Maximizing the Benefits of Collaborative Learning

In the College Classroom

by

Rachel J. Lam

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2013 by the
Graduate Supervisory Committee:

Kathryn Nakagawa, Chair
Samuel Green
Jill Stamm

ARIZONA STATE UNIVERSITY

May 2013

ABSTRACT

This study tested the effects of two kinds of cognitive, domain-based preparation tasks on learning outcomes after engaging in a collaborative activity with a partner. The collaborative learning method of interest was termed “preparing-to-interact,” and is supported in theory by the Preparation for Future Learning (PFL) paradigm and the Interactive-Constructive-Active-Passive (ICAP) framework. The current work combined these two cognitive-based approaches to design collaborative learning activities that can serve as alternatives to existing methods, which carry limitations and challenges. The “preparing-to-interact” method avoids the need for training students in specific collaboration skills or guiding/scripting their dialogic behaviors, while providing the opportunity for students to acquire the necessary prior knowledge for maximizing their discussions towards learning.

The study used a 2x2 experimental design, investigating the factors of Preparation (No Prep and Prep) and Type of Activity (Active and Constructive) on deep and shallow learning. The sample was community college students in introductory psychology classes; the domain tested was “memory,” in particular, concepts related to the process of remembering/forgetting information. Results showed that Preparation was a significant factor affecting deep learning, while shallow learning was not affected differently by the interventions. Essentially, equalizing time-on-task and content across all conditions, time spent individually preparing by working on the task alone and then discussing the content with a partner produced deeper learning than engaging in the task jointly for the duration of the learning period. Type of Task was not a significant factor in learning outcomes, however, exploratory analyses showed evidence of Constructive-type behaviors leading

to deeper learning of the content. Additionally, a novel method of multilevel analysis (MLA) was used to examine the data to account for the dependency between partners within dyads.

This work showed that "preparing-to-interact" is a way to maximize the benefits of collaborative learning. When students are first cognitively prepared, they seem to make the most efficient use of discussion towards learning, engage more deeply in the content during learning, leading to deeper knowledge of the content. Additionally, in using MLA to account for subject nonindependency, this work introduces new questions about the validity of statistical analyses for dyadic data.

This dissertation is dedicated to

Martine

for his continual love and support,

and Adelina

for graciously letting mommy

work "at the coffee shop again"

and again.

ACKNOWLEDGMENTS

I would like to thank my advisor, Kathy Nakagawa, for her continual support, encouragement, and mentorship over the last four years. The guidance that she provided has been invaluable and my accomplishments would not have been possible without her. I also thank Jill Stamm. Her belief in my potential that began nearly 10 years ago has encouraged me during the most challenging times. I am beyond indebted to her for her love and support. The ways that both of these women looked out for my best interests and supported my decisions will not be forgotten. I thank my third committee member, Sam Green, for sharing his expertise in statistical analysis and pushing me to "do things right." I am grateful for his time, patience, and kindness, from the development of my study through analysis and interpretations.

I would also like to thank a cadre of colleagues that I now consider friends. Ruth Wyle and Kasia Muldner have carried me through difficult periods of my graduate career. I am grateful to Ruth for her excellent guidance and mentorship, as well as her encouragement and positive outlook. I am grateful to Kasia for pushing me to stay one step ahead, being a great sounding board, and for her friendship over the last two years. I thank Michelle Jordan for her support, encouragement, and many conversations about research and the academic career path. Finally, thanks to Muhsin Menekse for his support and feedback.

There are several people I would like to thank that made my study possible. I thank Julie Morrison for sharing her students with me, providing feedback on my study materials, helping me to brainstorm at various stages of my process, and becoming a good friend. I also thank Diane McMillan and Katy Hisrich for allowing me to work with their students, and the administrators at Glendale Community College for believing in the value of my work. Thanks to Masumi Iida for her assistance in developing new analytic techniques to run my data. I thank Michelene Chi for her early influence on my work and for the rigorous training that I received at the Chi Learning and Cognition Lab. I also thank Ann Ewbank for her support during a pivotal change in my graduate career that influenced the direction of my work.

Lastly and most importantly, I thank my family and friends for their love and support. I am grateful to Martine for supporting my hopes and dreams and for the sacrifices that he made to make this all possible. Thanks to my daughter, Adelina, for putting up with a sometimes too-spread-thin mom and loving me anyway. Thanks to my parents, Diane, Mark, and Geanette, for always encouraging my educational and career pursuits, and my in-laws, Nashie and Salem, for their love and support. Thanks to Alexus, Dianne, Gloria, Allison, and Heather for being great friends throughout this long process.

This work has been partly funded by a GRSP grant from the ASU Graduate and Professional Student Association and a Spring 2013 Dissertation Completion Fellowship. I am grateful for the opportunities that both of these resources provided.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Chapter 1

INTRODUCTION

Collaborative learning activities have become common instructional strategies. Teachers in many educational settings use them, from the K-12 to graduate level. For instance, Jigsaw has become a well-known method of collaborative learning, where students engage in multiple phases of studying and learning about a particular aspect of a larger concept or a complex problem, and then share information with each other to learn all aspects of the concept (Aronson, Stevens, Sikes, Blaney, & Snapp, 1978). Problem-based or project-based learning methods are also widely used and involve students engaging in a group or team project that encourages problem-solving, creative thinking, and application of knowledge in real-world contexts (Barron et al., 1998). In addition to these more developed collaborative learning methods, teachers may provide less structured collaborative opportunities by having students work on group projects, talk with a partner in class, or write a group paper. Collaborative learning activities are often founded on socio-cognitive perspectives, and provide opportunities for students to learn from each other (Vygotsky, 1978).

Collaborative learning has been extensively studied. A search of “collaborative learning” in the research databases PsycINFO and ERIC from the last two decades produces over 5000 entries and many of these studies support the use of collaboration to improve student outcomes. Much of this work has been tested in both the classroom and laboratory, and in face-to-face as well as computer-mediated settings, and has found improvements in measures of student achievement, productivity, critical-thinking skills, motivation, and self-esteem (Johnson & Johnson, 1992, 2009; Johnson, Johnson, &

Smith, 1991). Although the work on collaborative learning can be situated in social learning theories, sociocultural views, or socio-cognitive perspectives, my work approaches the effects of collaborative learning from a primarily cognitive perspective. For instance, is there evidence that collaboration improves, corrects, or influences students' mental models? How are mental model improvements measured and associated with factors of collaboration? Learning at a mental model level may also be framed as acquiring conceptual understanding. Existing work provides some evidence that collaboration can improve conceptual understanding of certain topics, however, other work tells us that collaboration can fail to produce the positive learning outcomes expected.

Collaboration as Discussion

First and foremost, I refer to collaborating as engaging in verbal discussion with a partner or small group for the purpose of learning. This verbal discussion can take place in person or over computer-supported tools. Although many collaborative learning activities involve physical activity (such as building a model rocket (see work by Petrosino and colleagues as cited in Barron et al., 1998)), I acknowledge Chi's assertion that "learning seems to occur in the verbal discussion rather than in the motoric interactions" (2009, p. 80). Chi (2009) further identifies dialoguing as a main focus for assessing interaction between students in learning situations. Therefore, from this point forward, I may use the terms discussing, dialoguing, interacting and collaborating interchangeably.

There are a number of reasons why peer-to-peer discussion during learning activities should improve student domain-based understanding. It allows students to obtain immediate feedback from one another on which they can reflect upon, incorporate others' perspectives into their own thinking, and become better aware of their own (mis)understandings by being questioned and prompted to explain their ideas (Chi, 2000; Chi, de Leeuw, Chiu, & LaVancher, 1994; Coleman, 1998; King, 1994, 1999; Kneser & Ploetzner, 2001). Discussion with a peer provides opportunities for students to compare and contrast conflicting ideas, engage in debates and challenge each other, or catch each other's errors or mistakes (Andriessen, Baker, & Suthers, 2003; Asterhan & Schwarz, 2009, 2010; Engle & Conant, 2002; Hausmann, Nokes, VanLehn, & Van de Sande, 2009). It also allows students to co-construct ideas or solutions to problems, elaborate on each other's ideas in a way that they could not accomplish alone, and provides opportunities for students to create shared meaning and understanding (Hausmann, 2006; Hausmann & Chi, in preparation; Roschelle, 1992; Teasley & Roschelle, 1993). Ultimately, students may be more likely to experience cognitive conflict due to the presence of others' responses, contributions, rebuttals, elaborations, questions, etc., which may increase their attempts to resolve inconsistencies in their own thinking.[1] In other words, discussion during collaborative activities can increase the chances that students will cognitively engage with the domain content in a deep way, leading to improvements in learning. But, we know that collaborative activities in educational settings do not always result in such improved outcomes.

[1] I refer to cognitive conflict in the Piagetian sense, as an instance where a learner recognizes information that conflicts with his/her existing prior knowledge and becomes motivated to resolve the conflict (Piaget, 1977, 1985).

Cognitive Perspectives

Some work that has assessed students' domain-based conceptual understanding, and studies that have measured students' externalizations of their internal knowledge structures have provided evidence that collaboration improves learning by altering and/or creating new understanding (Coleman, 1998; Janssen, Erkens, Kirschner, & Kanselaar, 2010; Van Boxtel, Van der Linden, & Kanselaar, 2000a). In addition, other work attests that under certain conditions, collaboration seems to enhance student learning outcomes beyond what students are capable of by working alone (Chi, Roy, & Hausmann, 2008; Hausmann, Van de Sande, & VanLehn, 2008; Shirouzu, Miyake, & Masukawa, 2002). However, work that has focused more specifically on communication and dialogue factors of collaboration has found mixed results towards how collaboration improves learning. Factors such as how balanced students' contributions in a dialogue are (Volet, Summers, & Thurman, 2009), or how well collaborators can attend and relate simultaneously to one another's ideas and the task-at-hand (Barron, 2003), can influence students' conceptual understanding of a domain topic. Thus, researchers have become interested in determining the conditions under which collaboration is most fruitful for learning. This includes investigating how a variety of factors within collaborative learning settings are associated with positive outcomes, and what can be done to facilitate the kinds of interactions between students that are more likely to lead to deeper understanding of a concept. To quote Dillenbourg and Hong, "Collaborative learning is not always effective; its effects depend on the richness and intensity of interactions engaged in by group members during collaboration" (2008, p. 6).

In adopting a cognitive perspective towards examining collaborative learning, students' understanding of typically difficult concepts is of particular interest. I use the term "difficult concepts" to refer to topics that require deeper understanding to achieve learning success. (I acknowledge that there are topics that are difficult that do not require deep understanding, however, I do not address those here.) Using deep understanding as a measure of learning makes sense for concepts that involve complex processes that require learners to change or elaborate the structures of their prior knowledge. These kinds of difficult or complex concepts often give rise to misconceptions and errors in thinking, which can be detected through certain kinds of assessments, including those that require students to: (a) externalize their mental models in some form (freely writing, creating concept maps, drawing graphs/figures, etc.) (Haugwitz, Nesbit, & Sandmann, 2010; Schwartz, 1995; Schwartz, Sears, & Chang, 2007; Van Amelsvoort, Andriessen, & Kanselaar, 2007) or (b) answer questions and solve problems that cannot be accomplished without generating new inferences beyond the learning materials (Chi, 2000; Chi, Roscoe, Slotta, Roy, & Chase, 2012). In support of my current work, the literature reviewed here includes learning studies in the hard sciences, mathematics, and social sciences that has focused on the effects of collaboration on student learning and understanding of difficult material.

Goals of the Current Research

Considering the conditions under which collaboration has been found to improve student learning, conditions where collaboration seems to provide no added benefit, and some interventions that have been found to increase the chances that students will take advantage of the benefits that collaboration offers towards learning, my work investigates an alternative method for structuring collaborative activities in the classroom that has been less studied. This method involves intentionally activating students' existing knowledge structures (within a specific domain topic) by engaging them in particular individual learning activities that induce a state of cognitive conflict and awareness. This activation of prior knowledge may serve as a form of cognitive preparation prior to participating in a collaborative activity. Examination of this "preparing-to-interact" phenomenon sheds light on how to maximize the benefits of collaborative learning.

Firstly, to illustrate when collaboration positively affects student understanding of difficult material, a few key studies that have measured student outcomes resulting from individually engaging in a learning task compared to engaging in the task through discussion with a peer are reviewed. These studies measured improvements in student learning of difficult concepts or students' ability to create abstract principles from working with concrete problems or tasks. In both types of cases, mental model change is apparent because accomplishing the tasks requires correct structural (deep, principle-level) understanding of the material. Then, the work that has provided evidence for when collaboration fails to produce the positive outcomes expected is summarized. Collaboration provides a natural setting for students to ask each other questions, explain their ideas to one another, and to elaborate on the contributions made during discussion, which are all behaviors that have been shown to improve learning. However, students do not always take advantage of these opportunities.

To increase the likelihood that students do take advantage of the benefits that collaboration offers, researchers have investigated and recommended a number of practices and strategies for encouraging fruitful discussion during collaborative activity.

In review of this work, these practices can be categorized into three main types of interventions: (a) teaching students collaboration skills, (b) externally guiding student interactions, and (c) providing opportunities for students to work through ambiguous or open-ended tasks. These interventions are described more thoroughly, and the evidence for how each influences collaboration and why they are successful is explained. The challenges and limitations of these interventions is also described, leaving possibilities for other ways that collaboration may be maximized for student learning open to investigation. Therefore, the next area of work reviewed provides support for further study of a "preparing-to-interact" method of structuring collaborative learning activities.

Two cognitive-based approaches are described in support of examining a "preparing-to-interact" method of collaborative learning: (a) the Interactive-Constructive-Active-Passive (ICAP) framework and hypothesis (Chi, 2009; Fonseca & Chi, 2011; Menekse, Stump, Krause, & Chi, in press) and (b) the Preparation for Future Learning (PFL) paradigm (Belenky & Nokes, 2009; Froyd, 2011; Schwartz & Bransford, 1998; Schwartz & Martin, 2004; Schwartz et al., 2007). The ICAP framework provides a tool for categorizing learning activities and hypothesizes learning outcomes based on the cognitive engagement of the student during the activity. The PFL paradigm supports the phenomenon that prior knowledge can be deeply activated through specific cognitive activities that can then prepare students for learning in future activities. In combination, these two approaches support the idea that students can be cognitively prepared to collaborate more effectively during learning activities, leading to better learning outcomes. This perspective takes into consideration students' readiness for engaging in discussion to learn. In addition, some of the work that has included preparation-type phases in collaborative activities is reinterpreted under a "preparing-to-interact" perspective and work that shows indirect evidence of the effects of a "preparing-to-interact" method is reviewed. Although there is indirect empirical support for this method, it has not been directly tested and especially not under the approaches of the ICAP framework and the PFL paradigm.

Problem Statement

There are certain conditions under which collaborative learning is more successful. Because collaborative activities are so commonly used in a variety of educational settings, it is important to acknowledge these conditions and understand how they affect students' likelihood of acquiring deep understanding of typically difficult concepts. Although a number of effective practices for implementing collaborative learning in classrooms have been investigated, there is still room for improvement. Using two cognitive approaches, my research aims to answer the question: How can students better prepare to cognitively engage in collaborative activities, leading to deeper learning? I investigate the effectiveness of "preparing-to-interact" on learning, where cognitive preparation in a domain topic precedes collaboration.

Research Questions

1. Does individual cognitive preparation in a specific domain topic prior to engaging in collaboration have an effect on learning outcomes after collaborating?
2. How does the type of task in which individuals prepare prior to collaborating affect learning outcomes after collaborating?

3. How does the type of task in which individuals engage while collaborating affect learning outcomes?
4. As related to the three questions stated above, how is deep learning affected differently than shallow learning?

Chapter 2

REVIEW OF LITERATURE

Collaborative Learning Works… Sometimes

To address how collaboration improves students' conceptual understanding, studies that have found direct evidence that discussion produces improved outcomes for difficult-to-learn material are reviewed below. Some of this work has investigated less structured forms of collaboration (i.e. discuss/work with a partner), while others have used highly structured forms such as when students share complementary knowledge (i.e. Jigsaw methods). Highly structured methods of collaboration that took place in both classroom and laboratory settings are summarized first, and then a grouping of studies that tested student learning in laboratory settings are reviewed.

There are also a number of ways that collaboration seems to fall short of its promises. When considering the potential or ideal outcomes that collaboration should produce, some work has found a collaborative inhibition effect, suggesting that the presence of others while engaged in a task can actually hinder performance. Other work has more thoroughly examined the communication processes that occur during collaborative activities, showing that when certain aspects of communication are not present or are not utilized well, collaboration is not effective for learning. This work is reviewed next.

Sharing Information Through Collaboration

In a general sense, "sharing complementary knowledge" simply refers to the idea that when individuals each carry unique forms of knowledge that are not sufficient for learning a concept in its (relative) entirety, allowing them to share those forms of knowledge should improve learning for all involved. This lies at the heart of Jigsaw methods, which places students in “expert” roles where each learns and focuses on a particular aspect of a concept, and then they are positioned in complementary roles in a “jigsaw” group that allows them to share that knowledge with each other. Although the Jigsaw technique is one common example of how students can share complementary knowledge, there are other ways to design these learning activities (an example by Kneser and Ploetzner, 2001, is described later in this section).

It might be too obvious to state that learning of any concept will be hindered when students do not have access to the information that they need to make sense of the concept. For instance, let us say that to understand concept C, a person must know the information contained in aspect A and aspect B. If that person is completely missing the information that A provides, then he/she will not successfully learn C with B alone. Therefore, teachers and instructors may solve this problem by assuring that students have access to the information contained in both A and B. This can be done in a number of ways, such as through lectures, providing reading materials, and facilitating whole-class discussions that address both A and B. When provided with all the aspects of information needed to learn, it becomes possible that students are then able to generate the proper inferences in order to make sense of the concept. However, traditional lecture classes (that often include large group discussion and supplemental reading materials) for instance, can fail to elicit this generation of knowledge in many students, hence, the great amount of research supporting the use of active learning techniques in the classroom (Bonwell & Eison, 1991), and one of the common reasons for using collaborative learning strategies.

A strategy where students receive and learn pieces of information individually and then engage in dialogues to share them, as with a sharing-complementary-knowledge approach, has the chance of increasing the likelihood that students will create and/or modify their mental models because of the opportunities dialoguing provides. Being able to ask one's peers questions, to offer and receive different perspectives, and to provide and listen to alternative explanations makes it more likely that students will generate new inferences as they talk through what another peer's knowledge contributes to their own conceptual understanding of the topic as a whole (Roschelle, 1992). In other words, this sharing-complementary-knowledge approach (sometimes referred to as a cooperative learning method) is based on the assumption that learners will be more likely to collaborate effectively when they need information from their partners.[2] Effective collaboration/discussion should allow students to reach fuller understanding of the topic-at-hand, thus, more cognitively engaging students in learning.

A study by Doymus, Karacop, and Simsek (2010), conducted at a university in Turkey, showed that a sharing-complementary-knowledge method of instruction in a college course using the Jigsaw model improved student learning of concepts in electrochemistry, a difficult-to-understand domain according to Finley, Steward, and Yarroch (as cited in this work), compared to a traditional teaching method. In the traditional teaching method, all of the necessary information was provided to students through lectures and reading material, and lectures included whole-class discussions and opportunities for students to ask questions, however, predominantly consisted of the

[2] *Cooperative* learning is sometimes distinguished from collaborative learning to refer more specifically to the division of labor between group members, but I consider it one method of collaborative learning, as similar to Rummel and Spada's (2005) assertion.

instructor presenting information didactically. In the Jigsaw class, students were separated into expert groups in order to study and learn a particular subtopic both alone and collaboratively, which gave students knowledge unique to their assigned subtopic. Then, in a new phase, students were placed into jigsaw groups, with each new group including a student of a different expertise. In these groups, students worked together to create group presentations for the class. They basically had to integrate their complementary forms of knowledge and, in a sense, teach their individual expertise to the rest of their group, and then jointly create a presentation that consolidated the various forms of information.

The Jigsaw class outperformed the traditional class on a variety of standardized assessments of scientific reasoning and electrochemistry that included multiple-choice items, free response items, and tasks requiring students to make external representations (drawings). These assessments, in particular the students' drawings and scientific reasoning outcomes, indicate mental model changes in students since they are more direct measures of individual internal knowledge. Researchers attribute these learning effects of the Jigsaw class to the discussion that took place in the collaborative groups. However, it may be possible that students in the Jigsaw class simply had more opportunities to engage actively with the learning materials and that this more active engagement is what drove learning, rather than the sharing and discussing of complementary knowledge.

A study by Pozzi (2010), although it does not compare working individually with working collaboratively with a partner or group, does provide insight as to whether it is the collaboration itself in sharing-complementary-knowledge approaches that affects learning or if it is the opportunity to more actively engage with the learning material.[3] Using data collected from two online college courses at a university in Europe, her study compared effects of a Jigsaw method with a less structured collaborative method. In the less structured method, students studied all aspects of a topic (e.g. educational technology and instruction) individually and then discussed individual ideas about the material in a collaborative group.

Protocol analyses of student online dialoguing and message board communication show that the students in the Jigsaw class demonstrated "richer" discussion (p. 72). Overall, they produced a greater number of expressions explaining or presenting their points of view, accepting other's ideas and coming to consensus during group interaction, connecting ideas or synthesizing contributions from multiple group members, and reflecting on the learning process (metacognitive). The conclusions from this process analysis, along with those from the assessment measures used in the Doymus et al. (2010) study, provide evidence that discussion resulting from sharing complementary knowledge does not merely engage students more actively in learning, but that the dialogue behaviors and content of the discussion, in particular, affect students' understanding of the material. To summarize, this combined work shows that sharing-complementary-knowledge approaches to collaborative learning influence dialoguing behaviors that lead to better learning outcomes.

Considering the goal of my work to test the effects of a "preparing-to-interact" phenomenon, I should point out that Jigsaw-type approaches to collaborative learning

[3] This study focused on the process of learning, thus, learning is indicated through dialoguing behaviors during collaboration, rather than on knowledge/achievement outcomes.

include a similar prepare-before-collaborating structure. However, I make a distinction between this and what I refer to as “preparing-to-interact.” To prepare to collaborate by studying the learning material beforehand, whether it be in a Jigsaw fashion or otherwise, does not indicate that the preparation was meant to intentionally engage students in a cognitive process that allows them to activate existing knowledge. It is typically implemented in a more common way (e.g. study this material so that you will be able to talk about it in a group). The preparation that I am referring to would be with a specific kind of learning task that induces a state of cognitive conflict and awareness in students (as related to the domain content) prior to the collaborative activity. Giving students opportunities to share complementary knowledge does seem to engage them more deeply in a dialogue leading to better learning, however, this approach implies that the need for information from partners is what drives student engagement (student who studied aspect A needs information about B from another student). It is not necessarily driven by preparation that activates existing knowledge.

Another study by Kneser and Ploetzner (2001) directly addressed the question of how complementary knowledge affects student discussion. Their work is based on analyses of student dialogues of a prior laboratory study conducted through a university in Europe, where high school students were taught a lesson on classical mechanics with either a qualitative-based instructional unit or a quantitative-based unit, and then a student from each unit formed a dyad to collaboratively solve difficult (beyond their competence) mechanics problems (Ploetzner, Fehse, Spada, & Kneser as cited in Kneser & Ploetzner, 2001). Learning measures of the mechanics material were obtained through group-level and individual problem-solving, pretests, and posttests. For the purpose of comparison, a control group was included where students worked individually on several problems throughout all stages of the study.

Results showed that collaborative groups produced significantly more solutions to problems than the control participants. In addition, assessments of student knowledge were taken a various stages of individual and collaborative work for the experimental groups. Through these multiple stages of assessing student knowledge, there was evidence that collaborating with a partner significantly improved problem-solving performance when compared to students' individual work in prior stages. Detailed protocol analysis of some of the dialogue cases allowed the researchers attribute learning to specific instances of dialogic moves including question-asking and explaining, reflecting on each other's ideas and solutions, using a large proportion of reasoning moves, and overall, displaying coherent discussion. Finally, they found that the type of complementary knowledge of the individuals within a dyad also differentially affected learning outcomes. Qualitatively instructed students seemed to learn far more from their quantitative partners, suggesting that explicitly teaching students differing kinds of knowledge should be done with careful consideration, because certain kinds of knowledge may be better for learning from a partner than other kinds.

In essence, because the learning task in this study required students to solve problems beyond their competence level, students would not have been able to accomplish the task unless they generated new inferences. This kind of generative behavior is indicative of improvement of students' mental models of the concepts, and the protocol analysis showed that more generative activity took place when students engaged in discussion with a partner rather than when they worked alone. Again, the fact that students had complementary knowledge to share seemed to encourage meaningful discussion that affected learning.

To summarize, clearly students need access to all the relevant aspects of information to learn a concept. (By "relevant," I am referring specifically to the minimal pieces of fact-based information needed to construct meaning of a larger concept.) Learning these aspects during discussion, with an approach that gives students reason to share their complementary knowledge, seems to benefit students more than having full access to the information through lectures and/or readings, even when collaborative discussion follows the initial acquisition of that information (Pozzi, 2010). Compared to less structured forms of collaboration, it seems that sharing complementary knowledge through discussion is more likely to encourage students to relate alternative ideas to their own, generate new inferences (resulting from partner contributions), and elaborate and/or change their knowledge structures to make sense of new information. It is important to note that the studies referenced here were all conducted in European countries, and therefore, conclusions may not necessarily be generalized to other educational systems (such as in the United States). However, a number of other studies conducted in the U.S. and other parts of the world are highlighted throughout this review and provide more examples of how collaboration benefits student learning.

Working Alone Compared to Working with a Partner

A series of laboratory studies on solving difficult physics problems showed that peer collaboration for a college population in the U.S. improved learning outcomes above one well-known effective learning strategy called self-explanation (see Chi, Bassok, Lewis, Reimann, & Glaser, 1989, and Chi et al., 1994, for work on self-explanation).[4] These studies compared self-explaining to jointly-explaining with a partner, while using an intelligent-tutoring tool to solve problems (Hausmann et al., 2009; Hausmann, Van de Sande, Van de Sande, & VanLehn, 2008; Hausmann, Van de Sande, & VanLehn, 2008).

When students explain or ask questions, as with self-explaining or engaging in a dialogue, they are more likely to recognize gaps in their knowledge and generate the appropriate inferences that make connections among various knowledge components, leading to the construction of more accurate and complete mental models (Chi, 2000). The explanations that students make from engaging in a self-explanation task can be used as indications of how students' knowledge structures change or improve, as can the explanations that arise from discussion. Thus, these physics studies were able to single out the effect of peer discussion on student understanding by having students in control conditions self-explain (an already generative task that has been shown to affect students' mental models), and having students explain to each other in experimental conditions. Knowledge was assessed via domain-based tests and the student explanations themselves to indicate the direct effect of collaboration on learning.

The use of the intelligent tutoring tool in these studies allowed for additional assessments of student understanding at various stages of problem-solving to be made. For instance, the tutoring tool could provide "hints" to students when they reached an impasse. In other words, students could select a hint from the computer tutor when they were no longer able to make sense of the information or the next step to take in solving

[4] Self-explaining works by prompting students to explain out loud, in their own words, portions of what they read. It is meant to encourage students make sense of information according to their unique mental models.

the problem. One analysis of this work showed that students in self-explain conditions selected hints twice as often as those in joint-explain conditions. Thus, being able to discuss the problems and solution ideas with a partner allowed students to figure it out on their own, and generate the inferences needed without asking the tutor to provide the answers. Analysis of students' dialogues provided evidence that students did, in fact, generate new knowledge (that was not explicitly present in the information presented by the tutor or learning materials) and reasoned through their confusions and uncertainties by discussing them with their partners well enough to overcome impasses. The discussion between the students seemed to substitute as a sufficient replacement for the hints that could be chosen from the tutoring tool.

Another important result from these studies in physics found evidence for a strategy that appeared to prompt students to make deeper explanations during their discussion (which can serve as an indication of the depth of student understanding). Hausmann, Van de Sande, and VanLehn (2008) discussed how the instructions to complete the learning tasks, in retrospect, may have given students a form of preparation time for thinking more deeply about the domain content prior to discussing it with their partners. These instructions were to first solve a problem with the aid of the tutoring tool. Then, students were given the opportunity to study an expert's solution of an isomorphic problem. Subsequently, during the joint problem-solving task, students could reference both their confusions or struggles during the initial problem-solving with the computer tutor, as well as how those confusions related to the expert's solution.

The dialogue transcripts that showed deeper explanations of the physics concepts tended to include references to comparisons between the expert's solution step and the student's own alternative idea of the solution step. Considering that students in self-explain solo conditions also had the opportunity to take advantage of this preparation, and could have compared their confusions, alternative ideas/solutions, struggles, etc. with the expert's solution steps, a question remains: What was the benefit of being able to jointly-explain with a partner? One aspect of evidence addressing this question is that the students in the solo conditions asked for far more tutor hints during their problem-solving. Not having a partner with which to engage in discussion seemed to hinder students' likelihood of explaining in a deep way, as the self-explainers were quicker to ask the tutor for help, rather than to work through it. In other words, although the first two study tasks may have provided a form of cognitive preparation for all the students, it only translated to deeper learning for the joint-explain students.

Additional support for the idea that engaging in peer discussion leads to deeper and more useful explanations was found in Hausmann et al.'s (2009) work. Their analysis of verbal protocols showed that although there was no difference in the number of explanations produced between solo and dyad conditions, there was a negative correlation between explanation statements and number of errors in problem-solving in only the dyad condition. In other words, the discussion that took place during the problem-solving helped students to avoid errors, whereas students' self-explanations in the solo condition were "remarkably unhelpful" (p. 2). It is possible that this result would not have occurred if the students did not have access to the tutor hints (i.e. if they did not have an opportunity to obtain the right answers). What is of interest here is that even though there was an option to access the right answers, students avoided that strategy more often when they had the opportunity to engage in peer discussion.

Chi et al. (2008) evaluated student learning of difficult physics concepts in various collaborative and solo conditions where college students at a U.S. university were required to solve problems either with the aid of a textbook or by observing human tutoring sessions via video. Results showed that collaborative conditions were superior to solo conditions, as assessed by deep learning measures via pre- and posttests (i.e. assessment questions that specifically targeted deep understanding of the concepts, rather than correctness of more superficial, fact-based knowledge). Moreover, collaboratively observing was shown to be as beneficial to this kind of deeper learning as being tutored, suggesting that peer collaboration as an instructional strategy has the potential to reach the gold standard of one-to-one instruction.

More recent work has also found collaboratively observing to improve learning above solo activity on the topic of molecular diffusion in middle school students. Results from a classroom study conducted by Muldner, Dygvib, Lam, and Chi (2011) in a large U.S. city showed that collaboratively observing videos (of either a tutoring session or lecture) while studying diffusion produced higher learning gains than observing videos individually. It is important to note that students in the solo conditions were not just passively watching the videos. In all conditions, students were instructed to engage with digital simulations that demonstrated the process of diffusion, thus, collaboration proved beneficial above individual active learning.

Another example of work that compared partner and solo conditions found that college students became better at summarizing research articles (a challenging task because scholarly articles are difficult to synthesize and evaluate according to Taylor, as cited in this work) when they worked in dyads to detect errors in an experimenter-produced summary of an article (Gadgil & Nokes-Malach, 2010, 2011). Of particular interest regarding the cognitive processing of error detection, this summary included "structural" errors, which were more difficult for students to detect since they required a thorough and deep understanding of the article. For instance, students needed to understand the article relatively well in order to detect a misstated research question in a summary (2011, p. 4). (In contrast, a superficial error might have been an incorrect formatting of statistical findings.)

Student understanding was measured through their revisions of the error-prone summary and in a later summarizing task. Results showed that dyads found more structural errors than individuals during the error detection activity, and students who had worked in a dyad produced better article summaries in a later individual task. In addition, these results were compared against calculations of nominal pairs and showed equivalent outcomes between real dyads and nominal dyads. Nominal calculations of pairs or groups are used to determine the potential success that working with others can produce.[5] Thus, students learned how to better read and process difficult research articles in order to produce accurate summaries, by engaging in a collaborative versus a solo task.

Abstraction as Deep Understanding

Students' ability to develop abstractions is another way to assess deeper understanding of concepts. A few studies have used situated learning contexts (real-world situations/settings or learning-through-application, Greeno, Collins, & Resnick, 1996) to examine if students could develop abstract representations, principles, or rules by

[5] Nominal group measures are basically calculated by using specific statistical analyses that deal with all the possible random pairings of individuals' scores to create a mean representative of how those individuals would/should have performed, had they actually worked in a group.

engaging in concrete problem-solving or discovery-type tasks. In three separate experiments, Schwartz (1995) found that dyads that worked collaboratively were more likely to generate abstractions compared to individuals. In a laboratory experiment, high school students had to envision a horizontal chain of connected gears and figure out which direction the last gear would turn based on the direction the first gear was turned. This can be easily done in a concrete way by using hand gestures to "turn" each visualized gear when there are relatively few gears in the chain. This can also be done by using a parity rule, such that when there are an odd number of gears, the first and last will turn in the same direction and when there are an even number, they will turn in opposite directions. In this study, students were given several problems that used chain lengths of three to nine gears, and then a final problem that used 131 gears (which could not be solved unless the parity rule was discovered).

In assessing how many students induced the rule on their own, results showed that dyads induced the rule four times more than the individuals. Two additional experiments by Schwartz (1995) showed similar outcomes in middle and high school students in classroom settings using authentic lessons, such that students who worked in dyads created more abstract representations of concepts (graphs or matrices), while those who worked alone tended to draw pictures that concretely described concepts without representing the relations among them.

Similarly, Shirouzu et al. (2002) conducted a study in Japan where they asked college students to solve a fraction problem (i.e. What is three-fourths of two-thirds of the area of a square?) by using different concrete materials (origami paper, cardboard, acrylic board, pencil, marker) as the students saw fit. They were interested in assessing how quickly the students would recognize an abstract way to solve this problem by using a simple mathematical equation (3/4 of 2/3), which is a far more efficient way to solve the problem. They compared solo and pair conditions in two trials, one trial presenting it as three-fourths of two-thirds, then the other trial presenting it as two-thirds of three-fourths (to avoid order effects, the presentation of the two versions of the problem were equalized across conditions).

Nearly all the students used a non-mathematical strategy to solve the problem in the first trial (folded the origami paper, shaded-in and erased the board, etc.). However, by the second trial, 10% of students in the solo condition shifted to the mathematical formula, while 46.7% of the pairs shifted to the mathematical formula. In both this work and Schwartz's (1995) work, the proportion of dyads who used abstractions was also significantly higher than a nominal pair calculation (see footnote 5). Thus, engaging in discussion seemed to lead to a greater number of instances of developing or using abstractions in comparison to working alone.

So, why would pairs behave differently than individuals with regard to inducing or realizing an abstraction from concrete learning tasks? The Shirouzu et al. (2002) study provides further insight as to the benefit of being able to engage in a dialogue for considering ideas at an abstract level. They conducted protocol analyses of the conversations of student pairs and showed that the abstractions arose after particular instances of role-switching within an episode of discussion. They referred to these roles as "monitor" and "doer" roles. Basically, one student within the pair could adopt a doer role and work directly with the materials (the paper, the cardboard, etc.), while the other was left to monitor, or observe what his partner was doing. The monitor could then gain insight from watching his partner and then offer suggestions, or switch and then try out his own ideas directly on the materials. This would then leave the initial doer now in the monitor role. After several instances of this role-switching, the monitor of-the-moment would basically realize that the problem could be solved very efficiently by using a mathematical formula.

Thus, the benefit of the collaboration seemed to come from the fact that the students could work directly with the material, and then reflect on the content when watching their partner. In other words, the “forced” moments for reflection that occurred as a result of having to jointly work on the task led to the generation of knowledge that connected the concrete aspects of the task to something abstract. Individual students had no reason to stop and reflect, and also did not have the benefit of a partner’s perspective. It might be possible to obtain similar results for individuals if there was imposed time for reflecting and the access to alternative ideas. However, the main point is that collaborating, as was done in this study, provided a natural setting for reflection and consideration of alternative perspectives, which seemed to lead to more instances of thinking abstractly.

To conclude, there are many examples that support the use of collaboration and peer discussion over individual engagement in a learning task. However, the next section reviews work that has shown mixed results regarding the benefits of collaboration on learning. The kind of task in which students cognitively engage while collaborating, how deeply those tasks engage students in learning, and the quality of communication between collaborators are all factors that can influence the positive effects of interacting with others.

When Collaboration Falls Short of Its Potential

One example of mixed results shows that the presentation of the domain topic can effect how collaboration influences learning. Craig, Chi, and VanLehn (2009) examined solo and collaborative conditions in an observational learning task, where young adult students from the U.S. Navel Academy observed instructional videos on solving physics problems, while actually solving problems using an intelligent tutoring tool. Three conditions were compared: (a) solving problems individually while observing experts solving problems in worked-example videos, (b) solving problems in pairs while observing worked-example videos, and (c) solving problems in pairs while observing expert human tutoring videos. In this study, there was no difference in outcomes between solo and collaborative conditions when observing the worked example videos. However, collaboratively observing the human tutoring videos produced better learning than observing worked-example videos either alone or collaboratively (further supporting Chi et al.'s, 2008, work on the instructional value of collaborative observation of tutoring sessions). Thus, whether collaboration benefitted students depended upon the type of models students were able to observe.

Collaborative inhibition. The term collaborative inhibition has been investigated in the literature showing that collaborating actually impedes learning in some situations (Rajaram, 2011; Weldon & Bellinger, 1997). In most of this work, learning refers to memory (as opposed to understanding), and is measured by free recall assessments. One example of collaborative inhibition comes from a study by Blumen and Rajaram (2008) where undergraduates studied a random list of words and were then asked to recall the list either individually or collaboratively in groups of three. The researchers then created nominal groups of three from participants who recalled the words individually (i.e. their individual outcomes were pooled to indicate a potential or ideal outcome that collaboration should produce) and compared those nominal groups to the collaborative groups. In this work, the nominal groups outperformed the collaborative groups, recalling 68%-70% of the items compared to 54%-56%, respectively.

Considering my interest in "preparing-to-interact," the type of free recall assessments mentioned above would be poor indicators of learning. After all, "learning" from memorizing a list of meaningless words is not the same "learning" from figuring out how to apply the physics concepts of force and velocity to real-world problems, or from relating molecular behavior to how the process of diffusion works. Thus, the collaborative inhibition effect seems hardly relevant to studies of learning that address understanding, especially at the structural mental model level. However, using a collaborative inhibition perspective and considering nominal group calculation comparisons can be a useful indication for the potential best outcomes that collaborating should produce in a given domain. Work by Gadgil and Nokes-Malach (2010, 2011) borrowed from the literature on collaborative inhibition to assess how outcomes from collaborating compared to nominal group calculations in what can be categorized as surface-level versus deep learning tasks.

As mentioned earlier, Gadgil and Nokes-Malach (2010, 2011) compared solo and collaborative outcomes of error detection in a summary of a research article, and how that related to college students' ability to understand and summarize research articles in general. In addition, they also used nominal group calculations to compare how dyads performed relative to an ideal outcome. They found a collaborative inhibition effect for the detection of superficial errors. For instance, when they examined how many superficial or surface-type errors (incorrect formatting of statistical findings, etc.) were found, the real dyads found significantly fewer errors compared to nominal measures. This is not a surprising result if one considers the task of detecting superficial errors as requiring a similar depth of cognitive processing as memorizing a list of random words. Noticing superficial errors would not require any depth of understanding of the content of the research article. However, when the researchers looked at structural errors (a misstated research question), the real dyads detected as many errors as the nominal calculations showed. Thus, not only did collaboration lead to detection of more structural (deep) errors compared to working alone, the collaborative error detection was also comparable to an ideal for what collaboration should produce. In essence, there was a "failure" for students to take advantage of collaboration when the task was a surface-level task, but the students succeeded in taking advantage of collaboration when the task required deeper processing of the material. (This "failure" may have been due to students missing the opportunities to utilize collaboration effectively or choosing not to/being less motivated to do so; however, whatever the reason, the point is that collaborative inhibition occurred only in the surface-level task.)

Inadequate communication.[6] In studies that assess student achievement, ability, or understanding, peer collaboration has been shown to fail when students do not communicate in certain ways. Barron's (2000, 2003) work on "joint attention" provides insight regarding why students do not always benefit from collaborating. Joint attention

[6] The remaining studies reviewed in this section assessed meaningful learning outcomes, rather than superficial outcomes, to avoid confusing "learning" with "recall" or "memorization."

refers to group members' ability to attend to each other and to the instructional task at the same time, and especially during what Barron calls "critical problem-solving" moments. Through a classroom study, Barron assessed high-achieving 6th graders' ability to solve difficult mathematical word problems and showed that successful groups (those who produced more correct solutions and ideas in a problem-based task) showed more instances of jointly attending compared to unsuccessful groups. In this study, even though all students were highly capable of successfully working through the problems, some groups failed despite the individual students' typical prior achievement successes. These failures were directly attributed to the effectiveness of the groups' communication.

Not only should group members attend to each other, but each member also needs to provide substantive contributions towards completing the task. Barron's (2000, 2003) work found that instances of joint attention related to whether an individual group member ignored/rejected or accepted/"picked up" the ideas of others. When there were more instances of jointly attending, there were also more instances of group members acknowledging and incorporating fellow peers' contributions. Evidence for incorporating a peer's ideas into one's own understanding was found in verbal protocols, showing that a significant number of utterances were elaborations, modifications, or agreements of a group member's idea. The problem solutions provided by the students indicated better learning for groups who were better able "pick up" the ideas of others.

Volet et al. (2009) provide additional support for the collaborative success that occurs when all group members provide substantive contributions and consider the contributions of each of the members. In describing effective collaboration, they introduced the idea of "high-level co-regulation," which relates to joint attention and to the content-based contributions of group members. These researchers defined co-regulation as representing “episodes in which multiple group members made verbal contributions” and these verbal interactions had to have been related to the learning content of the task (p. 132), similar to joint attention as attending to group members and the task simultaneously.

These researchers examined case studies of collaborative groups from a university in Australia and showed that when individuals within groups all contributed a relatively equal number of ideas, they also used effective collaborative behaviors more often, such as explaining and questioning, and overall displayed better communication. Groups who had one to a few dominant members often left others out of the conversation or simply ignored their ideas (i.e. did not jointly attend), and seemed to treat the task as an individual activity rather than a collaborative one. These groups produced inferior performance measures (diagnoses and treatment of animals through authentic clinical case files in veterinary science) compared to groups whose members equally and substantively contributed to the discussion.

Fogel also supports the idea of mutual contribution in his writings on co-regulation, referring to instances when “individual joint actions” blend together “to achieve a unique and mutually created set of social actions” (1993, p. 6). Although his work is founded on theories of early development of social behaviors, it is applicable to collaborative learning because one goal of engaging in discussion to learn is to work together to gain meaningful understanding of a new idea. Some of the research on collaboration focuses on reaching shared meaning through discussion, where individual contributions “blend together,” or “converge” to create a joint discussion space

(Roschelle, 1992; Teasley & Roschelle, 1993). Although a person does not necessarily need to achieve shared meaning with another person to create his/her own meaning, effective communication is still necessary to obtain benefits of improving understanding through discussion. Fogel's "blended actions" refer to smooth communication, which collaborative learning work has operationalized as equal contributions of ideas from group members, a back-and-forth engagement of individual and joint work or reflection and action, and acceptance and elaborations of other's contributions.

Effective communication allows individual contributions within discussion to become relevant to the others in the group/partnership. This may lead to deeper engagement in learning and better understanding of concepts for each person involved. Through verbal protocol analyses, Kumpulainen and Kaartinen similarly concluded that "coordination of communication" during collaboration facilitates positive outcomes and that successful collaboration can be characterized by "symmetric interaction" (2003, p. 367).

Another factor of communication that has been found to affect learning from collaborating is in the coordination of group activity. Hermann, Rummel, and Spada (2001) and Rummel and Spada (2005) conducted laboratory studies with advanced medical and psychology students at a European university. They found that dyads who could better coordinate their group activity, such that the time-on-task and division of labor was decided upon and managed by both partners, produced better diagnoses and therapy plans for patients in a psychological cases activity. These studies assessed coordination management through the percentage of time devoted to individual versus joint activity, showing that when an equal proportion of time was spent across individually working and collaborating, learning outcomes were best. For instance, when partners interacted for a period of time, then shifted to work on individual tasks, then shifted back to joint activity, then back to individual work, and so on, they produced better outcomes than partners who spent the majority of their time working jointly. Collaboration was actually worse when there was "too much" joint activity occurring.

When too much joint activity is occurring, there may be little time for individuals to reflect upon the concepts. It may be that the relatively equal amounts of individual time that some groups coordinated provided opportunities for each partner to reflect upon the joint work (as similar to the Shirouzu et al., 2002, work on the reflection time available when partners switched between monitor and doer roles). Thus, it seems that a balance of joint work and individual work may better maximize collaborative learning outcomes.

Considering that one benefit of the individual "thinking time" is to reflect upon the domain content and also upon the joint work, the question that remains is: Does collaborating provide further benefit above providing opportunities for reflection? One simple answer is that engaging in a dialogue gives each individual involved in the conversation more content to think about. The presence and contributions of others brings about new perspectives, unique knowledge, and/or alternative ideas. Therefore, when the individuals within a group are all (relatively) equally contributing new ideas, engaging in the work, and sharing their prior knowledge, the occurrences of individual work time allow for reflection of not only one's own ideas, but also of others' ideas.

Some recent work offers an alternative hypothesis for why collaboration can be particular helpful above working alone, especially in reasoning tasks. Lin et al. (2012)

analyzed elementary students' classroom discussions in a rural area of the U.S. to find a "snowball" effect of children's use of analogies. Creating/using an analogy to reason through an idea is a "constructive" behavior (Chi, 2009), indicating the generation of new knowledge above and beyond the learning material. In this study, students engaged in discussion to answer open-ended questions about several stories (either containing a controversial issue, or difficult decision that a character needed to make, etc.). They found that once an analogy was introduced during discussion, more children began to use them, and they occurred with increasing frequency as time went on.

Thus, in collaborative situations, dialogues have the potential to kick-start effective learning behaviors (explaining, analogizing, elaborating, etc.) that may spread amongst group members. The symmetrical interacting and equal contribution of members of a group that have been shown to be present during successful collaboration, might also be indicative of a similar snowball phenomenon. When one person begins to deeply engage, and then another deeply responds, this can lead to meaningful ideas arising more frequently in conversation, which leads to more opportunities to give meaningful responses.

A review by Janssen, Kirschner, Erkens, Kirschner, and Paas (2010) addressed the process-oriented (qualitative) research on collaborative learning to further support the idea that collaboration will not meet its potential when communication breaks down in these aforementioned ways. The breakdown (or lack) of communication that often occurs in collaborative learning settings inhibits the meaning that can be discovered (individually or as shared meaning) through discussion. It is not enough that each person in a group or partnership engage in effective cognitive behaviors (explaining, questioning, arguing, elaborating, etc.), but that they are attending to, accepting, contrasting, comparing, etc. the content-based contributions of others. In other words, individuals can experience learning benefits from engaging in explaining, questioning, arguing, analogizing, etc. without necessarily engaging in a discussion. (Examples can be seen in self-explanation, Chi et al., 1989; Chi et al., 1994; Fonseca & Chi, 2011; self-questioning, King, 1992; and even self-arguing, Asterhan & Schwarz, 2007, coined the term "monological argumentation.") The specific benefits of collaborating come when individuals can incorporate the explanations, questions, arguments, analogies, of a partner into their own understanding, allowing their own mental models to be altered or improved by the contributions of another.

Summarizing When Collaboration Works

Peer collaboration is most likely to enhance learning when students can link their partner's contributions during discussion to their own prior knowledge, ideas, and claims. Thinking about the relationship between a partner's ideas and one's own and externalizing responses during discussion leads students to cognitively engage more deeply in learning. Dealing with difficult concepts or abstract principles and/or being in a position to share complementary knowledge provide greater opportunities for students to engage in meaningful discussion; discussion where students are incorporating another's contributions into their own knowledge structures. The benefits of collaboration are lost when partners do not consider how each group member's ideas relate to their own (in other words, when there is a low quality of communication amongst group members). Discussion then becomes more superficial and leads to shallow processing of the domain content. In addition, there seems to be little evidence for the benefit of collaboration when the goal is to memorize surface-level information.

To address how students can be helped to avoid poor communication during collaborative activities, and how they can be encouraged to connect a partner's contributions to their own ideas and thoughts, researchers have developed strategies to help students take better advantage of the benefits collaborative learning activities offer. The next section reviews studies that have attempted to improve student collaboration towards learning.

Taking Advantage of Collaborative Opportunities

Upon review of the work that has resulted in recommended practices for encouraging effective collaboration, three main types of interventions arise: (a) teaching students collaboration skills, (b) externally guiding student interactions, and (c) providing opportunities for students to work through ambiguous or open-ended tasks. How and why these interventions are successful in improving student learning from engaging in collaborative activities are described below. In addition, the limitations and challenges of each are addressed, providing support for investigating an alternative intervention, namely, "preparing-to-interact."

Teaching Collaboration Skills

One area of collaborative learning research that has found positive results has examined the effects of teaching students specific collaboration skills prior to collaborating. More specifically, this work has shown that teaching students certain skills can promote the use of those skills in subsequent student interactions, leading to enhanced learning. This is in contrast to directly guiding student interactions while collaborating or setting up certain task conditions to elicit effective collaborative behaviors, which are covered in later sections. This teaching approach typically involves directly training students in skills such as arguing, elaborating, explaining, and asking deep questions (which are behaviors that have been shown through prior work to be beneficial). Receiving instruction to develop or use such skills has been shown to increase the frequency of targeted collaborative behaviors, and to improve learning outcomes, compared to control conditions where students do not receive targeted instruction on how to collaborate.

One might argue that these skills are not necessarily “collaborative” skills, but simply individual behaviors that happen to be present (sometimes) in dialogues, and the benefits of these skills come to the individuals who use them. Thus, we come back to the question: What is it about interacting, *per se*, that is beneficial? I argue that imbedding these skills under a collaborative learning perspective allows us to see the benefit through the ways they are used in communication. For example, joint attention and co-regulation are partially operationalized by instances when group members elaborate on “each other’s” ideas, ask questions to “one another,” and respond with explanations “to someone else.” Therefore, these are collaborative skills when discussants are effectively communicating (listening to each other, considering each other’s points/perspectives, jointly attending, co-regulating the collaboration, etc.). The work covered in this section approached teaching these skills in a communicative context.

This instructional teaching-skills approach treats collaboration itself as a domain in which students can acquire knowledge. In other words, this approach implies that students can learn how to collaborate better when they are given some form of instruction about what makes collaboration work. Slavin's (1992, 1996) reviews on collaborative learning suggested that students must be taught or guided to effectively collaborate. The most straightforward way to teach students about collaboration is to give them specific instructions on how to interact. For example, work on argumentation has provided evidence that instructing students how to argue produces better quality argumentative dialogue compared to non-trained controls who engage in "natural" discussion (Asterhan & Schwarz, 2007, 2009).[7] Other work has supported instruction-to-argue by testing how specific instructions on isolated aspects of arguing (presented as differing goals such as persuasion, rebuttals, or exploratory talk) affect the quality of students' arguments (Nussbaum, 2005). In the simplest form, instruction may include providing students with information on what argumentative dialogue consists of (such as justifying claims with support or evidence, asking critical questions, or defending a position) and then asking students to engage in a subsequent dialogue in these specific ways.

Work by Hausmann (2006) and Hausmann and Chi (in preparation) took a similar approach in training dyads to elaborate by providing simple instruction about elaborative techniques (e.g. make incomplete ideas explicit, extend a partner's ideas), but bolstered training by including a warm-up session for students to practice the techniques they learned. In this study, subjects were college students at a U.S. university and the domain of interest was physics in an engineering context. In the warm-up session, student dyads engaged in an interaction while the experimenter assessed their use of elaboration and provided corrective feedback and intervened when necessary. Results showed that dyads

[7] Argumentation is described in more detail in the Guiding Peer Interactions section.

trained to elaborate produced significantly more elaborative statements than un-trained dyads and that the number of elaborative statements positively correlated with learning outcomes. In general, the instructional training positively affected student interactive behaviors and improved learning.

Other work has used training methods for reciprocal questioning with undergraduate and graduate students in an educational methods course (King, 1990), co-constructing math diagrams with junior high students (Uesaka & Manalo, 2011), and explaining (Bielaczyc, Pirolli, & Brown, 1995, on university students self-explaining during programming tasks and Wong, Lawson, & Keeves, 2002, on high school students self-explaining math word problems). They show similar outcomes of increased frequencies of targeted behaviors and enhanced learning. Thus, it seems that training students of a variety of ages in particular learning skills to use during collaboration improve domain-based learning outcomes compared to collaborating without training.

Some noteworthy work regarding how students can acquire knowledge about collaboration, by Rummel and Spada (2005), and Rummel, Spada, and Hauser (2009), directly measured collaborative knowledge through test assessments in addition to examining student interactive behaviors. These researchers took a vicarious learning instructional approach by training students to collaborate through observation of a model. They had students watch a video of two students engaging in what they deemed to be an "exemplary" dialogue prior to the collaborative task. Exemplary collaboration was referred to as a dialogue that included reaching common ground (Clark & Brennan, 1991), sharing/"pooling" information and complementary knowledge (Dillenbourg, 2002), and coordinating individual and joint work loads (Hermann et al., 2001). These factors of dialogue have been shown to improve learning.

Two main types of instructional interventions for college students from a European university were investigated: (a) observation of a model and (b) scripted guidance in a practice session prior to the main learning task (diagnosis and treatment of a psychological case). Compared to control dyads (those who did not receive training, but still worked in collaboration), the pairs who observed students collaborating in a video and who practiced with a script beforehand showed improved outcomes in collaborative knowledge, in actual discussion, and in their diagnoses and treatment plans for the psychological cases. These results show that either observing effective collaboration or practicing with a script allows students to acquire knowledge about collaboration, fosters effective collaborative behaviors, and leads to improved performance in domain-based tasks. The reasons why scripting can maximize collaboration are discussed in a later section. Below are speculations for why observation of a model might foster effective collaboration.

Regarding the application of collaborative knowledge, I discuss two possible reasons why observation of effective collaboration may have led to enhanced outcomes. One, because students were informed that they would watch a video of two students collaborating, this may have cued-them-in to the student behaviors. In fact, the researchers describe “enhancing” their interventions in the 2009 study by included prompts that directed participants to attend to certain behaviors, such as asking questions to clarify shared knowledge, and to reflect on and explain to themselves what made the observed collaboration successful (Rummel et al., 2009). However, a more general explanation for why observation of a model can teach collaboration skills is included next.

Students may inadvertently pick up helpful interactive strategies simply because they are available to observe. Chi (in press) and Muldner, Lam, and Chi (in revision) have found positive correlations between observed interactive behaviors and enacted behaviors. In these studies, observers were not instructed to attend to any particular student behaviors in the videos, nor told that videos should help them interact more effectively. In fact, they were only told that the videos would help them to learn the domain content. One result from this work showed that in a collaborative problem-solving task, the number of joint-explanations made by dyads positively correlated with the number of joint-explanations they observed in tutorial dialogue videos (Chi, in press). Similarly, Mulder et al. (in revision) found that the number substantive contributions dyads made correlated with the number of substantive contributions they observed. Thus, the observers may have unintentionally learned effective interactive skills and used them in their discussions. In fact, Rummel et al. (2009) found that observation of a model without prompts outperformed scripted conditions both with and without prompts.

Thus, observation of beneficial collaborative interactions can serve as an instructional approach that helps maximize subsequent collaboration. However, it is important to note that evidence for this collaboratively observing benefit has been found more strongly in college student populations. Mulder et al. (in revision) did not find as conclusive of an effect for middle school students, at least compared to being directly tutored. It is important to consider the age/grade level of subject samples when generalizing such findings.

To summarize, several instructional approaches have been used to train students how to effectively collaborate, and help maximize the benefits of collaboration, by focusing students on helpful behaviors that lead to learning. This can be done in a direct manner (instructing students about the targeted behaviors), through letting students practice behaviors prior to collaborating, or by allowing them to observe effective collaboration in action. Since other work has shown that students do not always naturally collaborate successfully (Barron, 2000, 2003; Volet et al., 2009; Yetter et al., 2006), providing instruction on collaboration can bolster effective interacting (Cohen, 1994). After all, students may not know which interactive behaviors lead to learning or they simply might not think to use them. Training/instruction informs students about effective collaboration, as well as triggers them to use effective collaborative skills during discussion. Benefits of training students in specific collaborative skills have been found in elementary to college populations, however, there are still questions as to how specific types of training or which particular skills are most effective in younger learners compared to adult learners.

Limitations of Teaching Collaboration Skills

Although these interventions have had positive effects on collaborative learning, other work has shown that effective collaborative behaviors can decrease after time. Studies by Webb and colleagues found that training middle school students in the U.S. in help-seeking and help-giving behaviors (such as in giving and receiving explanations) can improve collaboration and learning outcomes in some settings, but when the classroom culture is highly "teacher-centered," students are far less likely to adopt these behaviors for the long-term, even with intensive training (Webb & Mastergeorge, 2003;

Webb, Nemer, & Ing, 2006). Webb argues that teachers, as models, have a strong impact on student behaviors and that when teachers do not exemplify useful interactive skills and rather engage in primarily didactic forms of discourse with students, students will mimic this didactic discourse with one another despite collaboration training. Thus, a major question remains: Is it worth the time and effort to train students in collaborative skills if the use of these skills fades after time? Therefore, partly in response to addressing such a challenge, other work has investigated guiding peer interactions throughout the collaborative activity by use of scripts and prompts.

Guiding Peer Interactions

Guiding-peer-interactions may be defined as an intervention that provides external support to students to help them structure their discussion in particular ways. In studies examining these types of interventions, researchers typically focus on specific interactive behaviors or moves, such as question-asking, explaining, or arguing, and structure those behaviors using scripts or prompts. Scripts are predetermined sets of guidelines that may offer differing roles for students, specified phases of collaborative activity, explicit dialogue patterns to follow throughout discussion, or instructions to coordinate problem-solving (Dillenbourg, 2002). Prompts provide students with cues for continuing discussion and may direct students to ask their partner to elaborate, take turns speaking, explain their reasoning, or ask a partner to justify a claim (GE & Land, 2004; King, 1992, 1994; Palinscar & Brown, 1984). Scripts and prompts are sometimes used in tandem, and discussed somewhat interchangeably in many studies, but point here is that they are both external support devices for structuring student discussion.

The motivation behind externally structuring collaborative behaviors like explaining, questioning, and arguing is often driven by the evidence that these behaviors have been found to improve learning and performance. For instance, explaining one's ideas or assertions to another person drives him/her to construct and build knowledge, and this can lead students to repair inconsistencies in their thinking and develop deeper or fuller understanding of the domain concepts (Chi, 2000; Roscoe & Chi, 2007). Relatedly, questioning encourages students to focus attention, check their comprehension of learning material, organize new information, and integrate it with existing knowledge (King, 1992). Arguing allows students to consider, confront, and evaluate differing viewpoints, and can motivate students to learn through reconciling these viewpoints (Andriessen et al., 2003; Schwarz & Linchevski, 2007). Additionally, structuring student interactions should improve student-to-student communication by offering guidance when students may not know what to say next or what direction to turn the discussion (Coleman, 1998; King, 1999). The following sections review the literature that has investigated how scripts and prompts promote beneficial cognitive behaviors and can be successful interventions for improving learning through collaboration. Studies examining scripting and prompting are first covered, and then argumentation is addressed as a particularly beneficial form of guided discussion.

Scripting interactions. Scripting student collaboration has been studied for decades. O'Donnell (1999) described how highly structured interactions (such as with scripted cooperation), where students alternate between different roles (e.g. listener and speaker) based on specific cognitive activities like explaining, questioning, detecting errors, or summarizing, can enhance collaboration. In more recent work, Dillenbourg and

Hong (2008) described how macro-scripts can be used to scaffold student interactions towards better collaboration in computer-supported collaborative learning (CSCL) environments. Macro-scripts assign specific roles or knowledge/expertise to students or serve as a way to group students by individual opinions or traits. Macro-scripts take instructional strategies such as Jigsaw and reciprocal tutoring methods (Palinscar and Brown, 1984) and frame them in a computer-mediated interface that students can use via electronic communication (at separate computers) or face-to-face (sitting together at the same computer). Other work has also found support for using computer-mediated scripts to improve collaboration by basing their design and use in classrooms on cognitive knowledge building theories. For instance, Nussbaum et al., (2009) showed that computer-supported scripts could successfully scaffold students who were not familiar with working together to engage in co-construction and reach consensus to solve problems through shared understanding. This study assessed outcomes from middle school students in English, Art, and Math classes in the United Kingdom and from high school students in Science and Math classes in South America.

Scripts are beneficial because they can help prevent one partner from dominating the task and discourage "social loafing" (i.e. putting forth minimal effort while assuming the group or partner will pick up the slack, Latane, Williams, & Harkins, 1979). In addition, they encourage students to engage in effective learning behaviors that they may not otherwise undertake. Research has shown that scripts benefit students from elementary to high school in problem-solving domains and in comprehending text compared to working alone or participating in unstructured collaboration (O'Donnell & King, 1999). In essence, scripting generally places students in roles that encourage each student in a pair or group to contribute substantively to discussion. They can provide students with guidance in how and when to act (e.g. after your partner answers your question, switch roles so that your partner can ask a question), setting up a structure for collaborating that removes the uncertainty of what to do next. However, there are some limitations to scripting which are addressed after reviewing how prompting externally guides interaction.

Prompting interactions. In authentic elementary classroom settings in the U.S., King (1994) found benefits for prompting students (via written cards) to engage in particular cognitive moves during discussion of science topics. She tested the effects of "experience-based" questioning strategies, "lesson-based" questioning strategies, and explaining (without being given explicit question prompts) on student learning of human biological systems. Experience-based question prompts directed students to ask each other to connect new information to something they had already learned. For instance, students would be prompted to ask their partner, "How does … tie in with … that we learned before?" (p. 345). An example of a lesson-based question prompt was, "What are the strengths and weaknesses of …?" (p. 345). The explanation prompts basically asked students to ask each other questions and answer with explanations in a generic way. They provided no other guidance in what kinds of questions to ask or how to ask questions.

Results showed that students who used the specific questioning strategy cards performed better on posttests than students in the less guided explanation group. Furthermore, experience-based question prompts were superior in promoting retention compared to lesson-based question prompts. Qualitative analyses of verbal protocols revealed that the questioning prompts promoted more knowledge assimilation compared to the explanation prompts, and experienced-based questioning prompts showed the highest levels of knowledge construction. Additionally, students from the questioning strategy groups showed evidence of asking more deep, comprehension-type questions, compared to the explanation condition where students asked mostly factually based, clarification-type questions.

Because the experience-based question prompts specifically targeted the students' prior knowledge (by asking students to connect information to what they had previously learned, or to remember what they learned before), this may have tuned-in students to their own knowledge structures in a way that triggered deeper thinking. The explanation prompts were generic and merely asked students to explain the concept. At least for younger populations, more targeted approaches that specifically encourage students to access their prior knowledge may give them more to talk about. Again, as similar to the snowball phenomenon that can occur with analogies (Lin et al., 2012), snowballing might also occur with question-asking or explaining by directly activating one's prior knowledge. In other words, in groups that were prompted to access prior knowledge through the experience-based questioning strategy, the first few occurrences of these question-answer prompts may have kick-started deeper discussion.

A study by Coleman (1998) assessed how prompting elementary school students identified as "average intentional learners" (those who typically used rote-learning approaches) to use explanations during discussion of science concepts centered on photosynthesis could improve learning. The conditions of her study included a prompted average-intention group, a no-prompt high-intention group (students with a problem-solving orientation), and a no-prompt control (average-intention). In this study, students engaged in two separate tasks (creating a concept map and then discussing answers to problem questions) and worked in groups of three. Students in the prompt condition were given domain-specific prompts in each phase and switched roles as readers, writers, and explainers, so that each student participated in each role throughout the intervention. Students in the no-prompt conditions participated in each phase by "naturally" collaborating.

Results showed that students in the prompt condition generated more advanced explanations, produced concept maps that were more scientific, and performed better on posttests compared to their counterparts in the no-prompt control group. In addition, the prompted average-intention students performed similarly to high-intention students on all study tasks. These data also showed, via verbal protocol analyses, that explanation-prompting strategies can trigger students to connect their prior knowledge to new concepts, leading to more complex, sophisticated, and/or sustained discussions, which relates positively to learning.

To explain further, it appears that it was not necessarily the prompting alone that led to better learning. The advanced and deeper explanations could be considered the driving force behind student learning. The benefit of the prompting was that it gave external guidance to students to engage in better explaining behaviors. Again, similar to King's (1994) work, these prompts were specifically targeted to trigger students to access their prior knowledge. Thus, not only did the prompts trigger more effective behaviors (explaining), but also triggered students to activate their knowledge structures more deeply. Thus, it is possible that the combination of these two things attributed to successful learning.

With the increased use of technology in educational settings, much benefit has been found for using computer-based prompts, such as with Intelligent Tutoring Systems (ITS), to guide collaboration. This might occur when two students sit together at a computer (Hausmann, Van de Sande, & VanLehn, 2008), or when students collaborate remotely (i.e. at different computers in a classroom, Walker, Rummel, and Koedinger, 2011, or potentially any different geographic locations). Some promising interventions are testing "adaptive" support technology and its effects on improving student interactions and domain learning. Adaptive support prompts use Artificial Intelligence models to assess student contributions and provide targeted assistance and feedback to students on their errors, misunderstandings, and/or progress (Deiglmayr & Spada, 2010). Two studies assessing adaptive support prompts are described below.

Walker et al. (2011) tested adaptive support using a reciprocal peer tutoring design, where high school students sat at computers in separate areas of a classroom and worked together to solve difficult algebra problems on the same visible interface, with the capability to communicate via chat and work jointly on the problems. Students took turns playing tutor and tutee roles during the intervention, which took place over several class periods. In the treatment condition, a computer agent prompted students in the tutoring role with "targeted" support that guided them to help the tutee when appropriate. For instance, a targeted prompt would direct the tutor to ask the tutee to explain his/her answer or reflect upon an alternative answer, when the tutee provided an incorrect solution. In contrast, students in a fixed support condition received similar guidance in content via the computer agent, but the prompts were not directly adapted to the student answers/solutions. They partly occurred in fixed intervals or were random (came from the same list of hints, but did not reflect student progress or behaviors).

Results from this study showed that adaptive support encouraged effective collaboration during the student interactions compared to non-adaptive support. In particular, students who participated in the adaptive support condition evidenced giving more conceptual help to their partner during their time as tutor. Learning gains were found in both the treatment and control conditions, however, no differences were seen between conditions. Current work by E. Walker (personal communication, Fall, 2011) has shown correlational evidence that adaptive support during collaboration positively relates to learning outcomes. One interpretation of these results may be that the adaptive support prompted students to access their prior knowledge more deeply, since it was directly targeted to the individual's utterances and behaviors. This may have then provided students with more content to discuss, or in other words, provided the opportunity for students to bring more substantive contributions to the collaboration, thus, enriching the discussion.

Karakostas and Demetriadis (2011) did find direct support for domain learning (on the topic Learning Theories) by testing the effects of adaptive support prompts combined with scripted collaboration with computer science undergrads at a European university. In this work, two conditions were set up using the same scripted model, but in the experimental condition, the interaction was augmented with adaptive reminding prompts. The script structured student interaction by asking dyads to discuss and agree upon the answer to a "keyword question," then provide answers to an open-ended "learning question" using a text-based chat tool (called LAMS).[8] The script also designated roles for students as either author or reviewer, and asked students to switch roles upon new sets of questions. The adaptive prompts were designed to detect any missing keywords in students' chat discussions of the initial keyword question by comparing their utterances with a predetermined list of important keywords. When an important keyword did not appear in students' chat discussion, a prompt with brief information about that key concept was presented to the dyad. The adaptive prompts provided to the experimental group did not include any new information, but provided brief snippets of the information from the text that all students studied.

Results showed that the prompted students performed significantly better on domain-based posttests. Thus, computer-mediated prompts are further enhancing collaborative learning interventions by combining them with scripting, adaptive support, and peer tutoring designs. Prompting, scripting, and role-switching may help to set up a collaborative situation where both students are more likely to equally contribute to discussion. Adaptive support presses students to activate and externalize their existing knowledge. The combination of these external forms of guidance may allow students to provide more substantive contributions to a discussion, and engage more deeply in the interaction.

Guiding argumentation. Argumentation is a unique instantiation of guiding-peer-interactions that has the potential to bring students to conceptual change, which is often difficult to achieve. To borrow from Asterhan and Schwarz (2010), argumentation

[8] LAMS (Learning Activity Management System) is an open source authoring tool that can be used to create and manage online collaborative learning activities (www.lamsinternational.com).

may be defined as a series of verbal exchanges between two persons, where their dialogue involves reasoning for the purpose of supporting each person's own ideas or claims, and often includes a goal to persuade. Arguing, in this manner, comprises a set of cognitive behaviors that involve explaining, justifying, finding support through evidence, questioning, and/or challenging. It is important to note that argumentation rarely occurs spontaneously in educational settings, hence the recommendations to script or prompt argumentation behaviors, or directly teach them as mentioned earlier (for a full review, see Asterhan & Schwarz, 2010). In this section, studies that address how arguing relates to peer discussion and learning are reviewed, but I first explain why argumentation strategies can help maximize the benefits of collaboration.

Argumentation may be a particularly useful intervention when the instructional goal is specifically to bring students to conceptual change (Asterhan & Schwarz, 2009). Arguing is more than explaining in that it additionally requires one to justify (not simply clarify) an idea or claim, consider the opposing viewpoint, and convince another of his/her position. Argumentation has been found to improve conceptual knowledge (Schwarz & Linchevski, 2007), in part, because it naturally induces cognitive conflict since it provides students with opportunities to confront contradictory information while engaged in a dialogue. Andriessen et al., (2003) refer to this as a process of "confronting cognitions." In fact, the very nature of argumentation is to confront opposing or conflicting sides of a topic or issue and then attempt to make sense of them through discussion. Thus, arguing can lead to positive outcomes for the toughest kind of learning, that which fundamentally alters a learner's mental models. Interventions that encourage students to argue increase the potential for students to resolve misunderstandings and misconceptions, which can greatly impact conceptual understanding of a topic (Duschl & Osborne, 2002).

Schwarz, Newman, and Biezuner (2000) conducted a study that provided evidence of a learning situation based directly on student misconceptions that triggered arguing and led to improved learning. Their study, conducted in Israel, targeted high school students who had below-average math ability. These students were diagnosed by their misunderstandings of decimal/fractions (called "conceptual bugs"), and then strategically paired with a partner based on those diagnoses. These "conceptual bugs" represented a variety of incorrect problem-solving rules. The types of pairs were: R-W – a "Right" student (used a correct rule) and a "wrong" student (used an incorrect rule); W-W – two "wrong" students who had the same "bug"; and W1-W2 – two "wrong" students who had different "bugs." This can be seen as similar to positioning students in roles as with a script, but the roles are based directly on student prior knowledge. The students' resulting dialogues and learning outcomes were assessed.

Results showed that the W-W pairs performed the worst, which was expected. They would be likely to obtain incorrect answers and would have little reason to argue, since they used the same incorrect strategy, basically agreeing (incorrectly) about how to solve the problem. The researchers hypothesized that the R-W pairs' interactions would benefit the W student, but they found that the R student tended to dominate the discussion, suppressing a "genuine argumentative process," and ultimately leaving the W student in the same place where s/he started (p. 491). The W1-W2 pairs, on the other hand, tended to engage in several argumentative moves throughout their discussion, leading the students to infer new rules for solving the problems. In fact, in every W1-W2 pair, at least one student fixed their conceptual bug, while only a single W student from both other conditions adopted a correct strategy after peer discussion. Thus, students who were positioned in the roles of W1 and W2 had more reasons to argue, evidenced by the increase in their argumentative moves, and this helped them to overcome conceptual misunderstandings.

Another study that placed students in opposing roles, by Van Amelsvoort et al. (2007), tested how students across four high schools in the Netherlands used external representations in conjunction with arguing in a CSCL activity. They obtained an assessment of students' positions on the controversial topic of genetically modified organisms by asking them to individually construct a diagram or a text that represented their position on the topic and to use both arguments and counterarguments to support that position. (They all used the same reading material as a resource.) Students were then placed in pairs that set them in opposition to each other and instructed to collaboratively write a position paper on the topic (over chat and a computer-supported collaborative writing tool). Three conditions were assessed: (a) student generated diagram, where both students used their individually constructed diagrams during the collaborative activity (b) student generated text, where both students used their texts and (c) experimenter generated diagram, where both students received a diagram that was based on their individual texts. Final outcomes were assessed in students' revisions of their original diagrams or texts and through protocol analyses of students' discussion.

Results showed that students in the student-generated-diagram condition in general showed higher quality argumentative discussion and, in particular, evidenced engaging in deeper discussion (measured by how elaborate their arguments were). In addition, this group used more examples and explanations while chatting and used more arguments in their collaborative writing than students in other conditions. Not only did constructing diagrams and then using them during collaboration enhance discussion, but also seemed to deepen learning, as assessed by the students' revised individual products.

One interpretation of this work is that externalizing one's own graphical representation seems to better maximize collaboration above using a pre-created given diagram (even when that diagram is based on a student's own mental model). When an individual constructs a visual depiction externalizing his/her knowledge, s/he may be better able to offer explanations and justifications, which should encourage deeper discussion with a partner. A different interpretation is to say that individuals who created and used their own diagrams were better cognitively prepared to collaborate (this is more thoroughly addressed in a later section).

To test whether argumentation could be elicited without guidance, Veerman, Andriessen, and Kanselaar (2002) conducted three university classroom studies in the Netherlands that assessed how certain instructional designs (varying in guidance) would incite argumentative dialoguing. They examined the relationship between questioning and argumentation in three collaborative environments: (a) small groups of students working face-to-face with a tutor, (b) small groups working without a tutor, and (c) collaborating over a computer-mediated interface.

In the first study, students worked in small groups and used concept-mapping techniques to develop a plan that described, organized, and justified learning goals for an educational technology lesson, and then discussed their plan with a tutor. These tutoring sessions revealed that arguing was rare. As prior work on tutoring has shown (Graesser &

Person, 1994), the tutors asked most of the deep questions, while students mostly asked closed-ended questions for the purpose of clarification. There was limited evidence of question-asking relating to some forms of arguing, but questions aimed at inferring knowledge did not correlate with argumentation. The researchers concluded that small group tutoring did not elicit effective argumentative dialoguing.

In the small group without tutor design, researchers provided minimal instruction to students on critical question-asking and used more of a scripted approach by asking students to defend specific controversial claims and then try to "win the argument." These results showed more promising results in that students produced many more questions and argumentative moves. However, these moves generally operated in isolation. In other words, the moves didn't further along the dialogue, but merely served as reasons to support individual claims, as students were instructed to do. Question-asking did not elicit further argumentation for effective discussion.

In the third study, collaborating over a computer-mediated interface, the researchers then provided students with guidelines about how to engage in effective argumentation and structured the activity by giving some preparation time to discuss arguments and claims prior to engaging in the learning activity. Thus, students discussed the domain content and produced conflicting claims on specific aspects of the lesson. They then collaborated over a computer-mediated chat and diagram tool to complete their assignment. In this study, students evidenced more argumentative dialoguing and open-ended questioning (establishing a significant correlation between the two) compared to the other studies. The environment that was found to be most beneficial to student production of argumentative dialogue was one in which students had time to prepare for the collaborative activity by discussing and laying out arguments beforehand, were guided to construct a joint product, and communicated electronically.

To summarize, although argumentative dialoguing is a method of collaborative learning that has the promise of influencing deep, conceptual learning, it is difficult to elicit in educational settings from the high school to university level, evidenced by research across various parts of the world. It seems to require training for students, much guidance and instruction, or an intensive evaluation of student conflicting abilities, opinions, or knowledge so that they might be strategically placed in small groups or dyads. In other words, argumentation has great potential to maximize collaborative learning, but is difficult to attain without extensive instructional effort. In their review on the usefulness of argumentation to learning, Duschl and Osborne established that "argumentative discourse is possible when conditions are right" (2002, p. 62). To date, these conditions are difficult to create.

Limitations and Challenges of Guiding Peer Interactions

Guiding peer interactions has found positive effects for increasing effective collaborative behaviors, and some work has provided direct evidence for its benefit to domain learning. However, there are limitations and challenges with this intervention, suggesting that it does not fully maximize the benefits of collaborative learning. Some studies have shown that scripting, in particular, can reduce motivation if it constrains discussion too much (Dillenbourg, 2002), and can also hinder the potential for natural creative or flexible problem-solving to occur (Cohen, 1994; Dillenbourg & Tchounikine, 2007). Further challenges with interventions that externally guide peer interaction is that the instructional preparation may be quite intensive. Many of the studies on prompting are front-loaded with preparation such as student and/or teacher training (King, 1994; Palinscar & Brown, 1984), or development of computer-supported systems. In fact, some work has shown that it can take from 100-1000 hours to develop a computer-supported system for every hour of instruction (Anderson and Murray as cited in Muldner, et al., in revision). In addition, there is evidence that once guidance is removed, students no longer engage in targeted behaviors (Webb et al., 2006).

Taking Advantage of Ambiguous or Open-ended Tasks

Another way that the benefits of collaborating might be maximized is by providing opportunities for students to work through ambiguous or open-ended collaborative tasks, which can elicit more effective discussion towards learning (Cohen, 1994). These interventions might use ill-structured problems (meaning there is minimal external support to guide students in complex problem-solving) (Kapur, 2008), engage students in discovery-learning tasks (Bruner, 1961; Dewey, 1916), or require students to induce rules or generalities from real-world cases or contexts (Schwartz & Martin, 2004). This intervention can be differentiated from designs that teach collaborative skills or that externally guide interactions, although some of these studies that examined this intervention do include these elements. The difference with interventions that take advantage of ambiguity is that the learning tasks are specifically characterized as having multiple solution paths or multiple ways to achieve the task goal. They might include some minimal instruction to collaborate or discuss the learning material in a particular way, or might involve some guidance as in positioning students in certain roles, but they focus more on the open-endedness of the instructional task. Studies that emphasize this flexibility during collaborating are described below and the reasons why ambiguous tasks might maximize collaboration are addressed.

Ambiguity, as referring to the quality of being understood in multiple senses or ways, sets up a perception for learners that there is no one right answer or right way. It fosters a sense of exploration above following steps. It alludes to a goal of figuring it out or working through it, above getting it correct. The contrast to this is that a learner sees only one way to reach a solution. To illustrate why this matters in peer discussion, consider the following scenarios as paraphrased from work by Kapur and Bielaczyc (2011) in high school math classes in Singapore.

Scenario 1. Students need to learn the concept of variance in statistics. They are given an explanation of what variance is and the mathematical formula for solving variance; they observe the application of the concept as the teacher works through several data analysis problems at the board, while pointing out misconceptions and common errors in applying the concept; they then collaborate in small groups on more data analysis problems; afterwards, the teacher discusses solutions with the class; the students are given homework with similar problems. This represents a typical direct instruction method, with the added element of refuting misconceptions and common errors, plus an opportunity to engage in active learning by collaborating, and additional practice solving problems later on. Together, this scenario embodies good instructional practices and should produce positive outcomes in student learning. To contrast, consider the next scenario on same concept.

Scenario 2. Students who have not previously learned about variance are asked to solve a data analysis problem in small groups. Their task is to create a quantitative index of the most consistent soccer player, given a distribution of goals scored each year by three soccer players over 20 years. As students work on the task, they receive no scaffolding or instructional support; afterwards, they receive teacher feedback comparing and contrasting their solutions; they then are given direct instruction on variance, with explanations of the concept; they are not provided with homework practice problems. This also exhibits elements of good instructional practice, as with active learning via collaboration, subsequent feedback from the teacher, and direct instruction on the topic. However, students in this scenario must initially deal with ambiguity. In this scenario, since no one right way is provided at the beginning of the learning task, students are free to explore solutions, discover new rules or principles, and be more flexible during the discussion process to figure out the path to the answer.

To give a common interpretation, it can be easy to see how simply following a set of given solutions steps and explanations of a concept (as with Scenario 1) might lead to more constrained discussion, characterized by phrases like, "We are supposed to do it this way… the teacher showed it like this… you do this problem, I'll do that one… follow the steps on the board." Whereas having to figure out the path may lead to more open-ended discussion such as, "Let's try this… my idea was this… what if we do this?… why would you do it that way?" The lack of structure present in solving complex, ill-structured problems can encourage exploratory behavior and flexible discussion.

In Kapur and Bielaczyc's (2011) examination of these two instructional conditions, positive learning outcomes were seen in both, yet a significant difference was found between them with regard to conceptual knowledge. Despite the fact that students who engaged in the ambiguous task (Scenario 2) did not do homework practice problems and solved fewer problems during the learning task, they still performed better on conceptual questions compared to students of Scenario 1. Having the flexibility to discover the answer/solution positively related to learning of conceptual material. Thus, dealing with this ambiguity may provide a natural way for students to engage in meaningful dialogue that leads to deeper thinking and improved conceptual understanding.

Student dialoguing patterns have, in fact, been found to differ when students basically follow a set of instructions, such as in working with well-structured problems, compared to having to figure out how to solve problems that are ill-structured. Work on "productive failure" has shown that solving ill-structured problems leads to students to produce more complex dialogue sequences, involving feedback loops from solution evaluation to problem analysis and critique, while solving well-structured problems produces simpler discussions, typically representing solution development followed by "un-sustained" evaluation (Kapur, 2008). One critical point to mention is that students in ill-structured groups appear to "fail" compared to well-structured groups at first glance. Kapur found that discussions between students in ill-structured groups were not only complex, but appeared "chaotic" and "all over the place," and their solutions during the learning task were assessed by experts as inferior, compared to students in the well-structured condition (2008, p. 403). In earlier work, Kapur and Kinzer (2007) concluded that collaboratively solving ill-structured problems lowers the quality of discussion and group performance compared to solving well-structured problems, but this was based on measures of performance during the interaction, rather than on later posttests assessing retention and transfer. More current work shows that after the collaboration task is finished, students in ill-structured conditions outperform those in well-structured conditions on individual posttests that have both ill- and well-structured problems (Kapur, 2008).

Thus, the messy dialoguing that results from an ill-structured design seems to lead to more flexible learning and transfer of knowledge. Discussing through solutions/answers while dealing with ambiguity may look chaotic and "unproductive" initially, but there seem to be learning benefits that surface in later activities. Sometimes, the presence of too much structure, too clear a path to the answer, or the sense of "one-right-way," may restrict peer discussion in ways that fail to maximize the benefits of collaborating.

Engle and Conant (2002) have also addressed the point of ambiguity as beneficial to collaborating. In their work on identifying principles that foster "productive disciplinary engagement," they discuss how "problematizing content" can set up opportunities for students to think more deeply about the topics that they are learning. This principle encourages students to question, propose, and challenge information, rather than assimilate facts and "expert" answers. Problematizing has to do with inspiring curiosities to make sense of information. The researchers discuss the problem with setting up the perception for students that there is only one right way:

> …when learning environments communicate to students that there is a single valid response to every question and that students' job is merely to determine what it is… students may become highly engaged in getting possession of the right answer and having it validated by an authority,' thus, 'short-circuiting' productive [collaborative] engagement (pp. 408-409).

In other words, the setup of searching for "the" correct answer is different than the setup for discovering "a" correct answer. Searching-for is an active, hands-on type of learning activity that provides some benefits to students since they must do more than passively receive information (like hearing a lecture). Discovering, on the other hand, allows students to generate possibilities, make hypotheses or predictions, test out their ideas, and otherwise construct solution paths or avenues to achieve the learning goal, which presses students to cognitively engage more deeply in the learning material.

Engle and Conant's (2002) qualitative analyses portrayed an argument that formed over several class periods between elementary students who were completing a science group project in a Jigsaw-type fashion. (Although Jigsaw is typically considered highly structured, in this particular activity, the students had the freedom and were encouraged to engage in open-ended discussion.) For this argument that developed, there was actually a correct answer that solved the students' disagreement, but the teacher left it up to them to make their own decision based on the information the group had gathered. As a result, students engaged in a passionate discussion resembling many elements of argumentative dialogue, using evidence to support their claims and attempting to persuade the other to his/her side. They also developed more elaborate and sophisticated arguments throughout their discussion. In addition, prior to reaching a group consensus on the issue, students engaged in comparing and contrasting multiple sources of information and asking deep, conceptual questions about the topic. Leaving the solution path ambiguous fostered collaboration that highly engaged students in learning.

Another way in which ambiguity can be taken advantage of during peer discussion is by having students work with visual/graphical representations, which leave more room for interpretation compared to text-based representations, which are often linear and laid out in a sequential manner (Janssen, Erkens, et al., 2010). Visual representations tend to give a holistic view of a concept and its related subparts, allowing students to see multiple connections between several aspects of the concept and providing some freedom for students to determine how and possibly why they are related. Compared to text-based representations, graphical ones may be more open-ended in how students understand them, leaving multiple pathways open for making sense of information.

A study by Janssen, Erkens, et al. (2010) conducted in the Netherlands assessed how high school students in five history classes collaboratively recreated a historical debate with a graphical- versus a text-based computer-supported tool. They found that students who used the graphical tool had more complex interactions. Students displayed more instances of shared understanding, yet rated the social aspect of the task more negatively than students who collaborated with the text tool. The students who use the graphical tool had messier dialogues, but co-wrote better argumentative essays, co-constructed higher quality external representations, and scored higher on domain-based multiple-choice posttests. This messy dialoguing serves as another example of a productive failure (Kapur, 2008; Pathak, Kim, Jacobson, & Zhang, 2011). Although dealing with the ambiguous nature of the graphical tool (as non-linear and non-sequential) was less straightforward to discuss, it offered multiple ways to understand the debate, affording deeper engagement in the task and deeper discussion, and led to better post-intervention outcomes. Similar affordances have been found in collaboratively creating concept maps compared to collaborative-writing (Haugwitz et al., 2010).

One more study to mention regarding the open-endedness of graphical representations, also conducted in the Netherlands, is by Van Boxtel, Van Der Linden, and Kanselaar (2000b). They compared high school students' discussions in two conditions: (a) collaboratively creating a concept map with no additional resources (the students had just received a class lesson on the topic) and (b) collaboratively creating a concept map with two supplemental textbook chapters. Two opposing hypotheses were presented about the effect of the texts on collaboration:

- The texts should enhance collaboration since students can use them to as a resource when they are unsure how to proceed and this may support communication and negotiation activities.
- The texts will constrain student discussion since students see them as an authority and may be more likely to consult the textbook, rather than engage in discussion to reach consensus.

Results showed that students learned significantly in both conditions (assessed by pre- to posttest measures), but there was no different between conditions. In the no-textbook condition there was a positive correlation between the number of domain-related statements made during discussion and posttest scores, but in the condition with the textbooks there was no such correlation. Additionally, pairs in the no-text condition elaborated more, had more conflicts, and evidenced better reasoning throughout their discussions. Students who had access to the texts generally had limited discussion, and consulted the texts to find answers more often than using them to enhance their discussions. The researchers concluded that students who had the texts mainly used them to find answers, and seemed to do "less thinking for themselves" (p. 71), thus, constraining discussion. In other words, the availability of the texts removed some of the opportunity for students to take advantage of the open-endedness of the task. They, instead, reinforced a right way/right answer approach to the task, actually hindering meaningful discussion.

This is similar to the findings by Hausmann et al. (2009) and Hausmann, Van de Sande, Van de Sande, and VanLehn (2008), showing that when students have access to the correct solutions or answers (either through a tutoring tool or a textbook), they will tend to search for answers, rather than think through them on their own. Hausmann and colleagues (2008, 2009) found that in comparison to a solo group, it did seem that the opportunity to engage in discussion with a partner (jointly-explaining) encouraged students to try to figure out the answers themselves. Van Boxtel et al. (2000b) showed through their two collaborative conditions, that having access to the answers limited discussion compared to being "forced" to work through them.

In summary, giving students the opportunity to deal with ambiguity sets up a learning situation that encourages exploration above obtaining the right answer. With regard to discussion, it sets up an environment that focuses on what could be done, rather than what should be done. "What should" represents an inflexible learning situation, likely to constrain discussion towards obtaining the right answer, rather than being open-ended. On the other hand, "what could" opens up a flexible discussion space, where students can contribute a range of ideas towards figuring out a way that works. These types of open-ended learning environments seem to elicit effective dialoguing that improves learning.

Challenges with Open-ended/Ambiguous Tasks

Although work has shown that making learning tasks more open-ended by offering opportunities for students to work through ambiguous situations (such as working with concept maps or graphs or tackling ill-structured problems) leads to discussion that improves learning, there are still challenges that must be overcome. Considering student prior knowledge is of particular concern here. If students have limited to no prior knowledge in the domain, they may be far less likely to engage in meaningful discussion, especially if it is open-ended (Nokes-Malach, Meade, & Morrow, 2012; Wiedmann, Leach, Rummel, & Wiley, 2012). Without an adequate base level of existing knowledge related to a concept to-be-learned, there may be virtually nothing for students to talk about. My work addresses this issue by first cognitively preparing students (which may activate or improve their existing knowledge and more readily bring them to a state of cognitive conflict) prior to engaging them in a collaborative task.

Summarizing Taking Advantage of Collaborative Opportunities

Three major instructional approaches that have been investigated in a variety of classroom and laboratory studies have attempted to maximize the benefits of peer collaboration: (a) teaching collaboration skills beforehand, (b) externally guiding student discussion through scripts and prompts, and (c) using ambiguous and open-ended tasks to provide students a reason to talk through ideas or solutions. These approaches have shown success in cognitively engaging students in discussion, but also carry challenges and limitations. For both teaching skills and externally guiding interactions, skills are often lost after time or are no longer utilized once external guides are removed. Longer intervention periods or fading approaches may improve these strategies. For enhancing collaboration in open-ended tasks, assuring that students' domain knowledge is adequately prepared for discussion might better assure that students will collaborate to their potential. Using the lessons learned from each of these interventions can lead to the development of a better model for collaborative learning. The "preparing-to-interact" method avoids time spent on training students in collaboration skills and the cost and effort of developing scripts and prompts, and helps to assure that students' existing knowledge structures are activated in ways that spur on meaningful discussion. Thus, the next section addresses the role of cognitive preparation in collaboration using existing work as indirect support of a "preparing-to-interact" phenomenon.

The Role of Cognitive Preparation in Collaborative Activities

We know that collaborating during learning activities has the potential to boost student learning above working alone. We also know that students are not always effective collaborators, and therefore, miss out on learning opportunities that collaborative activities afford. Interventions such as teaching students collaboration skills, externally guiding and scaffolding their interactions, and designing open-ended collaborative tasks have all found some success in attempting to maximize the benefits of collaboration; however, they each carry their own challenges and limitations. The "preparing-to-interact" method of structuring collaborative activity is worthy of investigation, since it can avoid some of the existing challenges and limitations of other interventions, and it considers students' readiness for engaging in discussion to learn.

The "preparing" of this method refers to a cognitive preparation, where students' existing knowledge structures are activated by their engagement in a learning task prior to participating in a discussion with a peer. To provide theoretical support for the potential effectiveness of such a method, two cognitive approaches to learning are first described: the Interactive-Constructive-Active-Passive (ICAP) framework and hypothesis, and the Preparation for Future Learning (PFL) paradigm. Then, studies that used preparation phases in their designs of collaborative learning interventions are reviewed. Some of this work has already been described, however, a few studies are reinterpreted under a "preparing-to-interact" perspective.

The PFL paradigm is essentially a dual-phase learning model that first provides students with a highly engaging cognitive activity (typically an open-ended task requiring students to invent, induce or discover rules, patterns, or principles within situated contexts or applied problem-solving), and then follows this activity with a lecture-style presentation. The origin of the PFL model came from Schwartz and Bransford (1998), who found that certain kinds of preparation activities enhanced student learning from lecture. Currently, this model is generally limited to preparation to learn from lecture, but it might also be applied to preparation to learn from peer discussion. Before the PFL model and its potential application to preparing to collaborate are explained in more detail, I elaborate on the ICAP framework. ICAP is a single-phase model that can be used to predict learning outcomes from the way in which students engage in a task. This framework can be used to inform the types of tasks that might better prepare students to learn from collaborating.

The Interactive-Constructive-Active-Passive (ICAP) Framework

ICAP differentiates student engagement in learning tasks by categorizing overt student behaviors as Interactive, Constructive, Active or Passive, and is founded on theoretical assumptions about how those behaviors relate to different cognitive processes

(Chi, 2009; Fonseca & Chi, 2011). For instance, in an Interactive task, students may be discussing how to solve a problem that has various solution paths, asking each other questions and explaining their answers, or trying to reach consensus on a controversial issue through argumentation. Interactive tasks allow students to participate in dialogues, which have the potential to engage students in the process of co-creating knowledge. In a Constructive task, students may engage in discovery-learning or invention activities, generate inferences, use analogies, or draw/interpret graphs. Constructive tasks allow students to create new knowledge. In an Active task, students might follow step-by-step instructions, copy the solution steps to a problem, or underline/highlight written information. Active tasks allow students to physically engage in the learning material and link it to existing knowledge. In a Passive task, students might listen to a lecture, watch a presentation, or read from a textbook. Passive tasks merely present information to students, without necessarily triggering students to activate prior knowledge.

Based on the kinds of cognitive engagement that certain tasks are likely to elicit, ICAP provides a hypothesis for learning outcomes by categorization of the learning task: Interactive tasks produce better learning outcomes than Constructive tasks, which are better than Active tasks, and these are all better than Passive tasks, I > C > A > P. By better learning outcomes, I am referring specifically to evidence of deeper understanding. Chi (2009) reviewed prior literature that supports several pairwise comparisons based on this hypothesis. Recent work has found direct evidence for the complete hypothesis in both the laboratory and classroom (Menekse et al., in press; Menekse, 2012), which confirms the value of using the ICAP framework to design activities that are founded on different levels cognitive engagement. The table below summarizes the ICAP framework and hypothesis.

Table 1
Summarizing ICAP

	Interactive >	Constructive >	Active >	Passive
Predictions	activities produce better outcomes than…	activities, which produce better outcomes than…	activities, which produce better outcomes than…	activities.
Behaviors	Dialoguing	Generating	Selecting or emphasizing	Receiving
Cognitive Processes	Co-creating knowledge	Creating new knowledge	Assimilating knowledge	Storing knowledge
Examples	Arguing/debating; explaining to another; elaborating on a partner's ideas	Self-explaining; creating a concept map; inventing a rule; discovering a pattern	Underlining/ highlighting text; copying solution steps from the board; repeating verbatim	Listening to a lecture; reading silently; Watching a teacher

Note: Modified from Chi's (2009) work

One practical application of ICAP for instructional design is to modify existing learning activities by "bumping them up" to a higher engagement level. For instance, if the goal for students is to develop deep understanding of a concept, giving them a lecture or assigning textbook readings (both Passive activities) may not be enough. Reading a textbook chapter can be bumped up to Actively engage students by having them underline sentences or phrases that signify main ideas. This encourages student to think one step further in order to choose those sentences/phrases. This textbook activity could also be bumped up to Constructively engage students by having them create a concept map that links main ideas together along with related details or examples. This encourages students to think more deeply in order to assess the relations between parts of the concept, and represent them in a coherent structure. In this case, they must generate inferences, relations, and new ideas to make sense of information to achieve the task. (See Chi, 2009, and Menekse et al., in press, for extended explanations.) To make this activity Interactive, students could work in partners to create a concept map, with the hope that they would deeply discuss their ideas with each other.

I purposely state, “the hope” that students will deeply discuss ideas, because the collaborative learning literature is in general agreement that simply asking students to “work together” on an activity is not always effective, hence the various interventions that researchers have recommended to better assure effective collaboration. Thus, modifying a task to make it Interactive must do more than ask students to work together. Consideration must be taken to better assure that students interact in a way that leads to learning.

To expand on the example of bumping up the task of reading a textbook chapter, students could be asked to Interactively engage by discussing in pairs which parts of the text represent main ideas before underlining them. However, underlining is still essentially a selection activity and does not necessarily press students to generate any new knowledge beyond what is in front of them. Collaborative activities that are centered on Active tasks (like underlining) may not encourage students to engage in meaningful discussion as much as activities that push students to think more deeply, (i.e. to be Constructive). Some work has shown evidence of this (Kapur, 2008; Kapur & Bielaczyc, 2012, on comparing well-structured problem-solving (Active) to ill-structured problem-solving (Constructive); Schwartz & Martin, 2004, on comparing “tell and practice” strategies (Active) with invention strategies (Constructive)), however, no work has directly compared learning from collaborating in tasks that are manipulated by ICAP differentiations. The ways in which students interact in a collaborative learning activity may differ, depending on whether the tasks are requiring students to use the materials in a non-generative (Active) or generative (Constructive) way. Chi (2009) has theorized that to truly categorize engagement as Interactive, students must at least be constructive, but existing work has not yet examined learning tasks in such a way to test this.

Concept-mapping is a Constructive task and existing work supports it as an activity well suited for collaboration (Van Amelsvoort et al., 2007; Janssen, Erkens, et al., 2010). There are several ways to make a concept-mapping activity (or a variety of other Constructive tasks) Interactive, however, two general ways are highlighted here: (a) students can jointly create a concept-map, under the assumption that the open-endedness of the task will spur on deep and meaningful discussion, or (b) students can first create individual maps and then discuss their work with a partner or in a small group. In the latter case, students then have a frame of reference for comparing, contrasting, or evaluating their own and/or their partners' work and this might provide additional opportunities for students to engage more meaningfully in discussion, and ultimately enhance learning. To date, work on collaborative learning has not compared the learning outcomes resulting from these two ways to set up Interactive engagement activities based on the ICAP framework (with a notable exception by Van Boxtel et al. (2000a) to be described in a later section). The latter mentioned Interactive activity represents a kind of "preparing-to-interact" design, which is investigated in my work. Before some the studies that indirectly support this design are reviewed, the PFL paradigm is explained in more detail to provide the theoretical basis for why deep cognitive engagement in a preparation task is beneficial for learning in a future task.

The Preparation for Future Learning (PFL) Paradigm

The PFL paradigm considers the role of prior knowledge in light of early learning experiences and how those experiences can help or hinder future learning, and as schemas that can be activated appropriately (or not) for learning a new concept (Schwartz et al., 2007). The classic work that introduced PFL tested how students could be prepared to learn from classroom lectures. Despite a general consensus that active learning techniques are better than passive techniques (Bonwell & Eison, 1991), which lecture traditionally falls under, Schwartz and Bransford (1998) argued that there is still a place for lectures in the classroom. They found that a particular combination and order of learning activities (e.g. reading, receiving lecture, summarizing, and contrasting cases) could produce the best outcomes, namely, first contrasting cases, and then receiving a lecture. The contrasting cases activity required students to decipher on their own, the distinctions between two cases demonstrating different psychological phenomena. The researchers concluded that this activity tapped into student knowledge structures in a way that made them "ready" for the lecture. It created a "time for telling," suggesting that certain activities can positively impact students' readiness to learn from subsequent instruction.

Work that has tested the PFL model has shown that tasks that push students to make discoveries, invent rules, and figure out complex problems better prepare students to learn from a lecture (Belenky & Nokes, 2009; Kapur & Bielaczyc, 2011; Schwartz & Bransford, 1998; Schwartz & Martin, 2004; Schwartz, et al., 2007). One reason that these kinds of preparation tasks work well is because they have the potential to activate students' knowledge structures at a deep level. According to ICAP, these kinds of tasks can be categorized as Constructive, and one thing Constructive activities do is allow students to access their own prior knowledge deeply, making it more likely for students to recognize where their own misunderstanding lies. (Again, by "deeply," I am referring to a structural or higher principled mental model level, versus a superficial or "surface-feature" level, (Chi & VanLehn, 2012).)

As a result of the deep activation from engaging in a Constructive activity, students may be left with mental models that are more vulnerable to change. By leaving mental models "vulnerable," I mean the awareness that something is missing or incorrect in one's own thinking. It is, in a sense, the state of experiencing cognitive conflict or cognitive uncertainty. This is important to consider because students often (incorrectly) assimilate new information into existing mental models when, in fact, the structure of the models need to change (Chi et al., 2012). Assimilation is a path-of-least-resistance strategy, while accommodation (true mental model change) can require much cognitive work. Engagement in a Constructive activity, such as an invention task, a discovery task, a task forcing a person to reason-through, evaluate, synthesize, etc., might press a learner to be more ready to restructure, recreate, or completely change their mental models, rather than inaccurately assimilate new information. Thus, when students are in such a state of uncertainty or vulnerability with regard to their understanding, the chances for true accommodation to occur in future learning (as in a lecture) may increase.

Schwartz et al. (2007) discussed the idea of incommensurable pieces of knowledge, which are pieces of knowledge that cannot be reconciled because they do not “fit” into students’ existing schemas. In such cases, students are often left with incorrect or incomplete understanding without necessarily realizing it. (These pieces tend to stay isolated, as episodic forms of knowledge, or are simply linked in the “wrong” place.) However, a Constructive preparation task seems to discourage this incorrect fitting and helps students to reach cognitive conflict. The follow up task (such as the lecture, or possibly discussion as with collaboration, as my work examines) then serves as a way for students to restructure their thinking and understanding, since they are more ready to receive new information. Although they may not be able to fix misunderstandings during the preparation activity, they can at least become aware of them. This awareness may act as a catalyst that sets up students to better engage in future activities. Students may be more eager to receive new information, may want to avoid the feeling of disequilibrium, or may be more motivated to find solutions to problems.

Studies Supporting a “Preparing-to-Interact” Design

Some empirical studies that provide support for a few different ways that students can be prepared to better collaborate and engage in more meaningful discussion are reviewed here. Firstly, Froyd’s (2011) theoretical paper reinterpreted the PFL paradigm within a framework for preparing students for collaborating. His theoretical assumptions placed preparation as the process of activating knowledge structures, then placed collaboration as the way to hone those structures. I borrow from his connection of preparing-to-collaborate with Schwartz and Bransford’s (1998) preparing-to-receive-lecture to reinterpret collaborative learning studies that included individual preparation prior to a collaborative task.

Reinterpretations. Van Amelsvoort et al.'s (2007) study is briefly re-summarized using a "preparing-to-interact" perspective, where students first constructed diagrams or texts in a preparation phase prior to discussing a controversial issue in science. Students then used the products that they created for discussion in a generation-of-diagram versus a generation-of-text condition. The researchers included an additional condition where students' texts were converted (by the researchers) into a diagram, to see how discussion would be affected by diagrams that were not constructed by the students themselves.

What they found was that preparation by constructing a diagram produced better outcomes overall than constructing a text. However, with regard to "deep" discussion (they also assessed "breadth" of discussion), the difference was in comparisons to the (experimenter) given-diagram conditions. The given diagram condition did not produce as deep discussion as the other two conditions, which may lie in the fact that those diagrams were not directly generated by the students themselves. This provides further evidence about what constitutes a sufficient preparation task for future discussion. The best scenario might be for students to engage in a Constructive task prior to the interaction, and then use their own products of the preparation task during the interaction. Students' own created work may give them a better frame of reference for discussion, since they are direct external representations of students' knowledge structures.

Coleman's (1998) work on explanation prompts also included a preparation phase prior to student collaboration. She performed some analyses of how first preparing students for discussion with a concept-mapping activity affected later outcomes by assessing maps that were created individually, prior to collaborative work, and comparing them to joint-maps created during collaboration. These maps were used as indicators of scientific reasoning and domain learning. In this case, the manipulation the intervention task was in the use of explanation prompts students received and not necessarily in the type of preparation in which they engaged. Therefore, in that sense, all subjects individually prepared for future learning in a collaborative task.

These results showed that, in general, students improved their maps to a significant degree, in both prompted and no-prompt conditions. This provides support to conclude, firstly, that the Interactive task of jointly-creating a concept map produced better outcomes than the Constructive task of individually creating a map. A couple of interpretations of these results may be that students improved in the joint-task simply because they had a second chance to do it (i.e. the individual task provided practice), or that students who were less competent gained from highly competent partners and so the joint-product may have been an inflated measure of learning gains for those less apt students. However, I speculate that the reason why the Interactive task produced better outcomes was because the initial individual concept-mapping task served as a form of cognitive preparation that allowed students to engage meaningfully in discussion, which is supported by Coleman's verbal protocol analyses. As a "preparing-to-interact" method, this work provides some clues that students were able to restructure their mental models to reveal more accurate understanding after collaborating on an activity for which they cognitively prepared.

Instances during dialoguing. Another way that studies on collaborative learning might be reframed under a "preparing-to-interact" perspective is in the smaller instances of dialogue. Chin and Osborne (2010a, 2010b) conducted a qualitative study on the relationship between questioning and argumentation, concluding that the act of questioning prepares students to more effectively use argumentation strategies in discussion. This work was conducted with middle to high school students in the U.S. and Singapore on science topics. I should note that this study used quite an extensive intervention, including training in argumentation and questioning, developing self-generated questions in a preparation phase, and placing questions into a concept-map ("Question Web"). Thus, it may have been the entirety of this "preparation phase" that led to good arguing. However, their focus was specific to how questioning and self-questioning prepared students for arguing compared to the "norm," evidenced in protocol analyses on the roles questions played in shaping the students' dialogue. They found that questions (2010a, p. 893):

- "pushed students to be aware of and to articulate their puzzlement." Thus, questions prepared students by activating students' mental models and triggering a state of cognitive conflict. The articulation of "puzzlement" relates back to the idea that when knowledge structures are left in a vulnerable state, students may be more ready for a subsequent discussion in order to refine those structures.
- "prompted students to make explicit their beliefs, claims and (mis)conceptions." Questions prepared students by allowing them to externalize their knowledge structures for both themselves and their partners, improving the likelihood for deep and meaningful discussion.

Thus, according to Chin and Osborne (2010a), questioning can serve as a preparation tool that deeply activates knowledge structures, while the ensuing discussion can serve as a way to restructure knowledge through argumentation strategies. This was evidenced in the students' frequent behaviors of "challenging opposing viewpoints, critically evaluating ideas, and considering alternative propositions" (p. 893). The authors reasoned, "…asking questions about the phenomenon at hand is one means of establishing the nature and extent of students' domain-specific content knowledge prior to asking them to engage in argumentation" (p. 902). Recognizing students' domain-specific prior knowledge is not only helpful for teachers that are assessing student learning, but may helpful to the students themselves, helping them to be better aware of their own lack of understanding and making them more ready for subsequent learning.

Another type of collaborative activity that falls under a "preparing-to-interact" design can be seen in classrooms that have adopted Clicker technologies. Clickers are small remotes that students can use to anonymously answer questions that a teacher poses during lecture, and then their answers are immediately analyzed to indicate the overall results of their responses. Clickers are mainly used to enhance student engagement in lectures, however, one recommended strategy for using them is to have students answer a question individually, then take a few moments to discuss their answer with a classmate.

One study that provided evidence that individually responding to Clicker questions and then engaging in peer discussion could improve understanding of difficult concepts was conducted by Smith et al. (2009). They assessed student responses to questions prior to discussion with a partner, after discussion, and then to subsequent isomorphic questions (those that addressed the same principles, but used different cover stories) by using Clickers in a university lecture course on genetics. The results indicated that most students learned from their discussion of the first questions, and were also able to apply their knowledge in the isomorphic questions. Analyses also showed that it was not simply that less knowledgeable students were partnered with those who knew the answers. In many cases, both students could not answer the first question individually, thus, the authors concluded that the peer discussion directly affected student learning in a positive way. This provides further support for the benefit of using a “preparing-to-interact” strategy for collaboration, even in smaller instances of dialogue.

A study examining a “preparing-to-interact” method. A study by Van Boxtel et al. (2000a) provides strong support for the value of a “preparing-to-interact” method of collaborative learning. They essentially compared how two activities (concept-mapping and creating a poster) that high school students completed individually prepared them for subsequent discussion. This work was conducted in two physics classes in the Netherlands.

Their results showed that individually creating a map or poster and then collaborating in dyads to jointly recreate the item led to higher quality discussions than having students jointly-engage in the tasks without first preparing (but were given extra time during their collaborative task to equalize time-on-task). The dyads who prepared before collaborating spent the first part of the collaborative activity mainly discussing their individual designs, and further referred to their individual designs to support proposals, confirmations, and criticisms during discussion. Thus, the preparation provided an extra tool for accomplishing the joint task. Students who prepared also scored significantly higher on conceptual knowledge items of the posttest compared to students who did not prepare.

This work also provides insight regarding how the type of engagement in a task affects peer discussion. ICAP can be used to categorize the concept-mapping activity as Constructive, while the poster activity can be categorized as merely Active. The authors proposed that the concept-mapping activity should have required students to think abstractly, by considering the subparts of the concept-at-hand more holistically and deciphering the relations between those parts. The poster, on the other hand, was hypothesized to encourage concrete thinking and to serve more as a way to describe the concept, without necessarily requiring generation of new knowledge in the way that creating a concept map would.

Protocol analyses showed that dyads who engaged in the concept-mapping activity did, in fact, talk more about the relations between concepts, discussed abstractions (such as using formulas), and talked more "intensely." They evidenced using more elaborated conflicts and more frequently constructed reasons for their ideas compared to dyads who created a poster. In addition, the poster activity seemed to somewhat deter students from talking, protocol data showing that students in this condition engaged in more writing or drawing. Despite the differences these two activities had on student discussion, measures obtained via posttests were less conclusive about the effect of type of preparation task on learning. The researchers did find correlational evidence that certain aspects of interaction were positively related to learning, and that it was typically the students who created concept maps that engaged in dialogic moves that enhanced learning. Thus, my study extends Van Boxtel et al.'s (2000a) work by focusing more on the learning outcomes of a "preparing-to-interact" design, using assessments that measure deep and shallow learning. In particular, my interest is in the interaction between preparation and type-of-task within collaborative learning settings and how this affects learning.

Summarizing the Role of Cognitive Preparation

Upon review of this work, one important question to address is: Does learning within a “preparing-to-interact” design take place during the preparation, during the collaboration, or both? On one hand, there are certainly some learning benefits to collaborating without any kind of preparation (as mentioned in the work on comparing collaborative activities to working alone). On the other hand, work on preparing-to-teach speaks to the idea that much of the learning in such designs takes place during the preparation. I do not cover that literature here because I take a different angle. I see a unique benefit that preparation has on learning from collaborating. In referring back to the PFL paradigm, one might argue that some learning does take place in the preparation, but that understanding is left incomplete. Thus, the goal of the preparation task is not necessarily a learning goal, but a “readiness” goal. The future learning task (in most PFL work, a lecture or some didactic form of instruction) can then complete a students’ understanding, which positions it as the learning goal.

In light of the PFL model, when peer interaction follows a preparation task, there are chances that a number of things can happen: (a) students can share uncertainties with peers and may be more capable of externalizing their internal knowledge because they are better aware of it, (b) students can then better question each other, and (c) students may be more prone to engage in effective collaborative behaviors as these things may snowball. To say it differently, preparation may arm students with more to talk about, thus, providing more opportunities for students to experience cognitive conflict or uncertainty, and then further motivating students to resolve conflict through discussion, and so on.

The idea of "preparing-to-interact" is certainly not entirely new to the area of instructional design. The think-pair-share strategy, developed in the 1980s (Lyman, 1981; McTighe & Lyman, 1988), was founded on the idea that having students first think individually about an idea/concept in response to a teacher's question will better prepare them to discuss the concept with a partner (pair). Think-pair-share or questioning-then-discussing activities support the basic idea that cognitive preparation can improve subsequent discussion. I propose that we take fuller advantage of these activities by extending the "think" part into a preparation period of domain-based cognitive activity, and then extend the "pair" part into a subsequent period of peer discussion.

An extended period for preparation to think, and in particular, to allow students to activate their existing knowledge structures in a deep way that may induce cognitive conflict and/or an awareness of incomplete understanding, may provide a fertile ground for peer discussion. ICAP can be used to design preparation tasks that constructively engage students. The subsequent period of peer discussion may then boost learning because of the input that becomes available from a back-and-forth dialogue. Opportunities for questioning, hearing new perspectives, arguing, elaborating, and reflecting (all beneficial learning behaviors) may occur more naturally when students are more ready to interact. No work has yet attempted to develop preparation activities using the ICAP framework, and no work has specifically used the PFL paradigm to test the effect of preparing students to learn in future collaboration.

Finally, training students in collaborative skills that focus on cognitive behaviors can certainly be considered a form of preparation, and thus, also a type of “preparing-to-interact” design. However, studies on training or instruction to collaborate place the focus on the skill rather than on students’ own knowledge. My interest is in how preparing students by activating prior knowledge in an intentionally deeply engaging way may then translate to effective interactive behaviors and consequently affect learning. This kind of domain-based cognitive preparation may trigger students to more readily recognize where their understanding falls short, and leave their mental models more vulnerable or susceptible to change. When meaningful discussion ensues, this may better help students to restructure their thinking and result in enhanced learning. A “preparing-to-interact” method of collaborative learning has the potential to maximize the benefits of peer discussion and lead to deep learning in a domain.

Chapter 3

METHODS

This study tested an alternative method of collaborative learning, namely, how individual cognitive domain-based preparation affects learning from engaging in collaborative tasks. The Interactive-Constructive-Active-Passive (ICAP) framework and hypothesis and the Preparation for Future Learning (PFL) paradigm provide theoretical support that "preparing-to-interact" may be an effective way to structure collaborative learning activities towards maximizing student outcomes. Specifically, my study assessed the effects of two types of preparation on collaborative learning, measured via pre- and posttest in the domain of psychology. Assessments differentiated shallow from deep learning. In addition to type of preparation, results were compared against conditions where students worked collaboratively for the duration of the learning activity without first "preparing" through an individual task.

Research Questions

1. Does individual cognitive preparation in a specific domain topic prior to engaging in collaboration have an effect on learning outcomes after collaborating?
2. How does the type of task in which individuals prepare prior to collaborating affect learning outcomes after collaborating?
3. How does the type of task in which individuals engage while collaborating affect learning outcomes?
4. As related to the three questions stated above, how is deep learning affected differently than shallow learning?

Research Design, Domain, and Sample

Research Design

This study used a 2x2 experimental design examining the factors of Preparation (No Prep and Prep) and Type of Activity (Active and Constructive). Measures of both deep and shallow learning outcomes were obtained via pre and post domain knowledge tests, and domain-based activity worksheets. Basic demographic information, ratings of students' feelings about the activities, and students' ratings of the collaborative experience with their partners were also collected.

This experiment was run as a classroom study in several community college Psych 101 classes and, as opposed to a laboratory study, there were unique practical challenges that needed to be addressed. First, the time that students were given to complete all research activities was restricted to the specific class times. Thus, in some cases, students turned in their work before they finished and some of the analyses were conducted at reduced sample sizes due to the incompletion of materials. This is further explained in the Results chapter. Second, there was a challenge in balancing ecological and internal validity with regard to the assignment of experimental conditions. It was not feasible to run the No Prep and Prep conditions within a single classroom, because students would then have knowledge that only some students were being provided with individual work prior to collaborating. Such a setup could have confounded results because student behavior might be influenced by such knowledge. Therefore, students within a class were first randomly separated into two groups (i.e. No Prep and Prep) and each group worked in a separate classroom. Then, within those groups, students were randomly assigned to an Active or Constructive condition. Ecological validity was somewhat compromised since it was not typical for students to work in a different classroom, however, this allowed interval validity to be preserved as all four conditions could then take place within each participating class. (Figure 1 on page 91 illustrates the study design.)

Domain

The specific topic being assessed within the domain of psychology was "concepts of memory." Prior research attests to the difficulty that students have in deeply understanding the differences between various concepts of memory, in particular, for encoding- and schema-based concepts (Schwartz & Bransford, 1998). Thus, all assessment and student activity materials were centered on similar concepts of memory and were based on Schwartz and Bransford's (1998) materials.

Sample

The study was conducted at a community college in a large urban Southwestern city in the United States. Ninety students from four Psych 101 courses of two different instructors participated in the study. There were no criteria used to exclude any students. The sample included relatively young students (M = 21 years, 80% between ages 18-22) of a variety of ethnic backgrounds, and with diverse career interests. The sample was 46% Hispanic, 37% Caucasian, 10% African American, and 7% Asian, Native American, or Middle Eastern. Career interests were wide-ranging, including business, nursing, psychology, law enforcement, education, criminology, and graphic design. Fifty six percent of the sample was female, 44% male.

Students participated in the research-related activities of this study as part of their "regular" classroom activity. All activities and assessments supported the two participating instructors' existing lessons on memory concepts, thus, both instructors provided their students with a small amount of class participation credit for participating in this study. For all four classes, these materials served as the major form of instruction on memory concepts within the instructors' curriculum. Students were not provided with any other instructional material on memory (lecture, textbook readings, etc.) prior to the study; thus, it was assumed that students had limited prior knowledge of concepts of memory (also supported by low pretest domain knowledge scores, $M = 50.8\%$, $SD = 21.6$).

Procedure, Data Collection Materials, and Activity Tasks

Procedure

A pilot study was first conducted to obtain preliminary data. Twenty students from the same community college in a separate Psych 101 course participated in the pilot. Unfortunately, these students did not have time to fully complete the posttest, thus, their data were primarily used to inform final procedures and materials for the study. As a result of the information gathered from the pilot, the study was structured to run over the course of one week. The procedures are included below.

Day 1: Students were provided consent forms and were briefed about the study. They were informed that the purpose of the study is to investigate how different ways of doing collaborative classroom activities affects learning; therefore, the researcher would be examining their classwork and observing them as they worked on class activities with a partner. It was emphasized that their participation would be completely anonymous. Students were given a generic ID code to be used for all research-related activities and asked not to include their names on any of the research-related materials. They then took a pretest on memory concepts and filled out a pre-survey on their general preference for collaborative work. The maximum amount of time given for Day 1 activities was 15 minutes and activities were facilitated by the students' instructors.

Day 2: Students were randomly assigned to one of four experimental conditions: (a) No Prep-Active, (b) No Prep-Constructive, (c) Prep-Active, and (d) Prep-Constructive. (Further details on the activity tasks are provided in a later section.) They were told to follow the instructions on their activity packet and that the researcher would inform them as to when they could stop working (these packets are included in Appendix A). For No Prep conditions, students were randomly assigned to a partner and the instructions below were given:

> You will be working on an activity about concepts of memory with your partner. Please write both your ID numbers on the packet. You will fill out only one packet per pair. You will have approximately 30-35 minutes to complete this activity. If you have any questions, please raise your hand and I will come help you. Don't worry about writing "right" and "wrong" answers. Just do your best and share your ideas! You and your partner should try to come to agreement for each question/concept within the activity. You do not have to take turns writing, but you may if you choose to.

For the Prep conditions, the instructions below were given in the individual preparation phase of the activity:

> First, you will work individually on an activity about concepts of memory. Please write your ID number on your individual packet. Do not work on this individual work with anyone. Feel free to ask me any questions, but do not ask a classmate. You will have 15-20 minutes to work on this individual work. Don't worry about writing "right" and "wrong" answers. Do your best! You will use this work when you join a partner to work on the collaborative part of the activity.

After the individual preparation phase of the activity was complete, students were randomly assigned to a partner and given the following instructions for the collaborative phase of the activity:

> You will now work on the collaborative task with your partner. Please write both your ID numbers on the Collaborative packet. You will fill out only one packet per pair. You will have approximately 10-15 minutes to complete this activity. If you have any questions, please raise your hand and I will come help you. Again, don't worry about writing "right" and "wrong" answers. Just do your best and share your ideas! You and your partner should try to come to agreement for each question/concept within the activity. You do not have to take turns writing, but you may if you choose to.

One important note regarding these conditions is that students were always assigned to a partner within the same condition. In other words, a Prep-Active student was partnered with a Prep-Active student, a Prep-Constructive student with another Prep-Constructive student, a No Prep-Active with a No Prep-Active, and a No Prep-Constructive with a No Prep-Constructive. This was to ensure as pure conditions as possible, allowing for clean manipulations of the variables under investigation. In addition, general instructions were given about how to collaborate including asking students to try to contribute equally to the discussion and to discuss each part of the activity, rather than "divide-and-conquer" by sharing aspects of the task between themselves. Students were intentionally encouraged to focus on content-based discussion rather than task management or coordination of activity to maximize the opportunities for learning from the interaction with a peer (i.e. to prevent students from "wasting" learning time by trying to determine who does which part of the activity).

The activity materials across the four conditions were equivalent in content. They addressed the exact same concepts of memory and used the same examples. Further details are provided in the Activity Tasks section.

After completion of the conditional activity tasks, the students then filled out a post-survey regarding their opinions about the activities in general, and about working specifically with their partner. Day 2 activities took approximately 35-40 minutes and were facilitated by the researcher.

Day 3: The students individually completed a posttest.[9] The posttests were administered by the students' instructors, who reported that students spent 35-50 minutes completing the tests. Students were allowed to work on their tests until the end of the class period. Any students who completed the test within 30 minutes were asked to go over their answers one more time. Figure 1 below summarizes the study design and procedures.

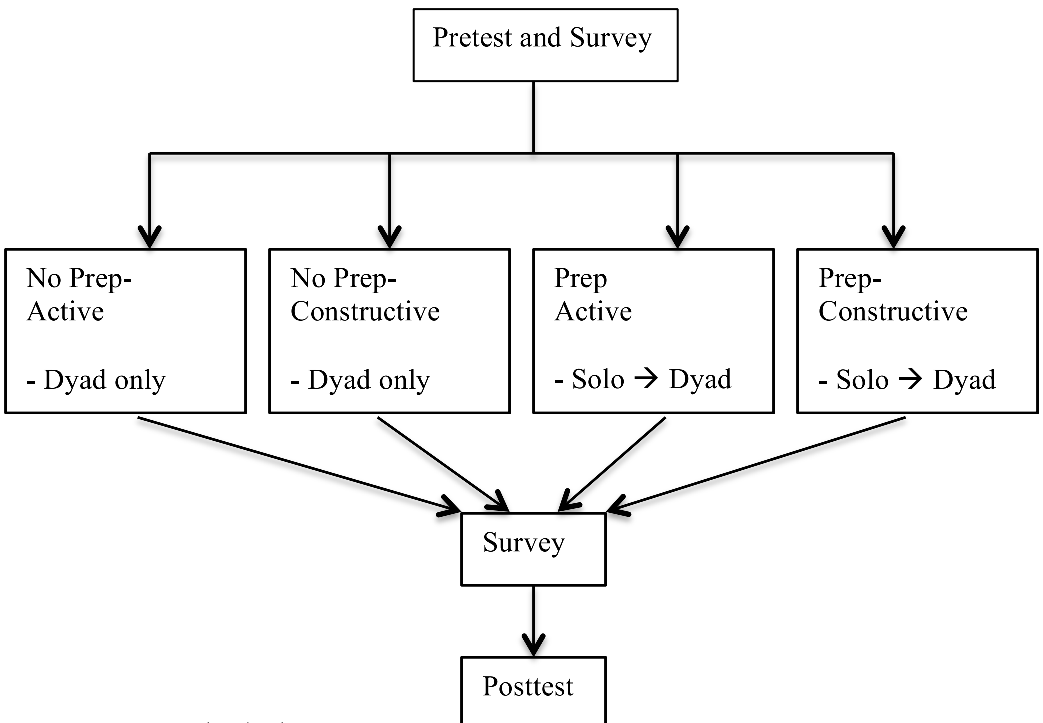

Figure 1. Study design

[9] The posttest was given on the same day of the learning activity for one class due to the instructor's course schedule. However, this class was one hour and 15 minutes long, compared to the other three classes that were 50 minutes long; therefore, there was sufficient time for students to complete all research-related activities. The survey was given in between the activity phase and posttest to serve as a distractor to avoid an immediate recall effect. There was no indication of any class effects for either shallow or deep learning outcomes, $F(3,32.2) = 1.40$, $p = .26$, and $F(3,41.5) = 1.13$, $p = .35$, respectively. In this case, the researcher administered the posttest.

Materials

The pretest and posttest used identical T/F questions that were very slightly modified from Schwartz and Bransford's (1998) verification measure, which was used in their work on concepts of memory. Two examples of T/F questions are as follows:

> When people understand something they read, they tend to recall exact sentences from it.
>
> True False I have no idea (circle one)
>
> When recalling a written story about a familiar event, people tend to remember the most stereotypical parts of the event.
>
> True False I have no idea (circle one)

A few multiple-choice questions from one instructor's existing weekly quiz on the topic of memory were also included on these tests. They covered the definitions of "retrieval," "storage," and "encoding," and were included as a request from the instructor. The results from these questions were removed from analyses since they did not directly pertain to the research questions of interest, and are not referred to from this point forth.

Although the T/F questions were identical, the order of the questions was different between the pre- and posttest and there were four to five days in between the tests to avoid a "testing effect" (i.e. learning solely attributed to the recognition of identical test questions at a later testing phase). (See work by Bjork, E. and Storm, 2011; Bjork, R.,1975; and Roediger and March, 2005, for further information regarding the conditions under which testing influences learning.) The T/F questions also included an "I don't know" choice, to prevent students from blindly guessing in order to yield an accurate measure of domain knowledge. To further prevent guessing, students were verbally instructed not to make blind guesses. These T/F questions were used to assess shallow learning, through gains from pre- to posttest.

The posttest included two additional prediction tasks that were used to obtain measures of students' deeper knowledge of memory concepts. These tasks were adapted from Schwartz and Bransford's (1998) materials on schema- and encoding-based memory concepts and involved reading through novel experiments (i.e. they were not present in any of the learning materials) and drawing conclusions about their data results. To complete these tasks, students had to study these experiments and synthesize what they had just learned about memory in order to apply their knowledge to novel conditions, generate new inferences about how memory works, predict the results of the experiments, and provide evidence of their reasoning for predictions. To provide an example, Experiment 1 presented the following scenario (next page):

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 40 participants to read the paragraph below and then they were asked to remember the paragraph as well as they could.

The Balloon Story

If the balloons popped, the sound would not be able to carry. Everything would be too far away from the correct floor. A closed window would also prevent the sound. This is because most buildings tend to be well insulated. The whole operation depends on a steady flow of electricity. A break in the middle of the wire would also cause problems. Of course, the fellow could shout. But the human voice is not loud enough to carry that far. An additional problem is that the string could break on the instrument. Then there could be no accompaniment to the message. It is clear that the best situation would involve less distance. Then there would be fewer potential problems. With face to face contact, the least number of things could go wrong.

Twenty participants read the paragraph by itself, and 20 read the paragraph with a picture attached [below].

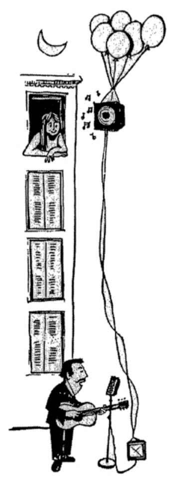

FIGURE A1 The picture seen by participants in the context-after and context-before conditions. (From Bransford & Johnson, 1972. Copyright 1972 by Academic Press.)

Figure 2: Sample question of posttest prediction task (a)

(The italicized paragraph and visual image were taken directly from Schwartz and Bransford, 1998.)

Question 2 then states:

Do you think that the participants who saw the picture will remember details from the paragraph differently than those who did not view the picture? Make 3 new predictions about what you believe participants that saw this picture will remember. Provide 2 separate explanations to explain each prediction.

1. Explain your prediction in general based on what you know about memory.

2. Explain how having the picture affects people's memories differently for each prediction you write.

The following table was provided for students to write their responses to Question 2:

Predictions – Details Remembered	Explanations for Predictions
a)	1.
	2.
b)	1.
	2.
c)	1.
	2.

Figure 3: Sample question of posttest prediction task (b)

The full versions of these prediction tasks are included in Appendix B.

In order to engage these kinds of cognitive behaviors (making inferences, predicting, reasoning) and generate correct ideas within a particular domain, one must typically have developed a corresponding mental model that is more accurate, more complete, and most likely better organized. In other words, one must have knowledge far beyond terms and definitions, or memorized facts, likened to structural accuracy at a mental model level. Chi and VanLehn (2012) refer to this as deep "structural" knowledge as opposed to shallow knowledge of "surface features." Thus, the degree to which students provided evidence of this structural knowledge through their written responses yielded a measure of deep learning. Another way to describe these prediction tasks is that they represented a measure of "transfer," which typically indicates learning at a deeper level. Because there are several definitions of transfer in the literature, I refer to the general idea that transfer occurs when people use prior learning from one kind of situation to create knowledge in a new situation (Bransford & Schwartz, 1999; Schwartz, Varma, & Martin, in press).

Because prediction tasks, such as the two used for this posttest, are by nature tasks that are more likely to be cognitively engaging, there was concern that including such tasks on a pretest might prime students in the Active conditions to actually engage constructively. To preserve the purity of the conditions (i.e. to avoid contaminating students in the Active conditions), prediction tasks were not included on the pretest. The downfall of this is that a measure of students' structural, deep knowledge could not be obtained for comparison from pre to post learning. However, because the students were assumed to have low-prior knowledge of the domain prior to the study (no formal instruction was given on the topic), obtaining a measure of deep knowledge at pretest was of far less concern than the possible contamination of conditions. In addition, the shallow knowledge pretest scores provided evidence of students' low prior knowledge, further supporting the fact that students had very limited deep knowledge of the domain prior to engaging in the study. (It is highly unlikely that students could have deep knowledge in a domain without having sufficient shallow knowledge, although the reverse could be true.) Thus, rather than a gain score, the deep learning measure used only the score obtained on the posttest prediction tasks.

The pre-survey included a set of Likert items used to obtain a measure of students' preference for working collaboratively in general. The post-survey included two Likert scales, one assessing students' feelings about the learning activities in which they participated and one assessing students' satisfaction in working with their partner. The surveys consisted of six items for each category, with two to three items that were reverse-scored. The scale ranged from one to seven, one indicating strong disagreement with the item and seven indicating strong agreement. These surveys were not adapted from any existing surveys, and were created solely for the purpose of this study. Two examples of learning activity items were:

> This activity helped me to understand the information better.
> I found this activity boring. [Reverse-scored]

Two examples of satisfaction-of-working-with-partner were:

> My partner and I made a good team.
> Doing this activity with a partner seemed to waste time. [Reverse-scored]

(The full surveys are included in Appendix C.)

Task Differentiation by ICAP

All four conditions were equivalent in domain content; however, the task categorization that differentiated Active from Constructive conditions varied according to the ICAP framework cognitive engagement definitions (Chi, 2009; Menekse et al., in press). Although the categorizations of Active and Constructive refer to students working alone in the original sense of the ICAP framework, they were used in the context of this study to refer specifically to the instructions of each task: (a) asking students to work within the existing learning material (Active) or (b) asking them to generate inferences beyond the existing material (Constructive).

For example, the Constructive task required students to invent ideas about memory (i.e. Why would people remember certain kinds of information, but not other kinds?) from studying a memory experiment and its results. In this case, students were required to decipher what the results of the experiment meant about how people's memory processes work. This is a task requiring generation of inferences, invention of concepts, and integration of various forms of information.

The Active task, on the other hand, required students to study memory terms and their descriptions. They then applied those terms to the same memory experiment included in the Constructive versions. However, the students doing the Active task did not have to generate any new knowledge. They simply had to "search and select" by writing in the memory term next to the appropriate result of the experiment provided in a list format. Because the Active task took much less time to accomplish (as indicated through the pilot study), it included a secondary experiment that was identical in structure to the first, but with a different "cover story." This was to control for time-on-task, which was equalized across all conditions. Similar strategies have been used in other work to equalize time-on-task, and essentially represent the comparison of an invention task (Constructive) with corresponding "tell-and-practice" tasks (Active) (Kapur & Bielaczyc, 2011; Schwartz & Martin, 2004). The excerpts below were taken from each of the activity types to illustrate the contrast between what is considered a form of Active instruction compared to Constructive instruction. In each case, students first read through a summary of an experiment on memory that included some data and the results. After reading the experiment summary, students were instructed to:

> [Active] … connect main ideas from the results with concepts and principles of memory [terms and descriptions included in packet]. For each [main idea] …write in the memory concepts or principles that are addressed.
>
> [Constructive] … invent some general principles and ideas about how memory works. Think about the following questions:
> Do you see any patterns in the data? …
> Why would people falsely remember information? …

> … try to come up with at least 8 ideas about how and why we remember certain kinds of information, but not other kinds.

The Prep versions of these activities provided time for students to first work on their task individually, and then discuss their work with a partner while jointly completing a collaborative worksheet. The No Prep versions simply did not include the individual preparation phase, thus, students jointly completed a collaborative worksheet for the entire 30-35 minutes.

Scoring Free-Response Data and Data Analyses

Free-Response Scores

Due to the open-ended nature of both the activity work and the prediction tasks in the posttest, several measures were taken to develop scoring procedures to objectively quantify the quality of student responses.

Activity worksheets. Because students were encouraged to write down their best answers, rather than be too concerned about obtaining the "right" answer, the activity worksheets were scored by student effort rather than by correctness of responses. This is similar to a "dynamic assessment" (Bransford & Schwartz, 1999), further described in the Results and Discussion chapters. For Active conditions, each dyad that completed 15-16 questions on the worksheet (94-100% completion) received two points, dyads that completed 12-14 questions (at least 75% completion) received one point, and dyads that completed fewer than 12 questions (under 69% completion) received no points. In addition, the Active tasks included a segment for each question where students could explain any disagreements that arose during discussion (this was included on both the Prep and No Prep activity sheets for the Active conditions). The number of dyads who had disagreements and the total number of disagreements in each condition were taken into consideration when assessing student effort and engagement (as supported by work that has found argumentation to improve learning, Asterhan & Schwarz, 2010; Schwarz et al., 2000).

The Constructive worksheets were scored differently, since these tasks were completely open-ended. In these conditions, students were instructed to develop ideas about how memory works, thus, the number of ideas presented in the collaborative worksheets was used to assess effort.

In particular, engagement in the learning task for the Active conditions and its influence on the prediction task results are addressed in detail in the Results chapter due to some unexpected outcomes. The Constructive task effort scores are briefly reported on, as these results supported the main outcomes.

Posttest prediction tasks. One might argue that because the prediction tasks were constructive by nature, those students who worked in Constructive activity conditions had an advantage. Perhaps, because those students had "practice" in engaging in constructive cognitive behaviors before the test, this could lead to improved performance. However, the specific cognitive behaviors elicited from these prediction tasks can be differentiated from the behaviors elicited from the activity tasks. In particular, the Constructive activity tasks required analysis of experiments and the "invention" of concepts. The prediction tasks at posttest, on the other hand, required the analysis of more complex experiments, the "synthesis" of domain knowledge, and the formulation of "hypotheses" in order to predict results. While these can be considered in the same realm of type of cognitive activity (constructive), they are still distinct behaviors. Thus, in light of framing the prediction tasks as transfer tasks, students in all conditions (not only the Active) needed to engage in some aspect of transferring knowledge in order to complete the posttests.

The student responses to the prediction tasks were scored by how well the following eight concepts were addressed: elaboration, schemas, generation effect, obstacle recall, gist, serial position effect, interference, and encoding failure. Table 2 provides general descriptions of each of these concepts.

Table 2
Terms and descriptions of memory concepts

Term	Description
Elaboration	People tend to remember new information when they are able to link or attach it to prior knowledge.
Schemas	People tend to remember information that supports their existing knowledge structures.
Generation Effect	People tend to remember information about an event that they generate themselves, rather than actual details about it.
Obstacle Recall	People tend to remember parts of events that hinder the completion of a goal.
Gist	People tend to remember the general overview or main idea of an event, an experience, or information.
Serial Position Effect	People tend to remember the beginning and ending details of an event or set of information.
Interference	People tend to forget information when new, incoming information disrupts or overtakes existing memories.
Encoding Failure	People tend to forget details that are given little to no attention at the time of encoding.

These are common concepts of memory taught in introductory psychology courses, as verified by both instructors who participated in this study, as well as the psychology textbook used by one of the instructors (Carter & Seifert, 2013). Each of these concepts was explicitly taught (through the Active “search and select” activities) or implicitly taught (through the “invention” of concepts in the Constructive activities).

Prior to a direct coding of these concepts and final scoring of student work, a generic method of triangulation was used to cross-examine the data. This was necessary because the student answers were rarely straightforward. A fairly clear-cut criterion such as, “student lists concept and correct definition,” could not be used since half of the sample (the students in the Constructive conditions) were not explicitly taught any terminology. Many of the students’ prediction task responses indirectly represented a concept and had to be interpreted within the context of their overall effort. For instance, to answer the question of how a visual image of a story may affect people’s memory, a student responded, “…the picture clarifies the idea since its very broad & would not be as easily remembered otherwise.” The student continues to explain at a later point, “…the picture of the man singing to the girl with the microphone just helps people put two & two together.” This student’s responses represent the concepts of gist (i.e. the picture portrays a “broad” idea about the story that is more easily remembered) and elaboration (i.e. associating or linking one piece of information to another by putting “two and two together” helps memory). Therefore, because most of the student work could not be scored in a straight-forward manner, three separate kinds of scorings were used, which then informed the development of a coding manual and rubric for a final scoring.

First, all student tests were given a general score of low, mid, or high to represent an overall sense of the quality of the work. Although these were very “rough” scores, they were recorded to serve as crosschecks for other scorings. Next, 20% of the tests were randomly chosen (with equal representation from each of the four conditions) to provide more fine-grained double-checks. The student responses from these tests were categorized by concept (the eight listed above plus a category of Other), and each student received a score based on the number of concepts represented (score from 0-9; one point for each concept and one point for Other). Finally, the same 20% of the tests were rank-ordered from an overall "best" to "worst" with regard to quality of student explanations and evidence of reasoning. The researcher was blind to conditions throughout this process of cross-examination.

The concept scoring corresponded to a 72% agreement with the initial scoring of low, mid, high, and the ranking from "best" to "worst" was in agreement with the low, mid, high scores when grouped into three chunks. Thus, based on the process of this cross-examination, a coding manual and rubric were developed in order to score the entire set of posttests. The coding manual served to produce a score representing student knowledge of a variety of memory concepts, while the rubric included an additional dimension measuring students' quality of reasoning. To score knowledge of concepts, three points were given for evidence of the presence of six or more concepts, two points for 3-5 concepts, and one point for 1-2 concepts. To score quality of reasoning, three points were given when the majority of the concepts were clearly linked to the predictions through students' explanations, two points when at least half of the concepts were clearly linked to predictions through students' explanations or the majority of the concepts somewhat related to predictions but were not explained in detail, and one point when concepts were loosely linked to predictions and minimal effort was used to explain reasoning. The two dimensions of this rubric yielded a score of 0-6 for each student. A copy of the rubric and coding manual may be found in Appendix D.

In addition to the cross examination of data that was done to develop a coding manual and rubric for objectively scoring the students' free-written responses, an interrater reliability check for two raters was computed on a randomly selected 20% of the data (with equal representation from each condition). One of the community college instructors scored 18 posttests using the coding manual and rubric. She was blind to conditions and had no knowledge of who the students' were, as the only identifying information on the tests were the generic ID codes. This instructor is a domain expert; she received a Ph.D. in cognitive psychology and has been a psychology professor for several years. She was given brief instruction on how to score the data outside of the coding manual and rubric, including explanation to try to score the work somewhat holistically, as a teacher might score a free-written essay. Consequently, she followed the coding manual and rubric, while considering how each final score "agreed" with her overall sense of that student's work.

Intraclass correlation (ICC) was used to compute the consistency between the two raters, the aforementioned instructor and myself (Shrout & Fleiss, 1979). The ICC(2,1) was .76, $p < .001$, which is typically considered acceptable for research purposes (AERA, APA, & NCME, 1999).

Data Analyses

MLA versus ANOVA. Analyses of both the shallow learning and deep learning scores involved a comparison of the means across conditions using a multilevel analytical (MLA) technique. For student shallow learning, adjustments for learning gains were made using "normalized change" (c) calculations to account for influences of pretest scores, yielding a more sensitive measure of gain scores (Marx & Cummings, 2007). The formulas for normalized change are as follows (next page):

c = post-pre / 100-pre, when post > pre
c = 0, when post = pre
c = post-pre / pre, when post < pre

Adjusted gain scores give more weight to high scores that increase from pre- to posttest. In other words, if a student scores relatively well on a pre-assessment, it makes it more difficult to improve compared to a student who scores relatively low. Thus, on a pretest score of 90%, an increase of 10% produces a higher gain score using a normalized change calculation compared to a 10% increase on a pretest score of 40%. Students' feelings about the activities and about collaborating were also compared across groups using a multilevel model.

ANOVA techniques have been commonly used in experimental collaborative learning studies to assess differences across conditions. This poses an issue; ANOVA is not robust to violation of the assumption of independence of subjects. Since the students in this study (all subjects in the sample) discussed the domain content with a partner for the purpose of improving learning, by design, there is dependency among subjects. Therefore, rather than using traditional ANOVA, a MLA technique that accounts for the dependency of students within dyads was used to compute F values. Because these models are less common in the supporting literature, some general background is provided below.[10]

Multilevel analytic methods yield more valid results for studies where subjects are interacting with one another and the resulting outcomes may be affected by that

[10] Note that both ANOVA and MLA were used to evaluate outcomes for the purpose of comparison. The significance tests between the ANOVA results and the MLA results were similar, such that with an alpha set at .05, there were no differences with regard to rejection of a null hypothesis. Thus, the only reason for using the multilevel model was to obtain the most reliable results, since it can account for the possibility of nonindependence. Although the two models essentially produced the same results, the multilevel analysis results are reported.

interaction (Kenny, Kashy, & Cook, 2006; Kirschner, Paas, Kirschner, & Janssen, 2011). Specific to dyadic data, one way to check for "nonindependence" is to determine if the outcome scores of individuals are correlated (Kenny et al., 2006). The question of interest here is: Is an individual's score more similar to his/her partner's score than another random individual in the sample? The intraclass correlation (ICC) is the appropriate measure to use for dyadic data that has "indistinguishable" partners (such as students in a classroom, equal-level coworkers, or same-gender friends, as opposed to husband-wife, parent-child, expert-novice, etc.). Using the Pearson *r* to compute a correlation between scores of indistinguishable partners poses an issue with the positioning of the partners' scores (i.e. which score appears in the first column of the data set). Since the positioning of scores of indistinguishable partners is completely arbitrary, the ICC provides a more accurate measure of the relationship.

For learning outcomes such as the ones used in this study, a positive intraclass correlation between partners' scores within each experimental group would be expected, thus, providing evidence of the dependency among the students within dyads. However, this positive relationship was not present in the shallow nor deep learning data, thus, the analytical path to take was less clear.[11] One might argue that the lack of relationship present in the scores means that independence can be assumed, and then in that case, there is support for using the more common ANOVA techniques. However, the assumption of independence actually lies in the design of the study, not necessarily in the results of any particular outcome measures. So others might argue for using a multilevel

[11] The ICCs computed on my data were not consistent across groups. None of the ICCs were significant and the majority were close to zero and/or negative. They are reported in Appendix E.

model, despite the lack of evidence of nonindependence. I take the latter view that, by design, this study violated the assumption of independence of subjects; therefore, it was most appropriate to use a MLA technique to be confident in the validity of the results. The dyadic design for the data from this study nested individual students within dyads (n = 2), with the level two factors of Preparation and Type of Activity. Figure 4 on the following page displays the structure of the model.

Level 2 Factors: Preparation and Type of Activity

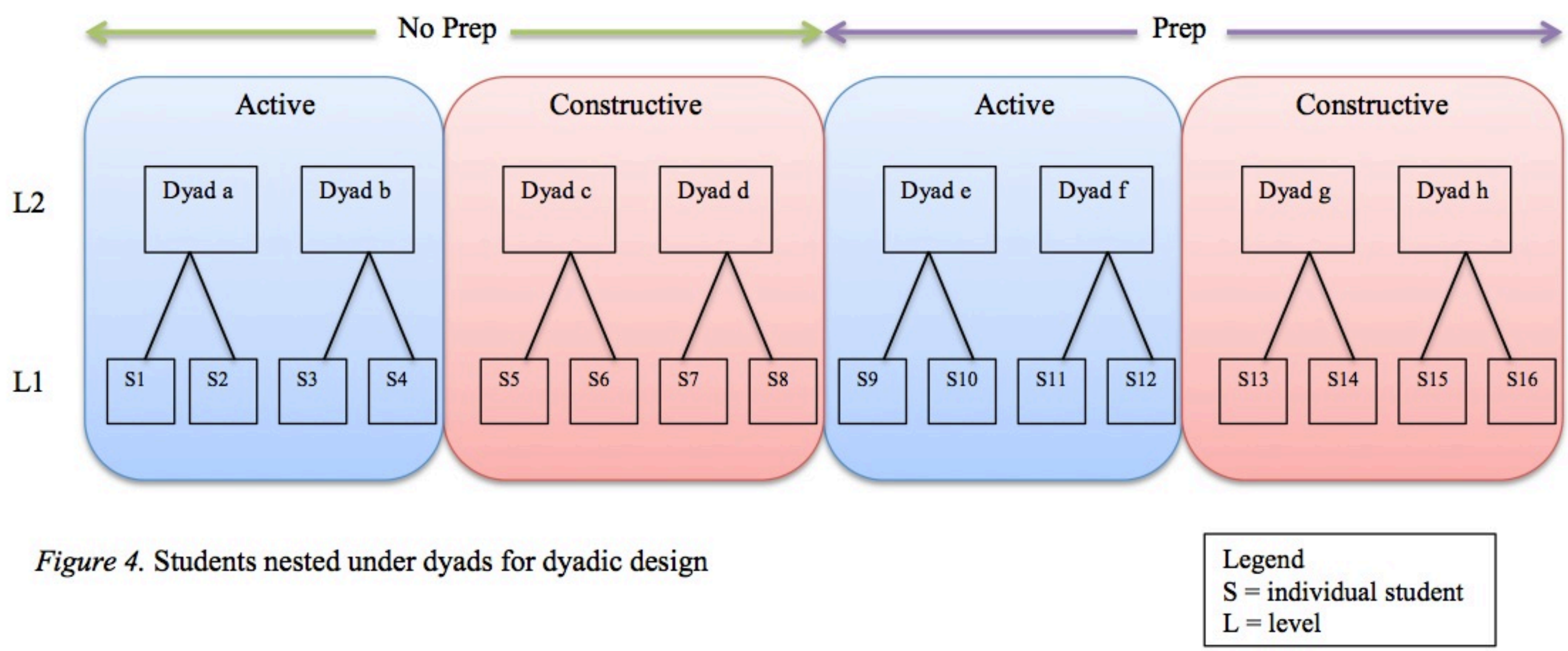

Figure 4. Students nested under dyads for dyadic design

Reliability of Survey Items. Because the survey items were created specifically for the purpose of this study and not taken from any standardized assessments, Cronbach's alpha was used to assess the correlations between the items. For the pre-survey items (measuring general preference for collaboration), Cronbach's alpha was .83. Every survey item was positively correlated with the other items. The post-survey items targeting students' positive feelings about the activity also were fairly consistent; Cronbach's alpha = .80. Every item was positively correlated with the others. The post-survey items measuring students' satisfaction with the collaborative experience with their specific partners did not have as high a consistency score, Cronbach's alpha was .70. Upon examination of the correlations between items, one ("I think that I would have done this activity better by working alone") was not as strongly correlated with some of the other items, relative to all other survey item analyses. After reviewing the wording of this item compared to the others, it was deemed somewhat qualitatively different. Removing the item increased Cronbach's alpha to .72, thus, the analysis of this particular construct was based on the five items. The correlations for the items of these reliability analyses are included in Table 3.

Table 3

Item consistency correlations of three survey constructs

Pre-survey general preferences for collaborating	2	3	4	5	6
1. I enjoy working with a partner.	.57	.71	.58	.61	.28
2. I like being able to share my ideas with someone.		.44	.30	.39	.47
3. I prefer to work alone.*			.44	.59	.26
4. I usually find a partner's feedback helpful.				.48	.33
5. I feel like I am on a team when I work with someone.					.29
6. I feel that working with others usually hinders my progress.*					
Post-survey feelings about the activity	**2**	**3**	**4**	**5**	**6**
1. This activity helped me to understand the information better.	.70	.41	.54	.32	.46
2. I enjoyed participating in this activity.		.24	.63	.34	.34
3. I was confused about the purpose of this activity.*			.29	.22	.45
4. I would like to participate in similar activities … in class.				.51	.37
5. I found this activity boring.*					.31
6. This activity was very difficult.*					
Post-survey satisfaction with partner experience	**2**	**3**	**4**	**5**	**6**
1. I enjoyed working with my partner.	.24	.32	.55	.71	.29
2. Explaining … helped me to understand the information.		-.04	.26	.26	.41
3. I … would have done this activity better by working alone.*			.28	.18	.30
4. It was helpful to receive input from my partner.				.48	.31
5. My partner and I made a good team.					.36
6. Doing this activity with a partner seemed to waste time.*					

Notes: Symbol (*) indicates reverse-scored items. Symbol (…) indicates removal of words for the sake of space. The full surveys can be found in Appendix C.

Chapter 4

RESULTS

The first set of analyses presented address the students' shallow learning through their gains from pre- to posttest. Then, results focus on the deep learning outcomes from the prediction task scores. The multilevel analytical (MLA) technique was a linear mixed model with the Restricted Maximum Likelihood method (REML) using SAS statistical analysis software. This model accounts for both fixed and random effects, and is especially useful to account for the dependency between subjects within clusters (in this case, dyads). Finally, the differences between shallow and deep learning outcomes are compared. The implications of these findings are covered more thoroughly in the Discussion chapter.

Shallow Learning

Adjusted gains from pre- to posttest on the T/F questions through "normalized change" calculations (Marx & Cummings, 2007) were used in the shallow learning analysis. Because a dyadic design was employed, *df* refer to the number of total dyads (rather than individual students). Although students' scores at posttest increased in all conditions, there were no reliable differences between conditions. There was no significant main effect of Preparation, $F(1,34) = .07$, $p = .79$, Type of Task, $F(1,34) = .01$, $p = .94$, nor an interaction effect, $F(1,34) = 2.29$, $p = .14$. Table 4 summarizes the pre- and posttest scores.[12]

[12] For descriptive data, *N/n* is reported at the individual level.

Table 4
Shallow learning scores

Condition	*n*	Means		
		Pretest%	Posttest%	Adj. Gain
No Prep-Active	14	53.6	72.6	.43
No Prep-Constructive	18	49.1	61.1	.21
Prep-Active	15	46.7	63.3	.28
Prep-Constructive	19	53.5	71.9	.40
Total	66	50.8	67.2	.33

Note: Due to absence or incompletion of either pre- or posttest, the total sample was reduced from 90 to 66 students.

These results are not surprising for a few reasons. First, it was expected that students in all conditions would improve at posttest. Even what could be considered the "control" condition (No Prep-Active) constituted an effective instructional strategy in several ways. Students in this condition were provided terms and definitions, the opportunity to apply those to real-world examples, and the benefit of engaging in discussion. The other three conditions provided "extra" benefits such as engaging more constructively and/or individually engaging as to better prepare to collaborate, so students in all conditions were expected to learn at least to a shallow degree. In other words, even the "worst" condition was good enough to improve shallow learning, relative to the other interventions.

Deep Learning

Ninety students individually completed the prediction task portion of the posttest. Tests were scored via the coding manual and rubric described earlier, and this score served as the overall measure of deep learning. Forty-six dyads (with two dyads missing data) were evaluated using MLA. The prediction task posttest scores were significantly different across conditions. There was a main effect of Preparation on learning,

$F(1,41.1) = 5.79$, $p < .03$, but no effect of Type of Task, $F(1,41.1) = .75$, $p = .39$, nor an interaction effect, $F(1,41.1) = .59$, $p = .44$. Students who engaged in individual preparation before collaborating showed evidence of deeper learning, as indicated by higher prediction task scores. The figure below summarizes these results.

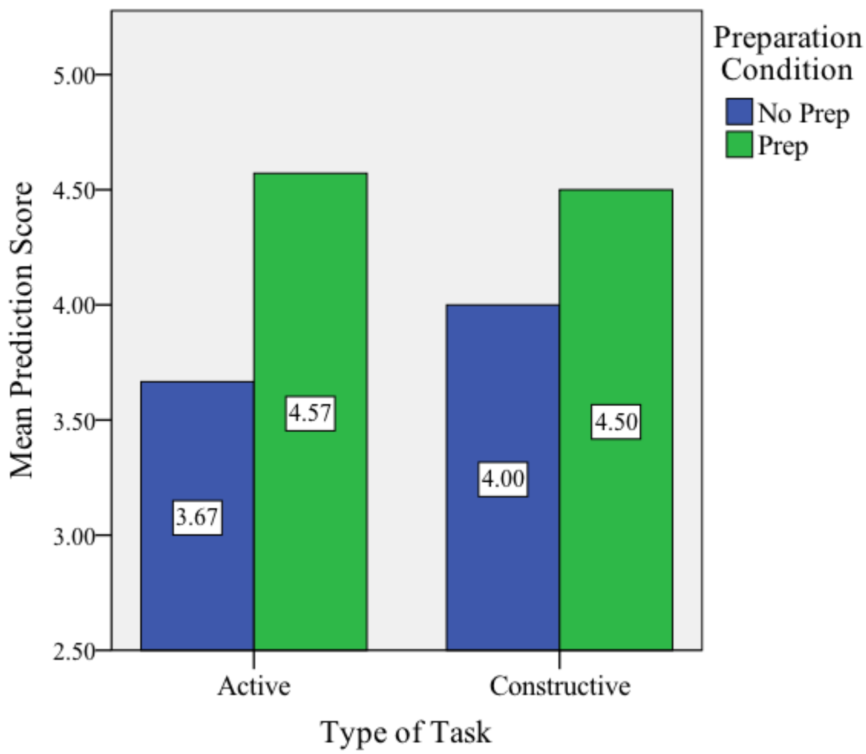

Figure 5. Prediction task results

The main effect found for Preparation in these data is theoretically supported by the Preparation for Future Learning (PFL) paradigm, in that individually working on the tasks may have made students more “ready” for learning in a future task, but instead of lecture, the future task in this study was peer discussion. The fact that there was essentially no difference in deep learning between the Prep-Active and Prep-Constructive conditions is contrary to PFL, because PFL work has found constructive-type tasks to improve readiness above others (tell-and-practice, worked example, etc.). In addition, the

ICAP hypothesis asserts that constructively engaging activities should produce improved learning above actively engaging activities (Interactive > Constructive > Active > Passive). There is a trend supporting ICAP within the No Prep conditions (the No-Prep Constructive group scored slightly higher than the No-Prep Active group), however, the difference between these groups was not significant.

One interpretation of these results is that the inclusion of preparing prior to discussion in a collaborative activity boosts learning, such that it overrides any effect of type of task. The lack of difference in the deep learning outcomes between the Prep-Active and Prep-Constructive conditions ($M = 4.57$ and $M = 4.50$ respectively) suggests that preparation may increase the likelihood that students will more deeply engage in a collaborative activity, regardless of the specific task instructions. In other words, it is possible that students in the Prep-Active condition actually engaged constructively in the task, which then explains the null result (i.e. the comparison becomes one Prep-Constructive activity against a different Prep-Constructive activity).

To check if preparation prompted students to engage constructively in the Active task, exploratory analyses of the student activity work were done. Analyses compared the No Prep-Active and Prep-Active conditions, as well as the No Prep-Constructive and Prep-Constructive conditions.

Analyses of Active Worksheets

Effort and prevalence of disagreement were used to quantify the level of engagement students had in the No Prep- and Prep-Active conditions. These indicators are analogous to those used in "dynamic assessments" (Bransford & Schwartz, 1999; Feuerstein, 1979). A dynamic assessment functions like a formative assessment measuring student progress, rather than mastery of content (such as a summative assessment). In particular, the worksheets in this study served as a measure of engagement as they indicated, to some level, student progress in learning.

Effort was scored on a 0-2 range and the average number of disagreements per dyad across groups was calculated. For these data, the dyad served as the unit of analysis, thus, traditional ANOVA techniques were sufficient as there was no dependency between dyads. Results showed that dyads in the Prep-Active condition produced a higher effort score (M = 1.73, SD = .93) compared to those in the No Prep-Active condition (M = 1.45, SD = .65), and had a slightly higher average number of disagreements (.55 compared to .45 respectively). However, none of these comparisons were significant (p = .44 for effort, p = .80 for disagreements). Although these differences are relatively small, put together they illustrate a pattern suggesting that preparation may have influenced students to engage more deeply in the activity, even though the activity itself did not require such engagement. In other words, although this Active task by instruction merely required the assimilation of new information by matching terms/descriptions to examples, students may have spontaneously engaged by creating new knowledge and thinking more deeply about the content (which is considered constructive). Thus, although the label "Active" reflected the instruction to engage, it may not have been an adequate reflection of the actual student engagement in the task. Of course, effort and number of disagreements are not completely comprehensive to assess engagement and these differences that were found are small, emphasizing the speculative nature of this interpretation. Analyzing the discourse between students in dyads could provide much insight, but that was beyond the scope of this work.

Constructive Activity Worksheets

For the sake of completeness, the Constructive activity work was also examined. For the Constructive conditions, the number of ideas about memory that students generated was used to indicate engagement across these two groups (Prep and No-Prep). The responses were also qualitatively used as an indication of students being on task, since off task behavior can affect learning outcomes.

For these data, the dyad also served as the unit of analysis, thus, traditional ANOVA techniques were sufficient. First, there was no indication of students being off task. Students' ideas were relevant to the activity materials, as evidenced by their explanations. Most dyads completed the task assignment (to try to come up with at least eight ideas), and many went beyond what was required. Regarding differences between the No-Prep and Prep conditions, dyads in the Prep condition produced more ideas on average ($M = 10.17$, $SD = 3.38$) compared to the No Prep condition ($M = 7.92$, $SD = 2.68$) and this result approached significance, $F(1,22) = 3.27$, $p = .08$. In this case, non-significance might be attributed to a lack of power due to small sample size. This result supports the deep learning outcomes (the Prep-Constructive group outperformed the No Prep-Constructive group, indicated by the deep learning scores).

Analyses of Survey Data

These data were used to assess how the factors of Preparation and Type of Activity influenced students' motivation to engage with a partner or in the activity in general, which can influence learning. The same dyadic model was used for these data and MLA was used to check for differences across conditions. In general, these variables did not seem to be highly influenced by the activity in which students engaged. There was a slight preference for the Active activities compared to the Constructive, for the post-survey measuring students' positive feelings about working with their partners, $F(1,40) = 4.00$, $p = .05$, and for the post-survey measuring students' general enjoyment of the activities, $F(1,40) = 10.00$, $p < .01$, (*df* are reported at the dyadic level). The table below summarizes the means for all of the survey-related variables. A score of one indicated low preference/positive experience/enjoyment, while a score of seven indicated high preference/positive experience/enjoyment.

Table 5
Survey outcomes by conditional group

Condition	Mean/SD		
	(Pre-survey) Preference for collaborative work	(Post-survey) Positive experience with partner	(Post-survey) Enjoyment of activity
Preparation			
No Prep	4.83/1.11	5.97/.78	4.46/1.23
n	40	46	46
Prep	4.67/1.36	6.00/.96	5.09/1.38
n	40	42	42
Type of Activity			
Active	4.93/1.23	6.19/.73	5.29/1.37
n	38	42	42
Constructive	4.59/1.24	5.80/.94	4.27/1.10
n	42	46	46
Total	4.75/1.24	5.99/.87	4.76/1.33
Total *N*	80	88	88

Note: *N/n* are reported at the individual level; *n* is reported as the total for each level within a factor. *N/n* vary due to incompletion of surveys.

In general, students had a slight preference for collaborative activities ($M = 4.75$), had highly positive experiences with their particular partners ($M = 5.99$), and seemed to enjoy the activities overall ($M = 4.76$). Regarding the higher means for the Active activities for both post surveys, this pattern is similar to the findings from studies that have examined productive failure (Kapur, 2008). In essence, the Active tasks are comparable to well-structured problems, while the Constructive tasks are comparable to ill-structured problems. Students tend to enjoy solving well-structured problems more because they are straightforward and less confusing; however, learning improves more after engaging in ill-structured problem-solving. Further interpretations are included in the Discussion chapter.

Additionally, a linear mixed model that included all three survey variables using the MLA technique was conducted. The pre-survey and post-survey on feelings about collaboration were shown to be significant predictors of the learning outcomes, $t(67.8) = -2.00$, $p < .05$, and $t(70) = -1.97$, $p = .05$, respectively. Essentially, every increased point towards preference for collaboration or satisfaction with partner predicted a slightly reduced deep learning score (regression coefficients were -.22 and -.33 respectively). In sum, deep learning was influenced more by cognitive preparation in the task than it was by students' preference for or positive experience with collaboration or enjoyment of the activities.

Differences in Shallow and Deep Learning

It is clear that shallow learning and deep learning were affected differently by the instructional interventions. The shallow outcomes did not differ significantly across conditions, while the deep learning outcomes showed a main effect of Preparation. In order to make further interpretations about the differences between the shallow and deep learning results, the raw scores were converted to standardized scores. The graph on the following page displays the z-score means for both shallow and deep learning.

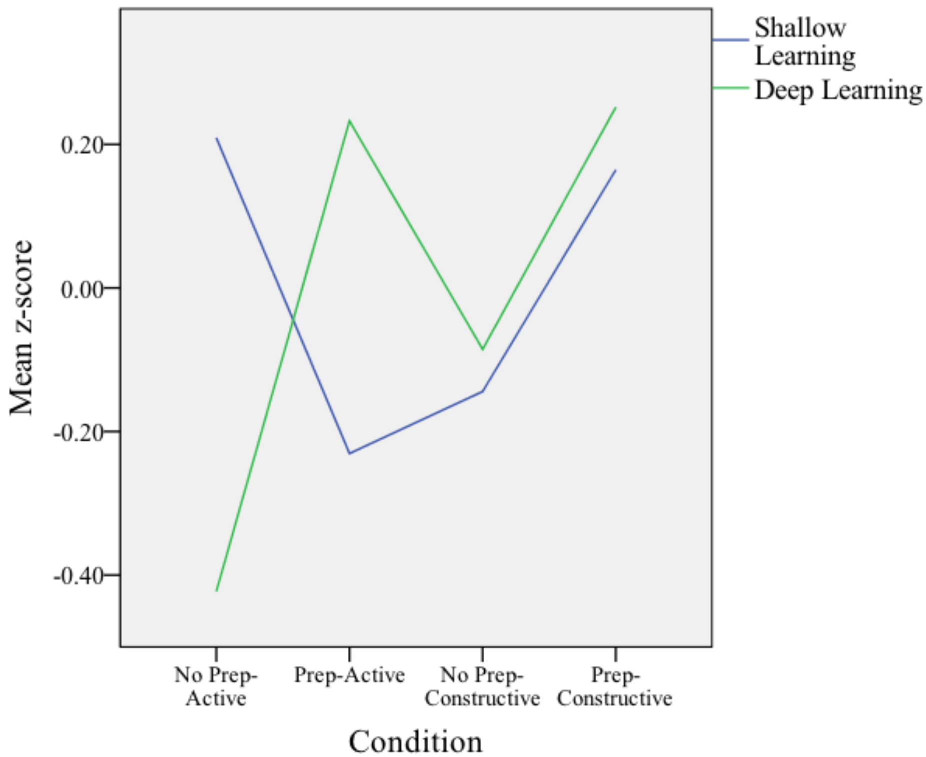

Figure 6: Comparison of shallow and deep learning results

Regarding the pattern of shallow and deep learning across conditions, most noticeable is the difference in the Active conditions across the Preparation factor compared to the Constructive conditions. The No Prep-Active instructional intervention essentially produced the highest shallow learning score and the lowest deep learning score. The opposite was true for the Prep-Active intervention. For the Constructive conditions, the pattern for both shallow and deep learning is similar. The work by Chi et al. (1994) on knowledge categories and the cognitive foundations of the ICAP framework provide some explanations for these differences, which are included in the Discussion chapter. In short, the instructional tasks were not designed to target shallow learning differently, and the null results are supportive of that. They were designed to activate students' mental models differently, which aligns to deep learning and structural-level understanding. Thus, the hypothesis that deep learning, in particular, would be significantly affected by the instructional interventions is supported.

Chapter 5

DISCUSSION

We know that peer discussion can enhance learning in collaborative activities. Both classroom and laboratory studies provide evidence that collaboration improves outcomes above working alone (Chi et al., 2008; Gadgil & Nokes-Malach, 2010, 2011; Hausmann et al., 2009; Hausmann, Van de Sande, & VanLehn, 2008; Muldner et al., 2011; Schwartz, 1995; Shirouzu et al., 2002) and extensive research has introduced a variety of ways that collaborative activities can be designed to maximize learning. These include training students to use effective collaboration skills (King, 1990, on question-asking; Rummel & Spada, 2005, and Rummel et al., 2009, on collaboration training through observation of a model; Uesaka & Manalo, 2011, on co-constructing diagrams); guiding students interactions during collaboration (Coleman, 1998, on prompting explanations; King, 1994, on prompting question-asking; O'Donnell, 1999, on scripting interactions; Walker et al., 2011, on reciprocal peer tutoring via computer-based adaptive support); and designing open-ended tasks that elicit meaningful dialogue (Engle & Conant, 2002, on productive disciplinary engagement; Janssen, Erkens, et al., 2010, and Van Boxtel et al., 2000b, on using graphical representations; Kapur, 2008, on productive failure; Kapur & Bielaczyc, 2011, and Schwartz & Martin, 2004 on invention tasks). However, a number of challenges and limitations that have been acknowledged in the existing literature also tell us that there is still work to be done. For instance, students can often engage helpful cognitive behaviors after just being instructed to do so or while being guided or prompted throughout their interactions, but these behaviors tend to fade away after time or disappear when the guidance is removed (Webb & Mastergeorge, 2003; Webb et al., 2006). Open-ended and ill-structured tasks fair well for naturally eliciting effective dialogic behaviors towards learning, but only when students have the appropriate prior knowledge to begin with (Nokes-Malach et al., 2012; Wiedmann et al., 2012). Additionally, when assessments do not differentiate between shallow and deep knowledge, it becomes difficult to know exactly how collaborative activities are affecting student learning (Chi, 2009, on ICAP; Chi et al., 1994, on question categories). Finally, with the development recent statistical models that are more appropriate for nonindependent data (Kenny et al., 2006; Kirschner et al., 2011), we must be cautious of the recommendations from prior studies that have used less appropriate analytical techniques.

My work adds to the existing literature on collaborative learning by testing the effects of cognitive, domain-based preparation on learning outcomes after collaborating, measuring student shallow and deep knowledge, and employing a cutting-edge multilevel statistical model for data analyses. This "preparing-to-interact" approach addresses the limitations of training students and guiding interactions because: (a) it focuses on readiness of learning as domain-based knowledge (Schwartz et al., 2007) versus acquisition of skills, avoiding the problem of students failing to use learned skills; and (b) it allows for tasks to be open-ended so that students are not restricted by highly structured scripts or instruction to interact, while also providing the chance for students to acquire the minimum knowledge necessary to engage in a substantive discussion. Additionally, the "preparing-to-interact" collaborative model combines two cognitive-based learning frameworks that have not yet been used together for instructional design, namely, the ICAP framework (Chi & colleagues, 2008-in press) and PFL paradigm (Schwartz &

colleagues, 1998-in press). Although Froyd (2011) discussed the possible benefit of using collaboration as a future task in a PFL model, it had not been tested empirically prior to my work.

Specifically in my study, students from Prep conditions engaged in either an Active or Constructive task individually, and then discussed their work with a partner. Their shallow and deep learning outcomes were compared against those of students who worked on the tasks with a partner the entire time (No Prep). For analyses, I used a specialized multilevel model that can cope with the dependency between partners within dyads. In the following paragraphs, I address each of the initial research questions that inspired this work, and then cover further interpretations and implications of the findings.

Addressing the Research Questions

Question 1. Does individual cognitive preparation in a specific domain topic prior to engaging in collaboration have an effect on learning outcomes after collaborating?

Preparation led to deeper understanding of concepts of memory, as evidenced by students' performances on the prediction tasks at posttest. These tasks served as measures of students' mental models of the concepts, in particular, their knowledge at a deep, structural level. In these tasks, students were instructed to freely write responses to open-ended questions that required the analysis of novel experiments on memory, the synthesis of their newly acquired knowledge about memory, and the formulation of hypotheses to make predictions about the results of these experiments. This type of measure is appropriate to assess deep knowledge because it is considered an externalization of students' internal representations of the concepts (Haugwitz et al., 2010; Van Amelsvoort et al., 2007).

This finding extends the work on the Preparation for Future Learning paradigm, showing that peer discussion can serve as a beneficial future learning task when students improve their readiness to learn through a preparation task (i.e. the future task may not necessarily need to be a form of didactic instruction). This result also supports and extends the conclusions of Van Boxtel et al. (2000a), who used a similar study design. They found that even a minimal amount of preparation improved outcomes compared to no preparation with extended time to work jointly on the task. While their study focused more on discourse processes as related to learning, my work specifically examined how the instructional interventions affected the depth of domain-based knowledge.

Question 2: How does the type of task in which individuals prepare prior to collaborating affect learning outcomes after collaborating?

The results were less clear about how the type of preparation task affects learning after collaborating. There was not a significant interaction effect, nor an effect of Type of Task in general, however, exploratory analyses of the students' collaborative work show that behaviors that can be categorized as more constructive seem to be better for preparation than behaviors that are merely active. This is speculative since none of these analyses was significant. The pattern that was detected from the exploratory analyses hint at a benefit of constructive engagement in preparation, which is supported in theory by both PFL and ICAP.

Existing research also supports the benefit of constructively engaging tasks versus those that are actively engaging for the purpose of preparation for collaborative learning. For example, Van Boxtel et al.'s (2000a) study found evidence that creation of a concept map (which they deemed an abstract task, which can be considered constructive) compared to a poster (a concrete task, considered active) produced dialogic behaviors that positively correlated with learning. In addition, the historical work that has been conducted on think-pair-share instructional strategies (Lyman, 1981; McTighe & Lyman, 1988) and more current studies on Clicker technologies (Smith et al., 2009) indirectly support the use of constructively engaging preparation tasks in the form of "thinking questions" for peer discussion. Using questioning, or deep questioning in particular, to prompt domain-based thinking tends to be a highly engaging activity that promotes effective discussion and learning (Chin & Osborne, 2010a/b; Gholson et al., 2009). However, the results obtained from my study were not sufficient to thoroughly answer this question.

Question 3: How does the type of task in which individuals engage while collaborating affect learning outcomes?

This question can be mainly addressed by examining the two No Prep conditions. The sample data support the ICAP hypothesis, showing that the Constructive condition produced slighter higher deep learning scores than the Active, however, this result was not significant. One argument can be made that because these are collaborative tasks, they should actually be categorized as Interactive. A null result in this case could then be explained by the comparison of one Interactive task with another Interactive task. However, this is highly unlikely because the analyses of the student work suggest that the students in the No Prep-Active condition were probably engaging actively, while the students in the No Prep-Constructive condition may have been engaging constructively, but not necessarily co-constructing knowledge. Chi (2009) discusses the idea that working in pairs does not automatically make engagement Interactive, and that to be truly Interactive, both students must at minimum be engaging constructively.

Another interpretation of this result is that although the outcomes were relatively equivalent, there is value-added for the Constructive condition. There was essentially no further "cost" to engaging students constructively in the task (time-on-task and content of learning materials were equalized across conditions), and yet they reaped the benefits of engaging in deeper thinking compared to the Active group (that was, in a sense, "given the answers"). This is comparable to the study by Van Boxtel et al. (2000b) who found that when students had textbooks available for a collaborative task, they tended to consult the textbook to complete the assignment, rather than try to think through the work on their own. Their performances at posttest were comparable to those of students who worked without the textbook as a resource. The researchers' interpretation of these findings was that the textbooks constrained discussion and prevented students from engaging deeply in the content. The students did "less thinking for themselves" when the textbook, which basically provided the answers, was available (p. 71). The Active learning materials used for my study may have functioned in a similar manner for these No Prep students by providing some "answers" without prompting students to think more deeply. A retention test could provide further insight that teases apart the effects of these conditions on learning.

Question 4: As related to the three questions stated above, how is deep learning affected differently than shallow learning?

It was critical to use an appropriate assessment for deep, structural knowledge to determine how learning was affected by these collaborative tasks. Recall the differences in the patterns for deep and shallow learning across the Preparation factor, in particular, for the Active tasks. Within these tasks, the No Prep activity produced the highest shallow learning scores and the lowest deep learning scores, while the Prep activity produced the highest deep learning scores, but the lowest shallow learning scores. The work on self-explanation by Chi et al. (1994), which included the categorization of test questions that aligned to depth of knowledge, in conjunction with the cognitive foundations of the ICAP framework can explain some of these differences. Basically, "verbatim" (shallow) knowledge can be well-developed by receiving and storing information, where Passive activities are sufficient instructional strategies, and can be measured by testing students' recall of fact-based (or surface-feature) knowledge. "Comprehension-inference" knowledge development, however, typically requires some assimilation and integration of information, and is better assessed by test questions or tasks that require students to make some low-level inferences based on their verbatim knowledge. Active (by the ICAP definition) activities are sufficient instructional strategies to develop comprehension-inference knowledge. "Knowledge-inference" development occurs when students can modify and improve their mental models at a structural level, and its assessment measures must tap into students' deep, conceptual understanding of a topic. The best kinds of instructional strategies for knowledge-inference development are most likely, at minimum, constructively engaging.

With regard to the shallow learning outcomes that were obtained from my study, the No Prep-Active instructional strategy was designed to at least engage students actively. The shallow learning T/F test was designed to measure verbatim and comprehension knowledge. Because these knowledge categories require only retrieval of fact-based knowledge and making low-level inferences, an Active strategy should be sufficient to improve learning. Thus, students did not necessarily need deep understanding of the concepts to do relatively well on the T/F test, which helps explain the lack of significant differences between conditions for the shallow assessment. The small differences in means across groups may have been due to sampling error. The main point here is that an Active task should improve shallow learning just as well as a Constructive task, which is supported by the lack of differences found in these results.

For the deep learning outcomes, the No Prep-Active strategy was expected to produce the lowest scores and the results obtained supported this hypothesis. The deep learning measure (i.e. the posttest prediction tasks) targeted structural knowledge (Chi et al.'s, 1994, knowledge-inference category), which is best developed by engaging students at least constructively in the learning task. As discussed in the exploratory analyses on the engagement of students in the Prep conditions, it makes sense that the Prep-Active instructional strategy, which was more likely constructively engaging, would produce better deep learning outcomes compared to the No Prep-Active strategy. A study by Gokhale (1995) also provides evidence that constructively engaging collaborative tasks (those that promote "critical thinking") lead to improved deep knowledge outcomes, but not necessarily improved shallow knowledge outcomes.

In other words, the posttest prediction tasks were likely extremely difficult for students who did not have structurally accurate and fairly complete mental models of the memory concepts, whereas the shallow T/F questions were not dependent on this depth of knowledge. Thus, when students engaged merely actively (and not constructively) in the learning tasks, they may not have been able to sufficiently develop mental models of the concepts that were needed to successfully complete the posttest prediction tasks. This shows the usefulness of distinguishing between deep and shallow learning, and that different kinds of assessments must be utilized to evaluate student knowledge of differing depths. The shallow learning test was not sensitive enough to detect differences in learning because of the way the instructional interventions affected knowledge at the structural mental model level. Only the deep learning posttest tasks were adequate assessments to detect differences in outcomes.

Further Interpretations of Findings

ICAP as a Tool for Instructional Design

To date, the ICAP framework has been shown to be a valuable tool for categorizing student cognitive engagement in learning activities (Chi, 2009; Fonseca & Chi, 2011; Menekse et al., in press), but has not yet been proven as reliable a tool for designing learning tasks. In particular, the Interactive level of the framework has many caveats about what it means to engage in an authentically Interactive way. For instance, Chi (2009, in press) claims that engagement should only be categorized as Interactive when both students are at minimum engaging constructively and there is evidence that the collaboration is encouraging co-construction of knowledge. Admittedly, I applied the framework outside of its original intent for design by using it more simply to encourage specific behaviors that aligned to cognitive engagement. For example, the Active task that I developed specifically asked students to “search and select” from a provided list of terms and descriptions. This is a fairly clear-cut example of Active instructions, but the original framework considers only solo engagement to be Active. I have used it in a collaborative situation, thus, it can be argued that this is not a truly Active activity.

What is interesting to note is the difference between the Prep and No Prep conditions for the Active tasks in particular. Within the No Prep condition, it seems that the task did engage students actively, even though they were collaborating rather than working alone. The addition of preparation before collaborating seemed to trigger more constructive engagement from students. I discuss some further interpretations below.

First, the generic instruction to "work with a partner" in class may be an effective strategy to help students learn shallow, surface-feature aspects of concepts. A basic example is two students drilling each other to learn the definitions of new terminology. It may simply be the repetition of the drilling that is reinforcing memorization of the information, which may be more natural and enjoyable to do with a partner, rather than alone (as shown in the survey data from my work, students tend to enjoy collaboration and Active-type activities). In this case, discussion *per se* may not be related to the learning outcomes (recall Gadgil and Nokes-Malach's work that showed a collaborative inhibition effect for superficial knowledge, 2010, 2011). Because this generic form of collaboration occurs in classrooms at many grade levels, it is important to understand the distinction between having a shallow learning or a deep learning goal. If collaborative activities are intended to help students acquire deep knowledge, one must consider other factors with regard to instructional design. Considering how well the activity might naturally elicit meaningful discussion (Cohen, 1994; Engle & Conant, 2002), whether students get a chance to individually engage with the material (Van Boxtel et al., 2000a), how much and what kind of guidance to use (Walker et al., 2011), etc. would be recommended, since these are factors that have been shown to affect how deeply students engage in a collaborative task and discussion (Janssen, Kirschner, et al., 2010).

Second, with regard to the ICAP framework, the question of what makes collaborative activities truly Interactive still remains. Including constructive-type instructions in the activities may not necessarily be needed. Chi and colleagues (Chi, 2009; Menekse et al., in press) have concluded that to be Interactive, both partners must be constructive, but the way that this translates to instructional design is left unclear. The students in my Prep-Active condition showed some evidence of actually engaging constructively, as demonstrated by the Active activity worksheet analyses. The students in the Prep-Active condition seemed to discuss the content more deeply (evidenced by a small increase in disagreements during discussion) and provided evidence of increased effort compared to the No Prep-Active students. This would suggest that other factors (such as preparation) are key in promoting truly Interactive engagement in learning tasks.

A better use for the ICAP framework for instructional design in its current state may be for solo activities. Again, the learning goals for students (and the assessments that match those goals) are critical for designing instructional activities. If students need only to memorize facts or terminology, Passive activities might suffice. Making a Passive activity Active, might help reinforce shallow knowledge (for example, by having students make flash cards that list terms and definitions verbatim). Constructively designed solo activities should probably be reserved for learning goals to improve students' deep knowledge. Typically, these kinds of activities take more effort to design and their assessments are either more difficult to create or the students responses are more difficult to evaluate. Regarding Interactive activities, there is not yet enough evidence for how to use the ICAP framework for instructional design.

PFL and "Readiness"

One aspect of the PFL work focuses on "dynamic assessments" (Bransford & Schwartz, 1999; Schwartz et al., 2007). These measure "readiness" to learn, rather than correctness of knowledge, and are often assessments of the work by students during increments of the entire learning process. In my study, the students' activity worksheets served the purpose of a dynamic assessment, which is part of the reason that I chose to evaluate students' effort and engagement over accuracy of the information they wrote. Assessing these activities in this way supports the interpretation that individual cognitive preparation may help to develop an enhanced readiness for learning from discussion. This "enhanced readiness" may have prompted students even in the Active condition to engage in the task more constructively in an individual sense, and/or may have led to better quality dialogue, which then improved learning. The specific way(s) that readiness was enhanced is left unclear, as my analyses were not sufficient to address this.

Because the model of the PFL paradigm places certain activities in a "readiness for learning" category, and other activities in a "learning" category, my work introduces new questions about the kinds of tasks, and combinations of tasks, that fit this paradigm. Most of the PFL work that has used collaboration in the instructional activities has positioned it the readiness phase, while the "learning" phase has typically been an individual activity (such as listening to a class lecture). As with Froyd's (2011) suggestion, the findings from my study show that individual preparation partnered with collaboration as the future task also works well to improve students' deep knowledge.

Enjoyment of the Collaborative Tasks

The results from the survey data mirror those found in studies on productive failure. That work shows that when students work on tasks that are open-ended, or ill-structured, they tend to show more evidence of confusion or lack of confidence in their learning and produce more complex patterns of discourse (Janssen, Erkens, et al., 2010; Kapur, 2008; Pathak et al., 2011). Students' work on and dialogues during ill-structured tasks tend to look "messier," and if their performances during learning were to be used as a summative assessment, the students would appear to be "failing" (Kapur & Kinzer, 2007). However, students who engage in ill-structured tasks during a learning phase tend to perform better on both kinds of tasks on transfer or retention posttests, compared to students who engage in well-structured tasks during learning (Kapur, 2008; Kapur & Bielaczyc, 2012).

In my study, the students seemed to like the Active activities better (comparable to well-structured problem-solving tasks). However, the deepest learning occurred for students who constructively engaged in the tasks; students in the Prep-Constructive, Prep-Active (where students may have spontaneously engaged constructively due to the inclusion of a preparation period), and No-Prep Constructive conditions produced higher deep learning scores than students from the No-Prep Active condition. Thus, the tasks that improved learning the most seemed to be those that elicited more complex kinds of engagement (cognitive, motivational, social). Discourse analyses in future work can better inform this interpretation.

Dependency in Data

There are relatively recent discussions about how to address dependency among subjects in learning studies, where either small groups or dyads work together during interventions (Kenny et al., 2006; Kirschner et al., 2011). For my data, the more traditional ANOVA and the multilevel analyses (MLA) essentially led to the same conclusions, thus, the issue of dependency did not "matter." However, since we use statistical testing as a cutoff for making decisions about results, suppose a difference had been between a $p = .06$ and $p = .04$? For an alpha level set at, $a < .05$, that matters. Differences can be more pronounced when outcome scores between partners are positively correlated. As mentioned, since the ICCs computed for my data were not positive, it was not as clear which analytical technique was best to use. I chose the MLA technique because dependency should be determined by the research design.

What does this all mean for studies that assess learning outcomes? One specific question that I find relevant is: If one views learners as having unique representations of knowledge, which can be affected differently by a variety of external forces (text, media, conversations, etc.), is the issue of dependency still relevant? To contrast this with a concrete example, consider my analyses of the differences in students' enjoyment in working with a partner across the collaborative tasks. Some of these scores per group were positively correlated (see Appendix F). In other words, if a student enjoyed the experience, his/her partner also tended to enjoy the experience. Partner scores did seem to be dependent upon one another, which may have affected the statistical outcomes of the group differences. In cases such as this, the multilevel model should be used to assess outcomes basically without question. For learning outcomes, however, if a partner's contribution towards one's learning could be like any other source of input or feedback, then consequently, each partner's knowledge structures may be changed quite differently. Under this perspective, it would not be surprising for outcomes to be uncorrelated.

To follow this line of thought, if dependency is not as "important" to consider for collaborative learning studies as far as how to analyze learning data, one might wonder what the actual purpose of collaboration is. In other words, what would be the point of collaborating if a partner has no influence on a student's outcomes, other than to serve as a feedback mechanism that elicits effective cognitive behaviors from that individual? Why not just find a way to elicit those same behaviors without having to bother with collaboration? Educational technology researchers are, in fact, addressing these questions in work on Intelligent Tutoring Systems and using Artificial Intelligence models to simulate human feedback. (See work in these fields for further information. Researchers S. D'Mello, R. Hausmann, K. VanLehn, and E. Walker are examples.) Considering the effort, cost, and time it takes to create adequate Intelligent computer-based "partners," I argue that collaborative learning with human-to-human interaction is a highly practical way to engage effective learning behaviors in students.[13] How simple it is for the classroom teacher to have students talk to each other about their learning and thinking, and reap the benefits of improving deep knowledge. Thus, research must not only discover the best ways to maximize collaborative learning in the classroom, but also

[13] I must mention that I do not disregard the work on distributed cognition, co-construction, shared-meaning, etc. I wonder how we might examine individual learning within the context of peer discussion, not discounting the effects communication, but embracing them in a way that draws conclusions about best practices in the classroom.

address this issue of dependency, to help assure that "evidence-based" practices are coming from studies that use statistical models yielding valid results.

Implications for Instructional Practices

According to the findings of my study, it appears that one way to maximize the benefits of collaboration on deep learning is to include a preparation task and, in particular, have students work individually on that task and then discuss their work with a partner (also supported by Van Amelsvoort et al., 2007). This task might engage students constructively at the individual stage, or at least promote constructive engagement during the discussion of one another's work. As Froyd (2011) implies, the PFL paradigm may also work for collaborative learning by placing peer discussion as the future learning task after students are sufficiently prepared and ready. It may not be necessary to design collaborative activities to more specifically engage constructive behavior. The preparation may provide enough "fuel" for fruitful discussion and effective learning.

In addition, this work shows that a "preparing-to-interact" structure for collaborative learning activities may make the most efficient use of peer discussion. The students who had the individual preparation time only spent half the amount of time collaborating compared to students in the No Prep conditions. This replicates the finding from Van Boxtel et al.'s (2000a) study, which showed that allowing students to prepare individually improved collaborative learning, compared to additional time to jointly work on the task. Thus, using collaboration as a future learning task, following a preparation task, is an efficient way for students to get the most benefits out of a discussion.

Limitations and Future Work

Addressing Immediate Limitations

One limitation of this work was that there was no true "solo" group with which to compare outcomes. The inclusion of an individual Active and an individual Constructive condition would have provided a way to assess how much of the individual phase of the Prep conditions contributed to the improved outcomes. In other words, it would have been beneficial to know that the improved outcomes that arose from the Prep conditions were not solely attributed to the individual time to work on the tasks. It begs the question: Is it possible that collaboration in the No Prep conditions somehow actually hindered learning for those students? I would argue that this is not of too much concern because collaboration occurred in every condition (in other words, it should have then hindered the Prep students as well). Additionally, I argue that the individual phase allowed for better discussion to occur, however, it was not feasible to do discourse analysis which would have better addressed this claim. Future work should analyze student discourse, as well as include solo conditions for the purpose of comparison to overcome these limitations.

Another possibility for future work would be to examine how "preparing-to-interact" compares to preparation-to-learn-from-lecture, as with most of the PFL work. One question of interest would be: How does peer discussion affect learning compared to lecture when students first are cognitively prepared and "ready" to learn? The findings from such a study could inform how communication, as an isolated factor, affects learning in preparation-type instructional designs.

Finally, it would have been interesting to have obtained data from a retention assessment. The posttest was given two days after the activities, however, this would not be considered by some a true retention test. End-of-semester learning outcomes on these specific topics, for instance, would have been a better measure. But again, it was simply not feasible to obtain this data within the limitations of the instructors' schedules and lesson plans.

Generalization to Other Domains and Populations

Because this work was done in a community college, it should not be generalized to younger student populations. Younger students tend to have more difficulty collaborating effectively without support, such as being scaffolded throughout their collaboration (Slavin, 1992; 1996). Thus, future work could address the question of how a "preparing-to-interact" collaborative activity affects learning of difficult concepts for younger students. With a general national interest in the STEM (Science, Technology, Engineering, Math) subjects for K-12 education, this collaborative learning method could also be tested in other domains. Future study should include the appropriate assessments for measuring for deep learning.

Although this work was tested on a relatively diverse population by ethnicity, gender, and career interests, it was not adequate to assess differences in learning by these factors. Such factors might impact the process of preparing and/or collaborating and communicating in educational settings. Future study is needed to better assess the influences of these factors on "preparing-to-interact" collaborative learning strategies.

Thinking Ahead: Computer-Supported Tools and Individual Differences

As an experimental study, this work did not address how any individual differences in students may have affected their engagement in the tasks, the way that they discussed the learning material, or their learning outcomes. Theoretically, I make the assumption that due to the random assignment to conditions, any individual differences are equalized across groups, thus, the findings are interpreted on a "majority-rules" basis. However, I do not deny that individual differences are important when considering teaching and learning in the classroom. Thus, future work can do more to examine how individual differences due to culture, class, ethnic group, personality, learning style, etc. may factor in to the best ways to design collaborative learning activities.

One field of work that is using computer-supported tools for educational purposes can help to better address questions about how individual differences can inform best-practice models. Some researchers are using computer programs that assess students' individual preferences, learning styles, prior knowledge, etc. and then using that information to personalize learning activities (Lazarinis, Green & Pearson, 2011; Popescu, 2010). We can borrow from the work on personalization of instruction to develop ways to assess "preparing-to-interact" collaborative learning activities by also considering individual differences. In addition, this work can help us incorporate computer-mediated technologies into the activities, as technology devices for learning are becoming more accessible and more common in classrooms at all grade-levels.

REFERENCES

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for Educational and Psychological Testing.* Washington, DC: AERA

Anderson, J.R. (1993). *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Andriessen, J., Baker, M. & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. In J. Andriessen, M. Baker, and D. Suthers (Eds.) *Arguing to Learn: Confronting Cognitions in Computer-Supported Collaborative Learning Environments*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Aronson, E., Stevens, C., Sikes, J., Blaney, N., & Snapp, M. (1978). *The Jigsaw Classroom*. Beverly Hills, CA: Sage Publications.

Asterhan, C.S.C. & Schwarz, B.B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology, 99(3)*, 626-639.

Asterhan, C.S.C. & Schwarz, B.B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialog. *Cognitive Science, 33*, 374-400.

Asterhan, C.S.C. & Schwarz, B.B. (2010). Argumentation and reasoning. In K. Littleton, C. Wood, & J.K. Staarman (Eds.) *International Handbook of Psychology in Education* (pp. 137-176). UK: Emerald Group Publishing Limited.

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences, 9(4)*, 403-436.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences, 12(3)*, 307-359.

Barron, B., Schwartz, D.L., Vye, N.J., Moore, A., Petrosino, A., Zech, L., & Bransford, J.D. (1998). Doing with understanding: Lessons from research on problem-and project-based learning. *Journal of the Learning Sciences, 7(3-4)*, 271-311.

Belenky, D.M. & Nokes, T.J. (2009). Motivation and transfer: The role of achievement goals in preparation for future learning. In N. Taatgen and H. Van Rijn (Eds.) *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, (pp. 1163-1168). Amsterdam, Netherlands: Cognitive Science Society.

Bielaczyc, K., Pirolli, P.L., & Brown, A. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction, 13(2)*, 221-252.

Bjork, E.L. & Storm, B.C. (2011). Retrieval experience as a modifier of future encoding: Another test effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(5)*, 1113-1124.

Bjork, R.A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R.L. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium* (pp. 123-144). Hillsdale, NJ: Erlbaum.

Blumen, H.M., & Rajaram, S. (2008). Influence of re-exposure and retrieval disruption during group collaboration on later individual recall. *Memory, 16*, 231-244.

Bonwell, C.C. & Eison, J.A. (1991). Active learning: Creating excitement in the classroom. *ASHEERIC Higher Education Report No. 1*. Washington, CD: George Washington University.

Bransford, J.D. & Johnson, M.K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717-726.

Bransford, J.D. & Schwartz, D.L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24*, 61-100.

Bruner, J.S. (1961). The act of discovery. *Harvard Educational Review*, 31(1), 21-32.

Carter, K. & Seifert, C.M. (2013). *Learn Psychology*. Burlington, MA: Jones & Bartlett Learning, LLC.

Chi, M.T.H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology* (pp. 161-238). Mahwah, NJ: Erlbaum.

Chi, M.T.H. (2009). Active-Constructive-Interactive: A conceptual framework of differentiating learning activities. *Topics in Cognitive Science, 1(1)*, 73-105.

Chi, M.T.H. (in press). Learning from observing experts. To appear in J.J. Staszewski (Ed.), Expertise and Skill Acquisition: The Impact of William G. Chase. New York, NY: Psychology Press.

Chi, M.T.H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Chi, M.T.H., de Leeuw, N., Chiu, M.H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439-47.

Chi, M.T.H., Roscoe, R., Slotta, J., Roy, M., & Chase, C. (2012). Misconceived causal explanations for “emergent” processes. *Cognitive Science, 36*, 1-61.

Chi, M.T.H., Roy, M. and Hausmann, R.G.M. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science, 32*, 301- 341.

Chi, M.T.H. & VanLehn, K. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, *47(3)*, 177-188.

Chin, C. & Osborne, J. (2010a). Students’ questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. *Journal of Research in Science Teaching, 47(7)*, 883-908.

Chin, C., & Osborne, J. (2010b). Supporting argumentation through students’ questions: Case studies in science classrooms. *Journal of the Learning Sciences, 19(2)*, 230-284.

Clark, H.H. & Brennan, S.E. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (pp. 127-148). Washington, DC: American Psychological Association.

Cohen, E.G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64(1)*, 1-35.

Coleman, E.B. (1998). Using explanatory knowledge during collaborative problem solving in science. *Journal of the Learning Sciences, 7(3-4)*, 387-427.

Craig, S.D., Chi, M.T.H. and VanLehn, K. (2009). Improving classroom learning by collaboratively observing human tutoring videos while problem solving. *Journal of Educational Psychology, 101(4)*, 779-789.

Deiglmayr, A. & Spada, H. (2010). Developing adaptive collaborative support: The example of an effective training for collaborative inferences. *Educational Psychology Review, 22*, 103-113.

Dewey, J. (1916). *Democracy and Education: An Introduction to the Philosophy of Education*. New York, NY: Macmillan.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. Kirschner (Ed.) *Three World of CSCL. Can We Support CSCL?*, pp. 61-91. Open Universiteit Nederland, Heerlen.

Dillenbourg, P. & Hong, F. (2008). The mechanics of CSCL macro scripts. *Computer-Supported Collaborative Learning, 3*, 5-23.

Dillenbourg, P. & Tchounikine, P. (2007). Flexibility in macro-scripts for computer-supported collaborative learning. *Journal of Computer Assisted Learning, 23*, 1-13.

Doymus, K., Karacop, A., & Simsek, U. (2010). Effects of jigsaw and animation techniques on students' understanding of concepts and subjects in electrochemistry. *Educational Technology Research and Development, 58(6)*, 671-691.

Duschl, R.A. & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38*, 39-72.

Engle, R.A. & Conant, F.R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction, 20(4)*, 399-483.

Feuerstein, R. (1979). *The Dynamic Assessment of Retarded Performers: The Learning Potential Assessment Device, Theory, Instruments, and Techniques.* Baltimore, MD: University Park Press.

Finley, N.F., Steward, J., & Yarroch, W.L. (1982). Teachers' perceptions of important and difficult science concepts. *Science Education, 66(4)*, 531-538.

Fogel, A. (1993). *Developing Through Relationships: Origins of Communication, Self, and Culture.* Chicago, IL: University of Chicago Press.

Fonseca, B. & Chi, M.T.H. (2011). The self-explanation effect: A constructive learning activity. In R. Mayer & P. Alexander (Eds.), *The Handbook of Research on Learning and Instruction*. Routledge Press.

Froyd, J.E. (2011). Problem-based learning and adaptive expertise. *Proceedings from the 41st ASEE/IEEE Frontiers in Education Conference* (pp. S3B-1-S3B-5). Rapid City, SD: IEEE.

Gadgil, S. & Nokes, T. (2010). Collaborative facilitation through error-detection: A classroom experiment. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2583-2588). Austin, TX: Cognitive Science Society.

Gadgil, S. & Nokes-Malach, T.J. (2011). Overcoming collaborative inhibition through error correction: A classroom experiment. *Applied Cognitive Psychology*, DOI: 10.1002/acp.1843

GE, X. & Land, S.M. (2004). A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educational Technology, Research, and Development, 52(2)*, 5-22.

Gholson, B., Witherspoon, A., Morgan, B., Brittingham, J.K., Coles, R., Graesser, A.C., Sullins, J., & Craig, S.D. (2009). Exploring the deep-level reasoning questions effect during vicarious learning among eighth to eleventh graders in the domains of computer literacy and Newtonian physics. *Instructional Science, 37*, 487-493.

Gokhale, A.A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education, 7(1)*, 22-30.

Graesser, A.C. & Person, N.K. (1994). Question-asking during tutoring. *American Educational Research Journal, 31*, 104-137.

Greeno, J.G., Collins, A.M., & Resnick, L.B. (1996). Cognition and learning. In D. Berliner and R. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 15-46). New York, NY: MacMillan.

Haugwitz, M., Nesbit, J.C., & Sandmann, A. (2010). Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learning and Individual Differences, 20*, 536-543.

Hausmann, R.G.M. (2006). Why do elaborative dialogs lead to effective problem solving and deep learning? In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Meeting of the Cognitive Science Society* (pp. 1465-1469). Vancouver, B.C.: Sheridan Printing.

Hausmann, R.G.M. & Chi, M.T.H. (in preparation). The impact of elaborative dialogs on an engineering design task.

Hausmann, R.G.M., Nokes, T.J., VanLehn, K., & Van de Sande, B. (2009). Collaborative dialog while studying worked-out examples. In V. Dimitrova, R. Mizoguchi, B. Du Boulay, & A.C. Graesser (Eds.) *Artificial Intelligence in Education*. Amsterdam, Netherlands: IOS Press.

Hausmann, R.G.M., Van de Sande, B., Van de Sande, C., & VanLehn, K. (2008). Productive dialog during collaborative problem solving. *Proceedings of the 8th International Conference of the Learning Sciences*.

Hausmann, R.G.M., Van de Sande, B., & VanLehn, K. (2008). Are self-explaining and coached problem solving more effective when done by pairs of students than alone? In B.C. Love, K. McRae & V.M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2369-2374). New York, NY: Erlbaum.

Hermann, F., Rummel, N. & Spada, H. (2001). Solving the case together: The challenge of net-based interdisciplinary collaboration. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.) *Proceedings of the First European Conference on Computer-Supported Collaborative Learning* (pp. 293-300). Maastricht: McLuhan Institute.

Janssen, J., Erkens, G., Kirschner, P.A., & Kanselaar, G. (2010). Effects of representational guidance during computer-supported collaborative learning. *Instructional Science, 38(1)*, 59-88.

Janssen, J., Kirschner, F., Erkens, G., Kirschner, P.A., & Paas, F. (2010). Making the black box of collaborative learning transparent: Combining process-oriented and cognitive load approaches. *Educational Psychology Review, 22*, 139-154.

Johnson, D.W. & Johnson, R.T. (1992). *Positive Interdependence: Activity Manual and Guide*. Edina, MN: Interaction Book Company.

Johnson, D.W. & Johnson, R.T. (2009). An educational psychology success story: Social interdependence learning theory and cooperative learning. *Educational Researcher, 38(5)*, 365-379.

Johnson, D.W., Johnson, R.T., & Smith, (1991). Cooperative learning: Increasing college faculty instructional productivity. *ASHE-ERIC Higher Education Report No. 4*. Washington, CD: The George Washing University, School of Education and Human Development.

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*, 379- 424.

Kapur, M. & Bielaczyc, K. (2011). Classroom-based experiments in productive failure. In L. Carlson, C. Hoelscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2812-2817). Austin, TX: Cognitive Science Society.

Kapur, M. & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences, 21(1)*, 45-83.

Kapur, M. & Kinzer, C.K. (2007). Examining the effect of problem type in a synchronous computer-supported collaborative learning (CSCL) environment. *Educational Technology Research and Development, 55(5)*, 439-459.

Karakostas, A. & Demetriadis, S. (2011). Enhancing collaborative learning through dynamic forms of support: The impact of an adaptive domain-specific support strategy. *Journal of Computer Assisted Learning, 27(3)*, 243-258.

Kenny, D.A., Kashy, D.A., & Cook, W.L. (2006). *Analysis of Dyadic Data*. New York: Guilford.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal, 27(4)*, 664-687.

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Research Journal, 29(2)*, 303-323.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and explain. *American Educational Research Journal, 31(2)*, 338-368.

King, A. (1999). Discourse patterns for mediating peer learning. In A.M. O'Donnell & A. King (Eds.), *Cognitive Perspectives on Peer Learning* (pp. 87-115). Mahwah, NJ: Lawrence Erlbaum Associates.

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*, 587-599.

Kneser, C. & Ploetzner, R. (2001). Collaboration on the basis of complementary domain knowledge: Observed dialogue structures and their relation to learning success. *Learning and Instruction, 11*, 53-83.

Kumpulainen, K. & Kaartenin, S. (2003). The interpersonal dynamics of collaborative reasoning in peer interactive dyads. *The Journal of Experimental Education, 71(4)*, 333-370.

Latane, B., Williams, K., Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37(6)*, 822-832.

Lazarinis, F., Green, S., & Pearson, E. (2011). Multi-criteria adaptation in a personalized multimedia testing took based on semantic technologies. *Interactive Learning Environments, 19(3)*, 267-283.

Lin, T., Anderson, R.C., Hummel, J.E., Jadallah, M., Miller, B.W., Nguyen-Jahiel, K., Morris, J.A., Kuo, L., Kim, I., Wu, X., & Dong, T. (2012). Children's use of analogy during collaborative reasoning. *Child Development, 83(4)*, 1429-1443.

Lyman, F.T. (1981). The responsive classroom discussion: The inclusion of all students. In A. Anderson (Ed.) *Mainstreaming Digest*, 109-113. College Park, MD: University of Maryland College of Education.

Marx, J.D. & Cummings, K. (2007). Normalized change. *American Journal of Physics, 75(1)*, 87-91.

McTighe, J. & Lyman, F.T. (1988). Cueing thinking in the classroom: The promise of theory-embedded tools. *Educational Leadership, 45(7)*, 18-24.

Menekse, M. (2012). Interactive-Constructive-Active-Passive: The relative effectiveness of differentiated activities on students' learning. (Doctoral dissertation). Retrieved from ProQuest. (3518872).

Menekse, M., Stump, G.S., Krause, S., & Chi, M.T.H. (in press). Differentiated overt learning activities for effective instruction in an engineering classroom. *Journal of Engineering Education*.

Muldner, K., Dygvib, K., Lam, R. and Chi, M.T.H. (2011). Learning by observing tutorial dialogue versus monologue collaboratively or alone. In L. Carlson, C. Hoelscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1340-1345). Austin, TX: Cognitive Science Society.

Muldner, K., Lam, R. and Chi, M.T.H. (in revision). Comparing learning from human tutoring and from observing.

Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education, 10*, 98-129.

Nokes-Malach, T.J., Meade, M.L., & Morrow, D.G. (2012). The effect of expertise on collaborative problem solving. *Thinking & Reasoning, 18(1)*, 32-58.

Nussbaum, E.M. (2005). The effect of goal instructions and need for cognition on interactive argumentation. *Contemporary Educational Psychology, 30*, 286-313.

Nussbaum, E.M., Alvarez, C., McFarlane, A., Gomez, F., Claro, S., & Radovic, D. (2009). Technology as small group face-to-face collaborative scaffolding. *Computers & Education, 52*, 147-153.

O'Donnell, A.M. (1999). Structuring dyadic interaction through scripted cooperation. In A.M. O'Donnell & A. King (Eds.), *Cognitive Perspectives on Peer Learning* (pp. 179-196). Mahwah, NJ: Lawrence Erlbaum Associates.

O'Donnell, A.M. & King, A. (1999). *Cognitive Perspectives on Peer Learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Palinscar, A.S. & Brown, A.L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1(2)*, 117-175.

Pathak, S.A., Kim, B., Jacobson, M.J., & Zhang, B. (2011). Learning the physics of electricity: A qualitative analysis of collaborative processes involved in productive failure. *Computer-Supported Collaborative Learning, 6*, 57-73.

Piaget, J. (1977) *The Development of Thought: Equilibration of Cognitive Structures*. Oxford, England: Viking.

Piaget, J. (1985). *The Equilibration of Cognitive Structures: The Central Problem of Intellectual Development*. Chicago, IL: University of Chicago Press.

Ploetzner, R., Fehse, E., Spada, H., & Kneser, C. (1999). Learning to relate qualitative and quantitative problem representations in a model-based setting for collaborative problem solving. *The Journal of the Learning Sciences, 8*, 177-214.

Popescu, E. (2010). Adaptation provisioning with respect to learning styles in a Web-based educational system: An experimental study. *Journal of Computer Assisted Learning, 26*, 243-257.

Pozzi, F. (2010). Using Jigsaw and Case Study for supporting online collaborative learning. *Computers & Education, 55*, 67-75.

Rajaram, S. (2011). Collaboration both hurts and helps memory: A cognitive perspective. *Current Directions in Psychological Science, 20(2)*, 76-81.

Roediger, H.L. & Marsh, E.J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(5)*, 1155-1159.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences, 2(3)*, 235-276.

Roscoe, R. & Chi, M.T.H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors explanations and questions. *Review of Educational Research 77(4)*, 534-574.

Rummel, N. & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences, 14(2)*, 201- 241.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *Computer-Supported Collaborative Learning, 4*, 69-92.

Schwartz, D.L. (1995). The emergence of abstract representations in dyad problem solving. *The Journal of the Learning Sciences, 4(3),* 321-354.

Schwartz, D.L. & Bransford, J.D. (1998). A time for telling. *Cognition and Instruction, 16(4)*, 475-522.

Schwartz, D.L. & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22(2)*, 129-184.

Schwartz, D.L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. Lovett & P. Shah (Eds.), *Carnegie Symposium on Cognition*, 319-344. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Schwartz, D.L., Varma, S., & Martin, L. (in press). Dynamic transfer and innovation. To appear in S. Vosniadou, *Handbook of Conceptual Change*.

Schwarz, B.B. & Linchevski, L. (2007). The role of task design and argumentation in cognitive development during peer interaction: The case of proportional reasoning. *Learning and Instruction, 17*, 510-531.

Schwarz, B.B., Newman, Y., & Biezuner, S. (2000). Two wrongs make a right… if they argue together! *Cognition and Instruction, 18(4)*, 461-494.

Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science 26*, 469-501.

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

Slavin, R.E. (1992). When and why does cooperative learning increase achievement? Theoretical and empirical perspectives. In R. Hertz- Lazarowitz & N. Miller (Eds.), *Interaction in Cooperative Groups: The Theoretical Anatomy of Group Learning* (pp. 145-173). New York: Cambridge University Press.

Slavin, R.E. (1996). Research for the future: Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*, 43-69.

Smith, M.K., Wood, W.B., Adams, W.K., Wieman, C., Knight, J.K., Guild, N., & Su, T.T., (2009). Why peer discussion improves student performance on in-class concept questions. *Science, 323*, 122-124.

Taylor, K.K. (1983). Can college students summarize? *Journal of Reading, 26(6)*, 524-528.

Teasley, S.D. & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S.P. Lajoie & S.D. Derry (Eds.), *Computers as Cognitive Tools* (pp. 229-258). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Uesaka, Y. & Manalo, E. (2011). The effects of peer communication with diagrams on students' math word problem solving processes and outcomes. In L. Carlson, C. Hoelscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 312-317). Austin, TX: Cognitive Science Society.

Van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (2007). Representational tools in computer-supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams. *Journal of the Learning Sciences, 16(4)*, 485-521.

Van Boxtel, C., Van der Linden, J., & Kanselaar, G. (2000a). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction, 10*, 311-330.

Van Boxtel, C., Van der Linden, J. & Kanselaar, G. (2000b). The use of textbooks as a tool during collaborative physics learning. *The Journal of Experimental Education, 69(1)*, 57-76.

Veerman, A., Andriessen, J., & Kanselaar, G. (2002). Collaborative argumentation in academic education. *Instructional Science, 30*, 155-186.

Volet, S., Summers, M., & Thurman, J. (2009). High-level co-regulation in collaborative learning: How does it emerge and how is it sustained? *Learning and Instruction, 19*, 128-143.

Vygotsky, L.S. (1978). *Mind and society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walker, E., Rummel, N., & Koedinger, K.R., (2011). Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *Computer-Supported Collaborative Learning, 6*, 279-306.

Webb, N.M. & Mastergeorge (2003). Promoting effective helping behavior in peer-directed groups. *International Journal of Educational Research, 39*, 73-97.

Webb, N.M., Nemer, K.M., & Ing, M. (2006). Small-group reflections: Parallels between teacher discourse and student behavior in peer-directed groups. *Journal of the Learning Sciences, 15(1)*, 63-119.

Wiedmann, M., Leach, R.C., Rummel, N., & Wiley, J. (2012). Does group composition affect learning by invention? *Instructional Science*. DOI: 10.1007/s11251-012-9204-y

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.

Wong, R.M.F., Lawson, M.J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction, 12*, 233-262.

Yetter, G., Gutkin, T.B., Saunders, A., Galloway, A.M., Sobandsky, R.R., & Song, S.Y. (2006). Unstructured collaboration versus individual practice for complex problem solving: A cautionary tale. *The Journal of Experimental Education, 74(2)*, 137-159.

APPENDIX A

LEARNING ACTIVITY PACKETS

No Prep-Active Packet (10 pages)

NP-A

Information to Begin!

This packet contains the materials you will need to complete today's activity. There are several concepts related to memory that this activity should help you to learn. Please read the instructions for each part carefully! You will complete this activity with a partner. Please turn in ONE packet and write both you and your partner's research ID numbers on the packet.

Please write your research activity ID number on all areas indicated!

Tips for Collaborating:

Work with your partner to contribute equally to your discussion. Listen to your partner's ideas and feedback. Share your own ideas and give your partner feedback. Ask each other questions and consider one another's perspectives. You should BOTH be talking and sharing! If you finish before the time allotted, review your work and make sure that you both agree that it is completed to your satisfaction.

Read this background information about memory before you start the activity!

Memory is a process that involves 3 kinds of actions:

1) *getting information "in" to memory, referred to as* ***Encoding***
2) *keeping it, referred to as* ***Maintaining*** *and* ***Organizing*** or ***Storage***
3) *and getting it back "out," referred to as* ***Retrieving****.*

People can encode information in several ways, maintain and organize it differently, and multiple factors can affect how people retrieve information from memory. Remembering can be a tricky thing! Various circumstances affect what we remember, why we remember some information well and some not-so-well, and why we forget information.

Why do we remember what we remember?

Please read through the descriptions of the concepts and principles about memory below (terms are **bolded**). Both you and your partner should look over these terms. Then read through the two experiments and the results of their data on the following pages. Further instructions are provided there. You can keep this sheet and use it for the later activities.

People tend to **remember**...

- information that they can connect to information they already know.
  - **Elaboration**
- information or inferences that they create or construct themselves.
  - **Generation effect**
- information that they can "fit" into their existing knowledge structures or patterns.
  - **Schemas**
- things that hinder the completion of a goal.
  – **Obstacle recall**
- the information from the beginning and the end of an event.
  – **Serial position effect**
- the general overview of an event, experience, or information, etc.
  – **Gist**

People tend to NOT remember or **forget** information when...

- new information disrupts or overtakes existing memories.
  – **Interference**
- it is given little attention at the time of encoding.
  – **Encoding failure**

Read through the description of the experiment and its results below[1]. Afterwards, you will connect main ideas from the results with concepts and principles of memory.

Experiment 1: Remembering a Doctor's Visit

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 8 participants to read several paragraphs about different topics. They then tested them on only one of the paragraphs (below).

The Doctor Visit

John checked in with the doctor's receptionist. While he waited he read magazines. The nurse called his name. John undressed. John talked to the nurse. The doctor came in to the examination room. The doctor was very friendly. The doctor prescribed some pills for John. John forgot his wallet. John left the doctor's office.

The participants' task was to rate the following 12 sentences as to how sure they were that they had actually read the sentence on a scale of **1** (not at all) to **7** (very). They were not told which sentences had appeared in the paragraph.

For your convenience, the sentences that were repeated from the paragraph are highlighted. The average scores for participants' ratings are included on the left.

Sentences	**Ratings**
A) John checked in with the doctor's receptionist.	6.63
B) John sat down.	5.88
C) While he waited he read magazines.	6.50
D) John followed the nurse.	2.75
E) John undressed.	4.38
F) John talked to the nurse.	4.00
G) The nurse tested John in the examination room.	5.50
H) The doctor examined John.	5.50
I) The doctor prescribed some pills for John.	4.00
J) John forgot his wallet.	6.75
K) John made another appointment.	2.50
L) John left the doctor's office.	6.75

[1] The activity example is slightly modified from activities Schwartz & Bransford (1998) used in several experiments.

NP-A

Student ID: ____________
Student ID: ____________

Read each main point below. For each one, work with your partner to write in the memory concepts or principles that are addressed. You may write in more than one term for each point. You may also use the same terms for multiple statements. Many of the concepts/principles are related and can overlap! Use the information on page 2 for help. See if you and your partner can reach a consensus about which concepts/principles each point addresses. (If you are completely split on an answer, please indicate so.) You will turn in ONE worksheet for your shared answers.

Main Points	Concept(s)
People were very sure that they read sentences A, C, and L. This is because people tend to remember the first and last details of an event pretty well.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	
But, many people were also fairly sure that they read, "John sat down," when that sentence was not actually included in the paragraph. This may be because people tend to remember the overall essence of an event rather than specific details.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	
Most people don't stand at a doctor's office reading magazines. So, you might think "John sat down," is an obvious part of the scenario. But this is only obvious because most of us have had experiences at the doctor's office. We assume this scenario matches with our prior experiences.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	

Main Points	Concept(s)
We also tend to remember our own inferences better than actual details.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	
People were pretty sure that they read sentences G and H, but these were also not included in the original paragraph. Again, being typical sorts of things that happen at a doctor visit, people may have "attached" their own inferences to their general knowledge structure of a doctor visit.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	
Sentences E, F, and I were those that people weren't completely sure, nor completely unsure that they had actually read. Since they had read many paragraphs, these statements may have been more easily forgotten simply because they were not encoded deeply.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	
Or, information from the other paragraphs that people read may have intruded on the memories of these statements.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	

Main Points	Concept(s)
Finally, most everyone was very sure that they read that John forgot his wallet. People tend to remember information that relates to hindering a goal.	

My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ______________________________

____

____

Continue to the next page for Experiment 2!

Read through the description of another experiment and its results below[2]. Afterwards, you will connect main ideas from the results with concepts and principles of memory.

Experiment 2: Going to the Grocery Store

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 8 participants to read several paragraphs about different topics. They then tested them on only one of the paragraphs (below).

The Grocery Store

Sam entered the store. He grabbed a grocery cart. He picked up fruit and vegetables first. Sam then headed to the dairy area. He got the usual things. A worker asked if he needed any help. The worker was stocking items on shelves. Sam went to check out. No one was at the checkout stands when he got there. Sam waited until someone arrived and then paid.

The participants' task was to rate the following 12 sentences as to how sure they were that they had actually read the sentence on a scale of **1** (not at all) to **7** (very). They were not told which sentences had appeared in the paragraph.

For your convenience, the sentences that were repeated from the paragraph are highlighted. The average scores for participants' ratings are included on the left.

Sentences	**Ratings**
A) Sam entered the store.	6.63
B) He looked for a cart.	5.88
C) He grabbed a grocery cart.	6.50
D) He picked up apples and bananas.	2.75
E) Sam then headed to the dairy area.	4.38
F) He got the usual things.	4.00
G) The store manager greeted Sam.	5.50
H) He was doing inventory.	5.50
I) Sam went to check out.	4.00
J) No one was at the checkout stands when he got there.	6.75
K) He looked at some magazines.	2.50
L) Sam waited until someone arrived and then paid.	6.75

[2] This activity example was created for the purpose of this activity, but is based on activities used in experiments by Schwartz & Bransford, 1998.

NP-A

Student ID: ____________
Student ID: ____________

Read each main point below. For each one, work with your partner to write in the memory concepts or principles that are addressed. You may write in more than one term for each point. You may also use the same terms for multiple statements. Many of the concepts/principles are related and can overlap! Use the information on page 2 for help. See if you and your partner can reach a consensus about which concepts/principles each point addresses. (If you are completely split on an answer, please indicate so.) You will turn in ONE worksheet for your shared answers.

Main Points	**Concept(s)**
People were very sure that they read sentences A, C, and L. This is because people tend to remember the first and last details of an event pretty well.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain) ____________	
But, many people were also fairly sure that they read, "He looked for a cart," when that sentence was not actually included in the paragraph. This may be because people tend to remember the overall essence of an event rather than specific details.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain) ____________	
You might think, "He looked for a cart," is an obvious part of the scenario. Most of us look for a cart that you can push easily or that isn't dirty. But we assume "looking for a cart" is obvious because we have had experiences at the grocery store. We assume this scenario matches with our prior experiences.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain) ____________	

Main Points	Concept(s)
We also tend to remember our own inferences better than actual details.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	
People were pretty sure that they read sentences G and H, but these were also not included in the original paragraph. Again, being typical sorts of things that happen at a grocery store, people may have "attached" their own inferences to their general knowledge structure of getting groceries.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	
Sentences E, F, and I were those that people weren't completely sure, nor completely unsure that they had actually read. Since they had read many paragraphs, these statements may have been more easily forgotten simply because they were not encoded deeply.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	
Or, information from the other paragraphs that people read may have intruded on the memories of these statements.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	

Main Points	Concept(s)
Finally, most everyone was very sure that they read that no one was at the checkout stands. People tend to remember information that relates to hindering a goal.	
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.) ☐ Yes ☐ No (Explain)	

Stop! You have completed the activity.

If you still have time, go over your answers with your partner one more time.

No Prep-Constructive Packet (8 pages)

Information to Begin!

This packet contains the materials you will need to complete today's activity. There are several concepts related to memory that this activity should help you to learn. Please read the instructions for each part carefully! You will complete this activity with a partner. Please turn in ONE packet and write both you and your partner's research ID numbers on the packet.

Please write your research activity ID number on all areas indicated!

Tips for Collaborating:

Work with your partner to contribute equally to your discussion. Listen to your partner's ideas and feedback. Share your own ideas and give your partner feedback. Ask each other questions and consider one another's perspectives. You should BOTH be talking and sharing! If you finish before the time allotted, review your collaborative work and make sure that you both agree that it is completed to your satisfaction.

Read this background information about memory before you start the activity!

Memory is a process that involves 3 kinds of actions:

1) *getting information "in" to memory, referred to as* ***Encoding***
2) *keeping it, referred to as* ***Maintaining*** *and* ***Organizing*** or ***Storage***
3) *and getting it back "out," referred to as* ***Retrieving****.*

People can encode information in several ways, maintain and organize it differently, and multiple factors can affect how people retrieve information from memory. Remembering can be a tricky thing! Various circumstances affect what we remember, why we remember some information well and some not-so-well, and why we forget information.

Why do we remember what we remember?

Read through the description of the experiment and its results below[1]. You and your partner will analyze the data from this experiment to try to invent some general principles and ideas about how memory works. Further instructions are provided after you study the experiment.

Experiment: Remembering a Doctor's Visit

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 8 participants to read several paragraphs about different topics. They then tested them on only one of the paragraphs (below).

The Doctor Visit

John checked in with the doctor's receptionist. While he waited he read magazines. The nurse called his name. John undressed. John talked to the nurse. The doctor came in to the examination room. The doctor was very friendly. The doctor prescribed some pills for John. John forgot his wallet. John left the doctor's office.

The participants' task was to rate the following 12 sentences as to how sure they were that they had actually read the sentence on a scale of **1** (not at all) to **7** (very). They were not told which sentences had appeared in the paragraph.

For your convenience, the sentences that were repeated from the paragraph are highlighted.

A) John checked in with the doctor's receptionist.
B) John sat down.
C) While he waited he read magazines.
D) John followed the nurse.
E) John undressed.
F) John talked to the nurse.
G) The nurse tested John in the examination room.
H) The doctor examined John.
I) The doctor prescribed some pills for John.
J) John forgot his wallet.
K) John made another appointment.
L) John left the doctor's office.

A Data Table from this study is included on the next page.

[1] The activity example is slightly modified from activities Schwartz & Bransford (1998) used in several experiments.

Participant Ratings of How Sure They Were That They Had Read Each Sentence (1 = not at all to 7 = very)

ID*	A	B	C	D	E	F	G	H	I	J	K	L
1	7	7	6	3	4	3	6	1	4	7	2	7
2	7	5	6	1	5	4	7	3	3	7	1	7
3	6	7	7	4	3	5	5	4	5	7	3	6
4	7	6	7	3	6	4	4	3	2	6	4	7
5	7	6	5	3	4	3	7	2	6	7	4	6
6	7	4	7	2	5	5	6	3	3	7	2	7
7	7	7	7	2	3	3	5	5	5	7	1	7
8	5	5	7	4	5	5	4	2	4	6	3	7
Average	*6.63*	*5.88*	*6.50*	*2.75*	*4.83*	*4.00*	*5.50*	*5.50*	*4.00*	*6.75*	*2.50*	*6.75*

* Each participant was assigned a generic ID number.

A) John checked in with the doctor's receptionist.
B) John sat down.
C) While he waited he read magazines.
D) John followed the nurse.
E) John undressed.
F) John talked to the nurse.
G) The nurse tested John in the examination room.
H) The doctor examined John.
I) The doctor prescribed some pills for John.
J) John forgot his wallet.
K) John made another appointment.
L) John left the doctor's office.

Summary of Results

You can see that sentences A, C, J, and L were rated very high compared to sentences E, F, and I, even though they were all included in the original paragraph.
Sentences B, G, and H were also rated high, but did not appear in the paragraph.
Sentences D and K had the lowest ratings.
(*Remember, a high rating means that people were very sure that they remembered reading the sentence.*)

What do you make of this data? Interpret what this data means regarding how memory works. Remember, a high rating means that people were very sure that they had read the statement and a low rating means that they were NOT sure that they had read the statement. The highlighted statements actually appeared in the paragraph, while the other statements did not appear.

Think about the following questions:

- Do you see any patterns in the data?
- What kinds of statements did people remember accurately?
- Why would people remember certain kinds of information, but not other kinds?
- What kinds of statements did people believe that they had read, but did not actually read?
- Why would people falsely remember information?

Look at the individual scores, the average scores, and the summary results to try to understand the ideas behind the data! Use the table on the next page to write out your ideas about memory.

NP-C

Student ID: ____________
Student ID: ____________

How do you think memory works? You and your partner are to try to come up with at least 8 ideas about how and why we remember certain kinds of information, but not other kinds. You can think of them as *concepts* or *principles* about memory. Use the data from the experiment to support your ideas. Do NOT worry about "right" and "wrong" answers! The full chart is on the next few pages and includes columns for what people CORRECTLY remember, INCORRECTLY remember, and FORGET.
Try to reach consensus for those you do not initially agree on. (If you are completely split on an idea, please indicate so.) You and your partner will turn in ONE collaborative worksheet.

People tend to CORRECTLY remember…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.
1.		
2.		
3.		
4.		
Table continues to the next page!		

People tend to CORRECTLY remember…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.
5.		
6.		
7.		
8.		
People tend to INCORRECTLY remember…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.
9.		
Table continues to the next page!		

People tend to INCORRECTLY remember…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.
10.		
11.		
12.		
13		
People tend to FORGET…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.
14.		

Table continues to the next page!

People tend to FORGET…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.
15.		
16.		
17.		
18.		

Stop! You have completed the activity. If you still have time, go over your answers with your partner one more time.

Prep-Active Packet (11 pages)

P-A

Information to Begin!

This packet contains the materials you will need to complete today's activity. There are several concepts related to memory that this activity should help you to learn. Please read the instructions for each part carefully! You will first complete a worksheet that you will do alone. Please do this to the best of your ability! You will be using this worksheet to participate in a partner activity. After completion of this individual worksheet, you then will be partnered with another student to complete the collaborative part of the activity. Your instructor will give you a 2-minute warning to wrap up each part of the activity.

Please write your research activity ID number on all areas indicated!

Tips for Collaborating:

Work with your partner to contribute equally to your discussion. Listen to your partner's ideas and feedback. Share your own ideas and give your partner feedback. Ask each other questions and consider one another's perspectives. You should BOTH be talking and sharing! If you finish before the time allotted, review your collaborative work and make sure that you both agree that it is completed to your satisfaction.

Read this background information about memory before you start the activity!

Memory is a process that involves 3 kinds of actions:

1) *getting information "in" to memory, referred to as* ***Encoding***
2) *keeping it, referred to as* ***Maintaining*** *and* ***Organizing*** or ***Storage***
3) *and getting it back "out," referred to as* ***Retrieving****.*

People can encode information in several ways, maintain and organize it differently, and multiple factors can affect how people retrieve information from memory. Remembering can be a tricky thing! Various circumstances affect what we remember, why we remember some information well and some not-so-well, and why we forget information.

Individual Worksheet

Why do we remember what we remember?

Please read through the descriptions of the concepts and principles about memory below (terms are **bolded**). After going through these, you will then read through two experiments and the results of their data on the following pages. Further instructions are provided there. You can keep this sheet and use it for the later activities.

People tend to **remember**...

- information that they can connect to information they already know.
  - **Elaboration**
- information or inferences that they create or construct themselves.
  - **Generation effect**
- information that they can "fit" into their existing knowledge structures or patterns.
  - **Schemas**
- things that hinder the completion of a goal.
  – **Obstacle recall**
- the information from the beginning and the end of an event.
  – **Serial position effect**
- the general overview of an event, experience, or information, etc.
  – **Gist**

People tend to NOT remember or **forget** information when...

- new information disrupts or overtakes existing memories.
  – **Interference**
- it is given little attention at the time of encoding.
  – **Encoding failure**

Read through the description of the experiment and its results below[1]. Afterwards, you will connect main ideas from the results with concepts and principles of memory.

Experiment 1: Remembering a Doctor's Visit

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 8 participants to read several paragraphs about different topics. They then tested them on only one of the paragraphs (below).

The Doctor Visit

John checked in with the doctor's receptionist. While he waited he read magazines. The nurse called his name. John undressed. John talked to the nurse. The doctor came in to the examination room. The doctor was very friendly. The doctor prescribed some pills for John. John forgot his wallet. John left the doctor's office.

The participants' task was to rate the following 12 sentences as to how sure they were that they had actually read the sentence on a scale of **1** (not at all) to **7** (very). They were not told which sentences had appeared in the paragraph.

For your convenience, the sentences that were repeated from the paragraph are highlighted. The average scores for participants' ratings are included on the left.

Sentences	**Ratings**
A) John checked in with the doctor's receptionist.	6.63
B) John sat down.	5.88
C) While he waited he read magazines.	6.50
D) John followed the nurse.	2.75
E) John undressed.	4.38
F) John talked to the nurse.	4.00
G) The nurse tested John in the examination room.	5.50
H) The doctor examined John.	5.50
I) The doctor prescribed some pills for John.	4.00
J) John forgot his wallet.	6.75
K) John made another appointment.	2.50
L) John left the doctor's office.	6.75

[1] The activity example is slightly modified from activities Schwartz & Bransford (1998) used in several experiments.

Student ID: ___________

Read each main point below. For each one, write in the memory concepts or principles that are addressed. You may write in more than one term for each point. You may also use the same terms for multiple statements. Many of the concepts/principles are related and can overlap! Use the information on page 2 for help.

Main Points	**Concept(s)**
People were very sure that they read sentences A, C, and L. This is because people tend to remember the first and last details of an event pretty well.	
But, many people were also fairly sure that they read, "John sat down," when that sentence was not actually included in the paragraph. This may be because people tend to remember the overall essence of an event rather than specific details.	
Most people don't stand at a doctor's office reading magazines. So, you might think "John sat down," is an obvious part of the scenario. But this is only obvious because most of us have had experiences at the doctor's office. We assume this scenario matches with our prior experiences.	
We also tend to remember our own inferences better than actual details.	
People were pretty sure that they read sentences G and H, but these were also not included in the original paragraph. Again, being typical sorts of things that happen at a doctor visit, people may have "attached" their own inferences to their general knowledge structure of a doctor visit.	
Sentences E, F, and I were those that people weren't completely sure, nor completely unsure that they had actually read. Since they had read many paragraphs, these statements may have been more easily forgotten simply because they were not encoded deeply.	
Or, information from the other paragraphs that people read may have intruded on the memories of these statements.	
Finally, most everyone was very sure that they read that John forgot his wallet. People tend to remember information that relates to hindering a goal.	

Read through the description of another experiment and its results below[2]. Afterwards, you will connect main ideas from the results with concepts and principles of memory.

Experiment 2: Going to the Grocery Store

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 8 participants to read several paragraphs about different topics. They then tested them on only one of the paragraphs (below).

The Grocery Store

Sam entered the store. He grabbed a grocery cart. He picked up fruit and vegetables first. Sam then headed to the dairy area. He got the usual things. A worker asked if he needed any help. The worker was stocking items on shelves. Sam went to check out. No one was at the checkout stands when he got there. Sam waited until someone arrived and then paid.

The participants' task was to rate the following 12 sentences as to how sure they were that they had actually read the sentence on a scale of **1** (not at all) to **7** (very). They were not told which sentences had appeared in the paragraph.

For your convenience, the sentences that were repeated from the paragraph are highlighted. The average scores for participants' ratings are included on the left.

Sentences	**Ratings**
A) Sam entered the store.	6.63
B) He looked for a cart.	5.88
C) He grabbed a grocery cart.	6.50
D) He picked up apples and bananas.	2.75
E) Sam then headed to the dairy area.	4.38
F) He got the usual things.	4.00
G) The store manager greeted Sam.	5.50
H) He was doing inventory.	5.50
I) Sam went to check out.	4.00
J) No one was at the checkout stands when he got there.	6.75
K) He looked at some magazines.	2.50
L) Sam waited until someone arrived and then paid.	6.75

[2] This activity example was created for the purpose of this activity, but is based on activities used in experiments by Schwartz & Bransford, 1998.

Read each main point below. For each one, write in the memory concepts or principles that are addressed. You may write in more than one term for each point. You may also use the same terms for multiple statements. Many of the concepts/principles are related and can overlap! Use the information on page 2 for help.

Main Points	Concept(s)
People were very sure that they read sentences A, C, and L. This is because people tend to remember the first and last details of an event pretty well.	
But, many people were also fairly sure that they read, "He looked for a cart," when that sentence was not actually included in the paragraph. This may be because people tend to remember the overall essence of an event rather than specific details.	
You might think, "He looked for a cart," is an obvious part of the scenario. Most of us look for a cart that you can push easily or that isn't dirty. But we assume "looking for a cart" is obvious because we have had experiences at the grocery store. We assume this scenario matches with our prior experiences.	
We also tend to remember our own inferences better than actual details.	
People were pretty sure that they read sentences G and H, but these were also not included in the original paragraph. Again, being typical sorts of things that happen at a grocery store, people may have "attached" their own inferences to their general knowledge structure of getting groceries.	
Sentences E, F, and I were those that people weren't completely sure, nor completely unsure that they had actually read. Since they had read many paragraphs, these statements may have been more easily forgotten simply because they were not encoded deeply.	
Or, information from the other paragraphs that people read may have intruded on the memories of these statements.	
Finally, most everyone was very sure that they read that no one was at the checkout stands. People tend to remember information that relates to hindering a goal.	

Stop here! Wait until your instructor tells you to move on.
If you have time left, go over your answers one more time!

Collaborative Worksheet

Student ID: ___________
Student ID: ___________

Now compare your individual work with a partner. Did you write the same concepts/ principles into the blanks? Were some of your ideas different than your partner's? Compare and contrast your ideas with your partner's ideas and discuss any differences. See if you can reach a consensus about which concepts/principles each point addresses and complete the new sheet below. (If you are completely split on an answer, please indicate so.) You will turn in ONE collaborative worksheet for your shared answers.

Main Points and Memory Concepts

A Doctor Visit:

People were very sure that they read sentences A, C, and L. This is because people tend to remember the first and last details of an event pretty well.
Concept(s): ____
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____

____

____

But, many people were also fairly sure that they read, "John sat down," when that sentence was not actually included in the paragraph. This may be because people tend to remember the overall essence of an event rather than specific details.
Concept(s): ____
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____

____

____

Most people don't stand at a doctor's office reading magazines. So, you might think "John sat down," is an obvious part of the scenario. But this is only obvious because most of us have had experiences at the doctor's office. We assume this scenario matches with our prior experiences.
Concept(s): ____
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____

____

____

We also tend to remember our own inferences better than actual details.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

People were pretty sure that they read sentences G and H, but these were also not included in the original paragraph. Again, being typical sorts of things that happen at a doctor visit, people may have "attached" their own inferences to their general knowledge structure of a doctor visit.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

Sentences E, F, and I were those that people weren't completely sure, nor completely unsure that they had actually read. Since they had read many paragraphs, these statements may have been more easily forgotten simply because they were not encoded deeply.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

Or, information from the other paragraphs that people read may have intruded on the memories of these statements.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

Finally, most everyone was very sure that they read that John forgot his wallet. People tend to remember information that relates to hindering a goal.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

____

____

The Grocery Store:

People were very sure that they read sentences A, C, and L. This is because people tend to remember the first and last details of an event pretty well.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

____

____

But, many people were also fairly sure that they read, "He looked for a cart," when that sentence was not actually included in the paragraph. This may be because people tend to remember the overall essence of an event rather than specific details.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

____

____

You might think, "He looked for a cart," is an obvious part of the scenario. Most of us look for a cart that you can push easily or that isn't dirty. But we assume "looking for a cart" is obvious because we have had experiences at the grocery store. We assume this scenario matches with our prior experiences.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

____

____

We also tend to remember our own inferences better than actual details.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

People were pretty sure that they read sentences G and H, but these were also not included in the original paragraph. Again, being typical sorts of things that happen at a grocery store, people may have "attached" their own inferences to their general knowledge structure of getting groceries.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

Sentences E, F, and I were those that people weren't completely sure, nor completely unsure that they had actually read. Since they had read many paragraphs, these statements may have been more easily forgotten simply because they were not encoded deeply.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

Or, information from the other paragraphs that people read may have intruded on the memories of these statements.
Concept(s): ____________________
My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____________________

Finally, most everyone was very sure that they read that no one was at the checkout stands. People tend to remember information that relates to hindering a goal.
Concept(s): ____

My partner and I fully agreed on our ideas. (Check the appropriate box. Explain all No answers.)

☐ Yes ☐ No (Explain) ____

____

____

Stop! You have completed the activity.

If you still have time, go over your answers with your partner one more time.

Prep-Constructive Packet (9 pages)

Information to Begin!

This packet contains the materials you will need to complete today's activity. There are several concepts related to memory that this activity should help you to learn. Please read the instructions for each part carefully! You will first complete a worksheet that you will do alone. Please do this to the best of your ability! You will be using this worksheet to participate in a partner activity. After completion of this individual worksheet, you then will be partnered with another student to complete the collaborative part of the activity. Your instructor will give you a 2-minute warning to wrap up each part of the activity.

Please write your research activity ID number on all areas indicated!

Tips for Collaborating:

Work with your partner to contribute equally to your discussion. Listen to your partner's ideas and feedback. Share your own ideas and give your partner feedback. Ask each other questions and consider one another's perspectives. You should BOTH be talking and sharing! If you finish before the time allotted, review your collaborative work and make sure that you both agree that it is completed to your satisfaction.

Read this background information about memory before you start the activity!

Memory is a process that involves 3 kinds of actions:

1) *getting information "in" to memory, referred to as* ***Encoding***
2) *keeping it, referred to as* ***Maintaining*** *and* ***Organizing*** or ***Storage***
3) *and getting it back "out," referred to as* ***Retrieving****.*

People can encode information in several ways, maintain and organize it differently, and multiple factors can affect how people retrieve information from memory. Remembering can be a tricky thing! Various circumstances affect what we remember, why we remember some information well and some not-so-well, and why we forget information.

Individual Worksheet

Why do we remember what we remember?

Read through the description of the experiment and its results below[1]. You will analyze the data from this experiment to try to invent some general principles and ideas about how memory works. Further instructions are provided after you study the experiment.

Experiment: Remembering a Doctor's Visit

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 8 participants to read several paragraphs about different topics. They then tested them on only one of the paragraphs (below).

The Doctor Visit

John checked in with the doctor's receptionist. While he waited he read magazines. The nurse called his name. John undressed. John talked to the nurse. The doctor came in to the examination room. The doctor was very friendly. The doctor prescribed some pills for John. John forgot his wallet. John left the doctor's office.

The participants' task was to rate the following 12 sentences as to how sure they were that they had actually read the sentence on a scale of **1** (not at all) to **7** (very). They were not told which sentences had appeared in the paragraph.

For your convenience, the sentences that were repeated from the paragraph are highlighted.

A) John checked in with the doctor's receptionist.
B) John sat down.
C) While he waited he read magazines.
D) John followed the nurse.
E) John undressed.
F) John talked to the nurse.
G) The nurse tested John in the examination room.
H) The doctor examined John.
I) The doctor prescribed some pills for John.
J) John forgot his wallet.
K) John made another appointment.
L) John left the doctor's office.

[1] The activity example is slightly modified from activities Schwartz & Bransford (1998) used in several experiments.

Participant Ratings of How Sure They Were That They Had Read Each Sentence (1 = not at all to 7 = very)

ID*	A	B	C	D	E	F	G	H	I	J	K	L
1	7	7	6	3	4	3	6	1	4	7	2	7
2	7	5	6	1	5	4	7	3	3	7	1	7
3	6	7	7	4	3	5	5	4	5	7	3	6
4	7	6	7	3	6	4	4	3	2	6	4	7
5	7	6	5	3	4	3	7	2	6	7	4	6
6	7	4	7	2	5	5	6	3	3	7	2	7
7	7	7	7	2	3	3	5	5	5	7	1	7
8	5	5	7	4	5	5	4	2	4	6	3	7
Average	*6.63*	*5.88*	*6.50*	*2.75*	*4.83*	*4.00*	*5.50*	*5.50*	*4.00*	*6.75*	*2.50*	*6.75*

* Each participant was assigned a generic ID number.

A) John checked in with the doctor's receptionist.
B) John sat down.
C) While he waited he read magazines.
D) John followed the nurse.
E) John undressed.
F) John talked to the nurse.
G) The nurse tested John in the examination room.
H) The doctor examined John.
I) The doctor prescribed some pills for John.
J) John forgot his wallet.
K) John made another appointment.
L) John left the doctor's office.

Summary of Results

You can see that sentences A, C, J, and L were rated very high compared to sentences E, F, and I, even though they were all included in the original paragraph.
Sentences B, G, and H were also rated high, but did not appear in the paragraph.
Sentences D and K had the lowest ratings.
(*Remember, a high rating means that people were very sure that they remembered reading the sentence.*)

What do you make of this data? Interpret what this data means regarding how memory works. Remember, a high rating means that people were very sure that they had read the statement and a low rating means that they were NOT sure that they had read the statement. The highlighted statements actually appeared in the paragraph, while the other statements did not appear.

Think about the following questions:

- Do you see any patterns in the data?
- What kinds of statements did people remember accurately?
- Why would people remember certain kinds of information, but not other kinds?
- What kinds of statements did people believe that they had read, but did not actually read?
- Why would people falsely remember information?

Look at the individual scores, the average scores, and the summary results to try to understand the ideas behind the data! Use the table on the next page to write out your ideas about memory.

P-C

Individual Worksheet **Student ID: ______________**

How do you think memory works? Try to come up with at least 8 ideas about how and why we remember certain kinds of information, but not other kinds. You can think of them as *concepts* or *principles* about memory. Use the data from the experiment to support your ideas. Do NOT worry about "right" and "wrong" answers! The full chart is on the next few pages and includes columns for what people CORRECTLY remember, INCORRECTLY remember, and FORGET.

People tend to CORRECTLY remember…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.
	Please continue to the next page!	

People tend to INCORRECTLY remember…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.
People tend to FORGET…	How does the data support this idea? Explain.	Does this relate or conflict with your own experience? Explain.

Stop here! Wait until your instructor tells you to move on. If you have time left, go over your answers one more time!

Collaborative Worksheet

Student ID: ____________
Student ID: ____________

Now go over your table with a partner. Did you have some of the same ideas? Were some of your ideas different? Compare and contrast your ideas and write down the ones that you both agree on. Try to reach consensus for those you do not initially agree on. (If you are completely split on an idea, please indicate so.) You and your partner will create ONE list of memory concepts/ principles together and turn in ONE collaborative worksheet.

What do you believe are some general ideas about the kinds of information people tend to remember? Go to the next 2 pages to fill in your answers! See if you can agree on at least 8 ideas about memory. Feel free to include more.

People tend to CORRECTLY remember…

1. ______________________________

2. ______________________________

3. ______________________________

4. ______________________________

5. ______________________________

6. ______________________________

Continue to the next page!

7. ____

8. ____

People tend to INCORRECTLY remember…

9. ____

10. ____

11. ____

12. ____

13. ____

People tend to FORGET…

14. ____

15. ____

Continue to the next page!

16. ______________________________

17. ______________________________

18. ______________________________

Stop! You have completed the activity. If you still have time, go over your answers with your partner one more time!

APPENDIX B

PRETESTS AND POSTTESTS

Pretest (2 pages)

Student ID _________

Memory Quiz

Answer to the best of your ability. If you think you know the answer, please choose it. Try not to make blind guesses. If you are making a completely blind guess, please select, "I have no idea," or "I don't know".

True or **False**? Circle your answer.

1) When people understand something they read, they tend to recall exact sentences from it.
   True False I have no idea (circle one)

2) When people recall a story they read about a typical event, they tend to remember obstacles that hinder the completion of the event.
   True False I have no idea (circle one)

3) When people understand what they read, they later have difficulty distinguishing their own inferences from what was originally written.
   True False I have no idea (circle one)

4) When tested after reading a paragraph, people tend to forget the first sentence.
   True False I have no idea (circle one)

5) When recalling a written story about a familiar event, people tend to remember the most stereotypical parts of the event.
   True False I have no idea (circle one)

6) When people recall what they read about a typical event, they tend to add in details from their own experiences.
   True False I have no idea (circle one)

Multiple-choice. Circle your answer.

7) Imagine memory is like a file cabinet. Taking information out to use is called _______. Putting information into the cabinet is called _______.
   a. encoding; retrieval
   b. storage; encoding
   c. retrieval; storage
   d. retrieval; encoding
   e. I don't know

Continue to the next page!

8) Organizing and maintaining information in memory is part of _______.
   a. encoding
   b. storage
   c. retrieval
   d. transfer
   e. I don't know

9) When you are taking a test and you know that you *know* the answer to a question but are drawing a complete blank, you are having a problem with _______.
   a. encoding
   b. storage
   c. retrieval
   d. transfer
   e. I don't know

Please fill out the information on the last page!

Posttest (11 pages)

Student ID ___________

Memory Quiz

Read through the experiment below to complete the short answer questions. Answer to the best of your ability.

Experiment 1: The Balloon Story

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 40 participants to read the paragraph below and then they were asked to remember the paragraph as well as they could.

The Balloon Story

If the balloons popped, the sound would not be able to carry. Everything would be too far away from the correct floor. A closed window would also prevent the sound. This is because most buildings tend to be well insulated. The whole operation depends on a steady flow of electricity. A break in the middle of the wire would also cause problems. Of course, the fellow could shout. But the human voice is not loud enough to carry that far. An additional problem is that the string could break on the instrument. Then there could be no accompaniment to the message. It is clear that the best situation would involve less distance. Then there would be fewer potential problems. With face to face contact, the least number of things could go wrong.

Twenty participants read the paragraph by itself, and 20 read the paragraph with a picture attached (included on the following page). **Before you view** the picture, answer question (1) below.

1) Make **3 separate predictions** about what you believe participants will remember. Your predictions should be specific to this paragraph. Explain **why** you believe they will remember these details/aspects of the paragraph using the information you have learned about how memory works. Each explanation should be detailed enough to indicate your knowledge of memory, link directly to the prediction, and you should have distinct explanations for each prediction.

Use the table on the next page to fill out your predictions and explanations.
Please write in one **prediction** and corresponding **explanation** per row of the table.

Student ID ___________

Make 3 separate **predictions** about what you believe people will remember about the balloon paragraph. Write an **explanation** for why you think they would remember that information for each prediction.

Predictions - Details Remembered:	Explanations for Predictions:
a)	
b)	
c)	

DO NOT view the picture on the next page until you answer this question!

Student ID __________

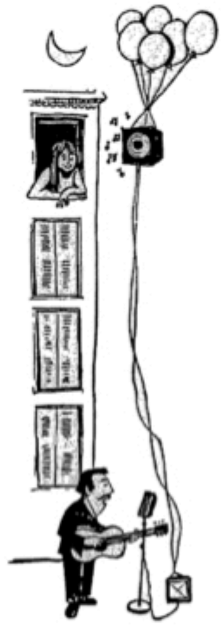

FIGURE A1 The picture seen by participants in the context-after and context-before conditions. (From Bransford & Johnson, 1972. Copyright 1972 by Academic Press.)

This picture was viewed by half of the participants.

2) Do you think that the participants who saw the picture will remember details from the paragraph differently than those who did not view the picture? Make **3 new predictions** about what you believe participants that saw this picture will remember. Provide **2 separate explanations** to explain each prediction:

1. Explain your prediction in general based on what you know about memory.
2. Explain how having the picture affects people's memories differently for each prediction you write.

Use the table on the next page to fill out your predictions and explanations.
Please write in one **prediction** and the two corresponding **explanations** per row of the table.

Student ID __________

Make 3 separate **predictions** and write 2 **explanations** per prediction.

Predictions – Details Remembered:	Explanations for Predictions:
a)	1.
	2.
b)	1.
	2.
c)	1.
	2.

Continue to the next page!

Student ID ___________

3) Make **3 separate predictions** about what you believe participants will **forget** by comparing the two groups of participants (those who saw the picture and those who did not see the picture). For instance, you might believe that one group would remember certain details better than the other group. Write this in the prediction column. Then, explain **why** you believe they will **forget** these details/aspects of the paragraph using the information you have learned about how memory works in the explanation column.

Predictions - Details **Forgotten** – Comparing the Two Groups:	Explanations for Predictions:
a)	
b)	
c)	

Continue to the next page!

Student ID __________

Read through the experiment below to complete the short answer questions. Answer to the best of your ability.

Experiment 2: A Scenario

Researchers wanted to test how well people could remember the details of a paragraph that they read. They asked 30 participants to read the paragraph below and then they were asked to remember the paragraph as well as they could.

The procedure was actually quite simple. First he collected all the items into one group. He might have had to use another place due to lack of facilities. But the usual facility was going to be enough. He arranged the items into different groups. Of course one pile might have been sufficient depending on how much he had to do. It is better to do too few things at once than too many. In the short run this may not seem important but expensive complications can easily arise. Red problems are the worst. He combined each group with the usual brand. At first, the whole procedure had seemed complicated. However, it had become just another fact of his life. He cannot foresee any end to the necessity for this task in the immediate future. But then, one can never tell. After the first procedures were completed, he moved all the groups and used the usual setting. When he arranged the finished materials into new groups, he was careful. He did not want to put an extra wrinkle in his work. He would just have more work later on. Finally, he put them into their appropriate places. Eventually they will be used once more and the whole cycle will then have to be repeated. However, that is part of life.

Fifteen participants read the paragraph by itself, and 15 read the paragraph with a title included. **Before you see** the title, answer question (4).

4) Make **3 separate predictions** about what you believe participants will remember. Your predictions should be specific to this paragraph. Explain **why** you believe they will remember these details/aspects of the paragraph using the information you have learned about how memory works. Each explanation should be detailed enough to indicate your knowledge of memory, link directly to the prediction, and you should have distinct explanations for each prediction.

Use the table on the next page to fill out your predictions and explanations.
Please write in one **prediction** and corresponding **explanation** per row of the table.

Student ID ___________

Make 3 separate **predictions** about what you believe people will remember about the scenario paragraph. Write an **explanation** for why you think they would remember that information for each prediction.

Predictions - Details Remembered:	Explanations for Predictions:
a)	
b)	
c)	

DO NOT read the title on the next page until you answer this question!

Student ID ___________

5) Do you think that the participants who read the title WASHING CLOTHES will remember details from the paragraph differently than those who did not see a title? Make **3 new predictions** about what you believe participants that read the title WASHING CLOTHES will remember. Provide **2 separate explanations** to explain each prediction:

1. Explain your prediction in general based on what you know about memory.
2. Explain how having the title affects people's memories differently for each prediction you write.

Predictions – Details Remembered:	Explanations for Predictions:
a)	1.
	2.
b)	1.
	2.
c)	1.
	2.

Continue to the next page!

Student ID __________

6) Make **3 separate predictions** about what you believe participants will **forget** by comparing the two groups of participants (those who saw the title and those who did not see the title). For instance, you might believe that one group would remember certain details better than the other group. Write this in the prediction column. Then, explain **why** you believe they will **forget** these details/aspects of the paragraph using the information you have learned about how memory works in the explanation column.

Predictions - Details **Forgotten** – Comparing the Two Groups:	Explanations for Predictions:
a)	
b)	
c)	

Continue to the next page!

Student ID __________

True or **False**? Circle your answer.

7) When people recall a story they read about a typical event, they tend to remember obstacles that hinder the completion of the event.
   True False I have no idea (circle one)

8) When people understand something they read, they tend to recall exact sentences from it.
   True False I have no idea (circle one)

9) When tested after reading a paragraph, people tend to forget the first sentence.
   True False I have no idea (circle one)

10) When people understand what they read, they later have difficulty distinguishing their own inferences from what was originally written.
   True False I have no idea (circle one)

11) When people recall what they read about a typical event, they tend to add in details from their own experiences.
   True False I have no idea (circle one)

12) When recalling a written story about a familiar event, people tend to remember the most stereotypical parts of the event.
   True False I have no idea (circle one)

Multiple-choice. Circle your answer.

13) When you are taking a test and you know that you *know* the answer to a question but are drawing a complete blank, you are having a problem with ________.
   a. encoding
   b. storage
   c. retrieval
   d. transfer
   e. I don't know

Continue to the next page!

Student ID ________

14) Organizing and maintaining information in memory is part of ________.
   a. encoding
   b. storage
   c. retrieval
   d. transfer
   e. I don't know

15) Imagine memory is like a file cabinet. Taking information out to use is called ________. Putting information into the cabinet is called ________.
   a. encoding; retrieval
   b. storage; encoding
   c. retrieval; storage
   d. retrieval; encoding
   e. I don't know

You are finished!

APPENDIX C

SURVEYS

Pre-Survey on Demographics and Preference for Collaboration

Student ID __________

Questionnaire

Please answer the following questions. You may leave any questions you prefer not to answer blank. All of your information is completely anonymous and confidential.

1. Gender: ___ Male ___ Female

2. Age: _____

3. Primary ethnic background:

___ American Indian or Alaska Native ___ White or Euro-American

___ Asian or Asian American ___ Native Hawaiian or other Pacific Islander

___ Black or African American ___ Hispanic or Latino

___ Other: ______________________

4. GPA: ___ 4.0-3.5 ___ 3.4-3.0 ___ 2.9-2.0 ___ Under 2.0

5. What career are you pursuing? ______________________

6. Rate your agreement with the following statements.

	Strongly Disagree						Strongly Agree
I enjoy working with a partner.	1	2	3	4	5	6	7
I like being able to share my ideas with someone.	1	2	3	4	5	6	7
I prefer to work alone.	1	2	3	4	5	6	7
I usually find a partner's feedback helpful.	1	2	3	4	5	6	7
I feel like I am on a team when I work with someone.	1	2	3	4	5	6	7
I feel that working with others usually hinders my progress.	1	2	3	4	5	6	7

Thank you! You are finished!

Post-Survey on Enjoyment of Task and Satisfaction With Partner

Student ID __________

Post Questionnaire

Rate your agreement with the following statements.

	Strongly Disagree						Strongly Agree
This activity helped me to understand the information better.	1	2	3	4	5	6	7
I enjoyed participating in this activity.	1	2	3	4	5	6	7
I was confused about the purpose of this activity.	1	2	3	4	5	6	7
I would like to participate in similar activities for other topics in class.	1	2	3	4	5	6	7
I found this activity boring.	1	2	3	4	5	6	7
This activity was very difficult.	1	2	3	4	5	6	7
I enjoyed working with my partner.	1	2	3	4	5	6	7
Explaining my ideas to my partner helped me understand the information.	1	2	3	4	5	6	7
I think that I would have done this activity better by working alone.	1	2	3	4	5	6	7
It was helpful to receive input from my partner.	1	2	3	4	5	6	7
My partner and I made a good team.	1	2	3	4	5	6	7
Doing this activity with a partner seemed to waste time.	1	2	3	4	5	6	7

10. Describe how well you and your partner worked together during this activity. ______________________

____

____

11. Had you ever worked with your partner before this activity? If so, describe how well you worked together in the past.

____

____

____

APPENDIX D

CODING MANUAL AND SCORING RUBRIC

Coding Manual (5 pages)

Instructions for Coding Free-written Responses – Posttest Prediction Tasks

General Guidelines:
Responses should indicate a prediction and a corresponding explanation. Each prediction and corresponding explanation shall be coded as a particular concept of memory. Eight total concepts (described below) shall be coded by name. Any predictions-explanations that cannot be coded as one of these 8 concepts shall be coded as "other". Any responses that do not represent a prediction-explanation shall not be coded.
A coded response indicates that a participant has demonstrated knowledge of that particular concept. If a response lists only a prediction and a term, the term must very obviously apply to the prediction (see examples for serial position effect*). If it does not, then it should not be coded (see examples for generation effect**).

Elaboration –
Written response demonstrates student understanding that people are more likely to remember new information that they can "link" or attach to their existing knowledge.
Examples:

"People would remember more when they see the picture because now they can know what the details of the paragraph are talking about."

"People would remember face to face content better when they see the picture because then they can see that this would really avoid all the problems."

"People would remember more details about the procedure because the title lets them remember it as laundry."

"People would remember the part about extra wrinkle because none of us want wrinkled clothes."

Schemas –
Written response demonstrates student understanding that people tend to remember information that supports their existing knowledge structures or patterns.
Examples:

"People would remember balloons. We have all had experience with balloons, so it's something they can remember from prior experience."

"People would remember a man singing to a woman because they can understand a romantic story of a man serenading a women. This is something they've seen before."

"People would remember that he had to sort a lot of piles because this is how you do laundry. We have all done laundry and we know you have to sort the clothes first."

"People would remember that this procedure is a part of life because we have heard this phrase many times in our life."

Instructions for Coding Free-written Responses – Posttest Prediction Tasks

Generation Effect –
Written response demonstrates student understanding that people are more likely to remember their own generated inferences rather than specific details of an event.
Examples:

"People would remember the smiling girl because this causes you to create a story from the picture and you can remember your own story better."

"People would remember the sound of a balloon popping because they make this association when they read that a balloon popped."

"People would remember the red problems because people tend to associate red with bad things, which we can easily remember."

"People would remember that the paragraph said the procedure was simple and complicated because that's a contradiction so we pay more attention to it."

** Examples of terminology use that should NOT be coded:

"People would remember arranging items. Generation effect."

"People would remember the fellow could shout. Generation."

** Without further explanation, there is no indication that these responses represent the link between memory and one's own inferences.

Obstacle Recall –
Written response demonstrates student understanding that people tend to remember parts of events that hinder the completion of a goal.
Examples:

"People would remember the wire breaking because this would mess up the guy's whole purpose and the girl wouldn't be able to hear his singing."

"People would remember the balloon popping because you can see from the picture that the speakers would not float if the balloons popped."

"People would remember doing few things is better than too many because we can imagine the problems that happen when we put too much clothes in the washer."

"People would remember that there were complications because we tend to remember things that cause problems."

Instructions for Coding Free-written Responses – Posttest Prediction Tasks

Gist –
Written response demonstrates student understanding that people tend to remember the general overview of an event or experience, or the main ideas of a passage, better than specific details.
Examples:

"People would remember sound because it was mentioned a lot so it must be a main idea."

"People would remember the overall story conveyed in the picture because otherwise the paragraph doesn't really make sense."

"People would remember groups and piles because this seems to be a main idea."

"People would remember laundry if they had the title, since this is easier to remember than a bunch of random statements."

Serial Position Effect –
Written response demonstrates student understanding that people tend to remember the beginning details and ending details of a passage, event, or set of information better than the details in the middle.
Examples:

"People would remember the beginning and end because this is more memorable."

"People would remember balloons popping because it's at the beginning of the paragraph."

"People would remember that the procedure was simple because it is the first sentence."

"People would remember it's a part of life because it was at the very end of the paragraph."

* Examples of legitimate codes for use with terminology:

"People would remember balloons popping. Serial position."

"People would remember things could go wrong. Serial."

* These actually are the first and last parts of this passage, thus the term serial position effect is correctly applied and shows understanding.

Instructions for Coding Free-written Responses – Posttest Prediction Tasks

Interference –
Written response demonstrates student understanding that people tend to forget information when new incoming information disrupts or overtakes the existing memory. Examples:

"When they didn't have the picture, they probably forgot most of it because there's too much information there for anything to stick out."

"Some people may forget some of the paragraph because the picture overtakes their memory."

"People would forget the usual facility because there's other information that is more important."

"People without the title would forget the paragraph because it's confusing and there's just a bunch of random sentences that are too detailed.

Encoding failure –
Written response demonstrates student understanding that people tend to forget details that do not seem important or meaningful because there is very little attention paid to them at the time of encoding.

"People would forget that the buildings are insulated because there is no reason to pay attention to this small detail."

"The picture group would ignore the details in the paragraph and forget them because they would focus on the story in the picture."

"People would forget the part about brands because no one cares what brands your clothes are when doing laundry. This is irrelevant information."

"Especially the middle parts would be forgotten by the no-title group. There's not much that catches their attention since it's random."

Other –

"People would remember the picture because it's always better to have a visual."

"People would remember happy feelings because we tend to remember emotions of characters."

"People would remember the man because visually seeing the man in the picture and reading the words let's you focus on it more in your memory."

Instructions for Coding Free-written Responses – Posttest Prediction Tasks

No Code –

"People would remember ______ because this is what I can remember."

Any prediction without an explanation connected to it.

Any terminology that is stated without it clearly linking to a prediction. Any terminology that is used incorrectly.

Scoring Rubric

Rubric – Student Understanding of Concepts of Memory
Two Prediction Tasks

	3	2	1	0	Score
Number of Distinct Concepts ("Other" counts at one concept)	Evidence of understanding of six or more distinct concepts	Evidence of understanding of 3-5 distinct concepts	Evidence of understanding of 1-2 distinct concepts	No evidence of understanding of any concepts	
Quality of Reasoning	Majority of the concepts are very clearly linked to predictions through student explanations.	At least half of the concepts are very clearly linked to predictions through student explanations. OR Majority of the concepts somewhat relate to the predictions, but are not explained in detail.	Student explanations loosely link predictions to concepts. Minimal effort is used to explain reasoning.	No evidence of understanding of any concepts.	
				Total (6 pts possible)	

APPENDIX E

INTRACLASS CORRELATIONS OF

STUDENT LEARNING SCORES

Table E
Intraclass correlations between partner scores on learning

	ICCs/ *p* values	
Condition	Adjusted Gains	Prediction Tasks
No Prep-Active	.25/ p = .51	-.08/ p = .80
No Prep-Constructive	-.29/ p = .48	.04/ p = .90
Prep-Active	-.01/ p = .99	-.18/ p = .64
Prep-Constructive	-.08/ p = .81	.20/ p = .50

APPENDIX F

INTRACLASS CORRELATIONS OF

STUDENT SURVEY SCORES

Table F
Intraclass correlations between partner scores on survey measures

Condition	ICCs/ *p* values		
	Preference for Collaboration	Satisfaction with Partner	Enjoyment of Task
No Prep-Active	.27/ p = .45	.09/ p = .78	.70/ p = .07
No Prep-Constructive	.36/ p = .31	.08/ p = .79	.40/ p = .21
Prep-Active	-.22/ p = .51	-.19/ p = .56	.72/ p = .08
Prep-Constructive	-.02/ p = .95	.20/ p = .52	-.09/ p = .77

APPENDIX G

SUBJECT CONSENT FORMS

Student Consent (3 pages)

****

CONSENT FORM
MAXIMIZING THE BENEFITS OF COLLABORATIVE LEARNING
IN THE COLLEGE CLASSROOM

INTRODUCTION

The purposes of this form are to provide you (as a prospective research study participant) information that may affect your decision as to whether or not to participate in this research and to record the consent of those who agree to be involved in the study.

RESEARCHERS

Dr. Kathryn Nakagawa, Associate Professor in the ASU School of Social Transformation, and Rachel Lam, graduate student in Educational Psychology, have invited your participation in a research study.

STUDY PURPOSE

The purpose of this study is to better understand how different methods of collaborative learning activities help students to learn. This understanding may be used to inform best educational practices for using collaborative learning methods in the classroom.

DESCRIPTION OF RESEARCH STUDY

If you decide to participate, then you will join a study involving the investigation of different ways that students can work with a partner or in a small group to complete learning activities in the classroom. Participation involves consent to allow the researchers to analyze your completed classwork and/or quizzes/tests pertaining to selected activities provided by your instructor. In addition, you will be consenting to allow the researchers access to your final course grades, however your name will not be linked to your grade for the purposes of this research, and your participation in this study will have absolutely no effect on your course grade. Since this research will be comparing different ways of doing collaborative classwork, you will be randomly assigned to a particular type of activity during the selected class periods related to this study.

If you say YES, then your participation will last for one semester, however only a few selected class periods will be utilized for this study. Your consent is completely voluntary. Your name or any other identifiable information will not appear on any data that is collected, ensuring full anonymity.
Up to 300 students will be participating in this study.

RISKS

There are no known risks from taking part in this study, but in any research, there is some possibility that you may be subject to risks that have not yet been identified.

BENEFITS

The possible/main benefits of your participation in the research are improved learning of course material and improved ability to collaborate in classroom settings.

CONFIDENTIALITY

All information obtained in this study is strictly confidential. The results of this research study may be used in reports, presentations, and publications, but the researchers will not identify you. In order to maintain confidentiality of your records, the research team will ask that you use only a generic ID code (to be assigned) on all data collection materials (classwork sheets, quizzes, surveys, etc.), rather than your name. In addition, any links to your course grades will use this generic ID, and no other identifying information such as your name. While we may audio-record some classroom activities, these materials will be kept in a locked cabinet on

ASU's campus and will only be accessible to the project research team, and will not be publically presented unless you give us consent to do so, and then, only in educational settings. The audio-recordings will be kept for the duration of the project and will be destroyed when no longer needed.

WITHDRAWAL PRIVILEGE
Participation in this study is completely voluntary. It is okay for you to say no. Even if you say yes now, you are free to say no later, and withdraw from the study at any time. Nonparticipation or withdrawal from this study will not affect your course grade in any way.

COSTS AND PAYMENTS
There is no payment for your participation in this study.

VOLUNTARY CONSENT
Any questions you have concerning the research study or your participation in the study, before or after your consent, will be answered by Dr. Kathy Nakagawa, School of Social Transformation, ASU, 480-965-0582. If you have questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk, you can contact the Chair of the Human Subjects Institutional Review Board, through the ASU Office of Research Integrity and Assurance, at 480-965 6788. Or, you may also contact Maricopa's Institutional Review Board Office 480-731-8701 with any questions or concerns.

This form explains the nature, demands, benefits and any risk of the project. By signing this form you agree knowingly to assume any risks involved. Remember, your participation is voluntary. You may choose not to participate or to withdraw your consent and discontinue participation at any time without penalty or loss of benefit. In signing this consent form, you are not waiving any legal claims, rights, or remedies. A copy of this consent form will be offered to you.

* Your signature below indicates that you consent to participate in the above study.

___________________	___________________	__________
Subject's Signature	Printed Name	Date

* By signing below, you are consenting to being audio-recorded during research activities in the classroom.

___________________	___________________	__________
Subject's Signature	Printed Name	Date

* By signing below, you are granting to the researchers the right to use your performance - whether recorded on or transferred to audio files - for presenting or publishing this research.

___________________	___________________	__________
Subject's Signature	Printed Name	Date

INVESTIGATOR'S STATEMENT

"I certify that I have explained to the above individual the nature and purpose, the potential benefits and possible risks associated with participation in this research study, have answered any questions that have been raised, and have witnessed the above signature. These elements of Informed Consent conform to the Assurance given by Arizona State University to the Office for Human Research Protections to protect the rights of human subjects. I have offered the subject/participant a copy of this signed consent document."

Signature of Investigator_____________________________________ Date_____________

APPENDIX H

INSTITUTIONAL REVIEW BOARD APPROVALS

Arizona State University

ASU Knowledge Enterprise Development

Office of Research Integrity and Assurance

To:	Kathryn Nakagawa WILSN
From:	Mark Roosa, Chair Soc Beh IRB
Date:	08/17/2012
Committee Action:	**Exemption Granted**
IRB Action Date:	08/17/2012
IRB Protocol #:	1208008120
Study Title:	Maximizing the Benefits of Collaborative Learning in the College Classroom

The above-referenced protocol is considered exempt after review by the Institutional Review Board pursuant to Federal regulations, 45 CFR Part 46.101(b)(1) .

This part of the federal regulations requires that the information be recorded by investigators in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects. It is necessary that the information obtained not be such that if disclosed outside the research, it could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.

You should retain a copy of this letter for your records.

Maricopa County Community College District
2411 West 14th Street
Tempe AZ, 85281
TEL: (480) 731-8701
FAX: (480) 731 8282

DATE:	August 23, 2012
TO:	Nakagawa, Kathryn, Education Lam, Rachel, Education
FROM:	MCCCD Institutional Review Board
PROTOCOL TITLE:	Maximizing the Benefits of Collaborative Learning in the College Classroom
FUNDING SOURCE:	NONE
PROTOCOL NUMBER:	2012-08-215
APPROVAL PERIOD:	Approval Date: August 23, 2012 Expiration Date: July 31, 2013
FORM TYPE:	NEW
REVIEW TYPE:	EXPEDITED

Dear Principal Investigator,

The MCCCD IRB reviewed your protocol and determined the activities outlined do constitute human subjects research according to the Code of Federal Regulations, Title 45, Part 46.

The determination given to your protocol is shown above under Review Type.

You may initiate your project.

If your protocol has been ruled as *exempt*, it is not necessary to return for an annual review. If you decide to make any changes to your project design which might result in the loss of your exempt status, you must seek IRB approval prior to continuing by submitting a modification form.
If your protocol has been determined to be *expedited or full board review*, you must submit a continuing review form prior to the expiration date shown above. If you make any changes to your project design, please submit a modification form prior to continuing.

We appreciate your cooperation in complying with the federal guidelines that protect human research subjects. We wish you success in your project.

Cordially,
MCCCD IRB

BIOGRAPHICAL SKETCH

Rachel J. Lam was born in 1979 in Indianapolis, IN. She completed her Bachelor of Science (BS) in Family Studies/Child Development and Master of Arts (MA) in Educational Psychology at Arizona State University, graduating Summa Cum Laude. Professionally, Rachel has worked for public libraries and non-profit educational institutions developing programs and curricula for children ages birth to 18 years and conducting trainings and workshops for parents and educators in early childhood education and development. During her first year of graduate school, Rachel became an instructor for a variety of educational psychology topics and taught both undergraduate and graduate courses for over five years. During the last two years of her doctoral program, she worked as a full time research assistant for projects funded by the Institute of Education Sciences and Spencer Foundation. She has had research experiences in the areas of collaborative learning, vicarious learning, teacher education, and parent education, and has designed and implemented learning and assessment materials for classroom and laboratory studies in biology, chemistry, research methods, psychology, and early brain development. In addition to experimental research and quantitative analyses, Rachel also has experience conducting interviews with children and adults, doing classroom observations, and analyzing qualitative data. Rachel has presented at several local and national conferences and was a student member of the Cognitive Science Society.