command. This is part of the process they are using to lay the foundation leading to the development of a contingency plan. Each group is expected to develop and deliver a mission analysis brief to the other groups as the product of their group efforts. The rubric below (Table E1) can be used to evaluate the results of the group effort.

Table F1

Mission Analysis Group Product Rubric

Item	Fails to Meet Standards	Meets Standards	Exceeds Standards
Known Facts/Conditions (4)	Listed facts are not facts (0)		Listed facts are facts (1)
	The list of facts is greatly incomplete. (0)	The list of facts is mostly complete (1)	The list of facts is complete. (2)
	No evidence that the assumptions from higher authority were analyzed and are simply listed as fact (0)		Assumptions from higher authority analyzed and placed in appropriate category. (1)
	0	0	0
Assumptions (10)	Assumptions are not logical (0)	Some of the listed assumptions are not logical (1)	Assumptions are logical (2)
	Assumptions are not realistic (0)	Some of the listed assumptions are not realistic (1)	Assumptions are realistic (2)
	Assumptions are not essential to the planning effort (0)	Some of the assumptions are not essential to the planning effort (1)	Assumptions are essential to the planning effort (3)
	Assumptions do not address gaps in	Assumptions only partially address gaps	Assumptions address gaps in knowledge

	knowledge (0)	in knowledge (1)	(3)
	0	0	0
Military Objectives (6)	Each objectives is not clearly defined (0) Each objectives is not decisive (0) Single objectives addresses multiple goals (0) Objectives are not specific and attainable (0) Objectives do not link to higher-level objectives (0) Objectives infer ways or means (0)		Objectives are clearly defined (1) Objectives are decisive (1) Objective address single goal (1) Objectives are specific and attainable (1) Objectives are linked to higher-level objectives (1) Objectives do not infer ways or means (1)
	0	0	0
Military End State (5)	End state does not link to the strategic end state (0) End state does not describe military conditions necessary to achieve strategic objectives (0) End State does not address problem (0)	Many of the conditions listed in the end state are not military in nature (1) End state partially addresses problem (1)	End state links to the strategic end state (1) End state describes the military conditions necessary to achieve strategic objectives (2) End state addresses problem (2)
	0	0	0
Limitations (5)	Constraints are not identified Restraints are not identified Fail to identify all guidance and environmental variables that limit freedom of action.	The list of constraints is incomplete. (1) The list of restraints is incomplete (1)	All constraints are identified (2) All restraints are identified (2) Have identified all guidance and environmental variables that limit freedom of action. (1)
	0	0	0
Theater Tasks (4)	Specified task are not clearly identified Implied tasks are not		The specified task list is complete (1) The list of Implied

	identified		tasks is complete (1)
	Essential tasks that are absolutely necessary, indispensable, or critical to accomplish the mission are not identified	Some of the identified essential tasks are not essential to accomplish the mission (1)	Essential tasks that are absolutely necessary, indispensable, or critical to accomplish the mission are identified (2)
	0	0	0
Mission Statement (10)	Mission statement is missing more than one element of the basic who, where, and when (0)	Mission statement is missing one of the key elements of who, where, or when (1)	Mission statement answers the basic who, where, and when (3)
	Mission statement does not include mission essential tasks (0)	Mission statement is missing most of the essential tasks (1)	Mission statement includes mission essential tasks (3)
	Mission statement does not describe the reason for conducting the tasks (0)	The purpose of the mission does not support attainment of the strategic end state (1)	The purpose of the mission supports attainment of the strategic end state (3)
	Mission statement explains how to accomplish the tasks (0)		Mission statement does not capture how to accomplish the tasks (1)
	0	0	0
CCIR's (5)	CCIR's do not support resolution of assumptions (0)	CCIR's partially support resolution of assumptions (1)	CCIR's support resolution of assumptions (2)
	CCIR's do not reflect important information needed for CDR's decision-making (0)	CCIR's partially support important information requirements needed for the CDR's decision-making (1)	CCIR's reflect important information needed for CDR's decision-making (2)
	0	0	0
Force Structure Analysis (7)	Force structure analysis is not based on assigned forces (0)		Force structure analysis is based on assigned forces (1)

	Force structure analysis does not identify shortfalls (requirements for future IPL/ Force Posture Plan or other long range strategic resourcing venues) (0)	Force structure analysis partially identifies shortfalls (requirements for future IPL/ Force Posture Plan or other long range strategic resourcing venues) (1/2)	Force structure analysis identifies shortfalls (requirements for future IPL/ Force Posture Plan or other long range strategic resourcing venues) (3)
	Force structure analysis does not provide recommendations and justification	Force structure analysis provides recommendations but no justification (1/2)	Force structure analysis recommendations address identified shortfalls and justification (3)
	0	0	0
Strategic Risk Assessment (10)	Risks (threats) are not clearly identified (0)	Risk list is missing some major threats (1)	Risks (threats) are clearly identified (2)
	Risk levels are not appropriate	Some risk levels are inappropriate (1)	Risk levels are appropriate (2)
	Threat severity is not identified	Threat severity, while indicated is incorrect (1)	Threat severity is identified and correctly assessed (2)
	Risk mitigation actions are not identified	Risk mitigation are inadequate to address the risk (1)	Risk mitigation actions are identified and appropriate to address the risk (2)
	Opportunities within the theater are not identified (0)	Some opportunities available in the theater of operations are not identified (1)	Opportunities within the theater are identified (2)
	0	0	0
Centers of Gravity (10)	Operational Centers of Gravity not identified (0)	Friendly and Enemy Operational Centers of Gravity identified (1/1)	Critical Factor Analysis conducted on the Operational Center of Gravity, identifying Critical Capabilities, Requirements, and Vulnerabilities (2)
	Strategic Centers of Gravity not identified (0)	Friendly and Enemy Strategic Centers of Gravity Identified (1/1)	Critical Factor Analysis conducted on the Strategic Center of Gravity,

			identifying Critical Capabilities, Requirements, and Vulnerabilities (2))
	Critical factor Analysis not conducted on Centers of Gravity (0)	Partial or faulty Critical factor Analysis conducted (2)	
	0	0	0
Description Initial Operational Approach(10)	Current conditions are not identified (0)	Current conditions are identified (1)	Desired conditions address the current conditions (2)
	Lines of Operation (LOO)/Effort (LOE) are poorly developed or absent (0)	Lines of Operation(LOO)/Effort (LOE) are clearly identified (1)	
	Defeat mechanisms are not identified or not listed for every (LOO/LOE) (0)	Defeat Mechanisms are identified (1)	Defeat mechanisms are reflected in the supported objectives (2)
	Supported objectives are not arrayed along LOO/LOE's (0)	Supported objectives are identified and arrayed along the LOO/LOE's (1)	Supported objectives will result in attainment of the desired outcome (2)
	LOO/LOE's do not indicate the desired conditions as outcomes (0)	Desired conditions for each LOO/LOE are identified (1)	Desired conditions support the end state (2)
	End State is not identified (0)	End state is identified (1)	
	0	0	0

Appendix G: Student Contribution Survey Instrument

Contingency Planning Mission Analysis

Lesson: CNP 3

Name: ______________ Seminar: _____ Group: ______ Student ID: ____

Assessment of others is an important skill. Please take time to think about the questions, and answer as honestly and frankly as you can. The individual responses will be kept confidential.

Part 1

In the column headings in the table below, record your ID number and those of each member of your group. For each member of the group, award a mark of 1 to 5 for that group-mate's level of participation in each of the tasks listed in the following table:

Use the following grading scale and enter your responses in the following table:

1. didn't contribute in this way
2. willing but not very successful
3. average
4. above average
5. outstanding

	Yourself Student ID	Name #1 #1 ID	Name #2 #2 ID	Name #3 #3 ID	Name #4 #4 ID	Name #5 #5 ID
Identification of facts and development of assumptions						
Defining the tasks and mission						
Defining the problem (End states/objectives/termination criteria)						
Identification and critical factor analysis of the Center of Gravity						

Force structure analysis	
Identification and analysis of risk	
Delivering the group presentation	
Overall	

From (Goldfinch & Raeside, 1990) as modified by Li (2001)

Part 2

1. How well did the group function and what were the challenges (if any) encountered by the group?

2. Describe how well the workload distributed among the group members? How effective was the distribution?

3. Describe the level of conflict/discourse within the group. How did it contribute/detract from the overall group success.

4. Do you have any other comments to provide concerning the group effort?

Appendix H: Cognitive Load

You will be asked to rate the mental effort you expended during multiple parts of the lesson. Please rate your own mental effort for a specific section. Upon completion of the lesson, you will also be asked to provide an overall mental effort for the mission analysis and operational design effort as well as for the exam itself. For each section, using the scale below to rate each section, place an "X" in the box that best represents the mental effort you expended.

	Very, Very little mental effort	Very low mental effort	Low mental effort	Rather low mental effort	Neither low or high mental effort	Rather high mental effort	High mental effort	Very high mental effort	Very, Very high mental effort
Development of facts & assumptions									
Specified/Implied/Essential Tasks									
Termination Criteria Development									
Center of Gravity Analysis									
Objective Development									
Force Structure Analysis									
Risk Matrix									
Overall									
Exam									

Appendix I: Calculation of the Score for an Individual from the Group Score

The example is for a group of three students; Alice, Bill and Christine. Each student is rated for effort by the other two members in the group Table H1).

Table I1

Sample Group Score Data

Group Mark	75.5					
	Alice		Bill		Christine	
	(B)	(C)	(A)	(C)	(A)	(B)
(A) Literature search	3	2	3	2	4	5
(B) Analyzing the literature	3	2	4	3	4	4
(C) Writing the report	4	3	4	3	3	4
(D) The group presentation	4	3	3	3	4	4
	14	10	14	11	15	17
Individual Effort Rating		24		25		32

Using the data in the table above, the below items will step the researcher through the process to use peer data to calculate a normalized group score for each individual based on the participation of that group member as perceived by themselves and their peers.

1. Calculate the Individual Effort Rating (IER) by adding scores for each person.

    a. Alice – 24

    b. Bill – 25

    c. Christine - 32

2. Calculate the Average Effort Rating (AER) by adding the IER's for all group members and dividing by the number of group members.

(24 + 25 + 32)/3 = 27

3. Calculate the Total Ratings (TR) provided by each group member on their peers by adding the scores that each rater provided.

a. Alice – 29

b. Bill – 31

c. Christine - 21

4. Calculate each group members Bias Factor by dividing the Total Rating provided by each group member and dividing by the Average Effort Rating.

a. Alice – 29/27 = 1.07

b. Bill – 31/27 = 1.15

c. Christine – 21/27 = 0.78

5. Calculate the Normalization Factor for each group member by dividing 1 by the group members' Bias factor.

a. Alice – 1/1.074074 = .93

b. Bill – 1/1.14815 = .87

c. Christine – 1/0.78 = 1.29

6. Calculate normalized data by multiplying each rating by the normalization factor of each rater.

7. To calculate the normalized Individual Effort Rating, add the Normalized Ratings for each group member.

a. Alice – 25.05

b. Bill – 27.18

c. Christine – 28.78

8. Calculate the Individual Weighting Factor by dividing the Normalized Rating for each group member by the Average Effort Rating (AER).

a. Alice – 25.05069/27 = .93

b. Bill – 27.1773/27 = 1.01

c. Christine – 28.772/27 = 1.07

9. To calculate the normalized grade for each group member, multiply the group mark by the Individual Weighting Factor of each group member.

a. Alice – 75.5 X .93 =70.05

b. Bill – 75.5 X 1.01 = 76.00

c. Christine – 75.5 X 1.07 = 80.45

Table H2 shows completed calculations for normalization of group scores based on student-rated participation.

I2

Sample Normalization

Group Mark	75.5					
	Alice		Bill		Christine	
	(B)	(C)	(A)	(C)	(A)	(B)
(A) Literature search	3	2	3	2	4	5
(B) Analyzing the literature	3	2	4	3	4	4
(C) Writing the report	4	3	4	3	3	4
(D) The group presentation	4	3	3	3	4	4
	14	10	14	11	15	17
Individual Effort Rating		24		25		32
AER	27					
IWF		0.89		0.93		1.19
Total Ratings		29		31		21
Bias factor		1.07		1.15		0.78
Normalization factor		0.93		0.88		1.29
Weighted grade		70.29		65.76		97.07

Normalized data			
	Alice	Bill	Christine

	(B)	(C)	(A)	(C)	(A)	(B)
(A) Literature search	2.61	2.57	2.79	2.57	3.72	4.35
(B) Analyzing the literature	2.61	2.57	3.72	3.86	3.72	3.48
(C) Writing the report	3.48	3.86	3.72	3.86	2.79	3.48
(D) The group presentation	3.48	3.86	2.79	3.86	3.72	3.48
	12.19	12.86	13.03	14.14	13.97	14.81
Individual Effort Rating		25.05		27.18		28.77
AER	27					
IWF		0.93		1.01		1.07
Weighted and Normalized grade		70.05		76.00		80.45

CURRICULUM VITA

NAME: Gary L. Roemmich
ADDRESS: 7800 Hampton Blvd
Joint Forces Staff College
Norfolk, VA 23511

EDUCATION:

M.M.A. in Marine Affairs, University of Rhode Island, 1984)
B.S. IN Fisheries Biology, University of Washington, 1976

EXPERIENCE:

6/2010 – Present	Associate Professor, Director of Electives, Joint and Combined Warfighting School, Joint Forces Staff College, National Defense University
6/2006 – 6/2010	Assistant Professor, Joint and Combined Warfighting School, Joint Forces Staff College, National Defense University
7/2004 – 6/2006	Instructor, Director of Curriculum Development, Joint and Combined Warfighting School, Joint Forces Staff College, National Defense University
6/1990 – 3/1994	Instructor, Course Developer, Joint and Combined Warfighting School, Joint Forces Staff College, National Defense University
10/1981 – 6/1983	Instructor, Navy Officer Candidate School

TEACHING:

Course Number	Title	Responsibility
JCWS 1S40 now JCWS 6450	Deliberate Planning	Design, Develop, Deliver
JCWS 1S50 now JCWS 6460	Crisis Action Planning	Design, Develop, Deliver
JCWS 1S20 now JCWS 6410	Campaigning	Develop, Deliver
1S70	PURPLE SUNSET Exercise	Design, Develop, Deliver
JFSC 6016	US Transportation Command	Design, Develop, Deliver
JFSC 6066	Operationalizing Knowledge Management	Design, Develop, and Deliver

HONORS, AWARDS AND PRIZES:

2012 ODU Mandell Prize in Instructional Design and Technology

MEMBERSHIP IN PROFESSIONAL SOCIETIES:

American Educational Research Association, 2009 – Present
Association for Educational Communication and Technology, 2009 - Present

UNIVERSITY SERVICE:

Treasurer, Instructional Design and Technology Graduate Student Organization, 2010-2012

COMMUNITY SERVICE:

Chairman of the Board, Neighborhood Youth Soccer, Virginia Beach, 1992-1993.