COMPLEX INSTRUCTION:
GIVING STUDENTS THE EDUCATION THEY DESERVE

by

Cindel K. Tobias

A Project Submitted to the Faculty of

The Evergreen State College

In Partial Fulfillment of the Requirements

for the degree

Master in Teaching

2010

This Project for the Master in Teaching Degree

by

Cindel K. Tobias

has been approved for

The Evergreen State College

by

Jon Davies

Jon Davies Ed.D., Member of the Faculty

June 2010

ACKNOWLEDGEMENTS

I would like to thank my friends and family for the endless support and encouragement.

ABSTRACT

This paper explores the details of complex instruction, student academic and social achievement within a complex instruction classroom, as well as implications for adopting this model. Many students do not enjoy math classes and do not feel it connects with their identity. Through a history of inequity within the public school system, teachers implementing complex instruction strive to make the classroom environment a place where all students are valued, interact with each other in equitable ways, and can have deep learning experiences. Through an analysis of the literature, complex instruction shows promising results in lessening achievement gaps, creating classrooms that reinforce sociomathematical norms, and improving conceptual learning. The research suggests that teachers adopting this model need to strive to comply with each facet of complex instruction as well as utilize support networks. Further research needs to be done with teachers trying complex instruction without the support of their math department.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Introduction

Many students have negative views of math and even adults almost proudly tout, "I am not a math person." For some students, math seems far removed from their everyday life experiences and their self-identity is very divergent from the identity perceived necessary to be a good math student. In traditional math classes, memorization skills, good penmanship, and diligent note taking result in math success. These skills do not represent the variety of skills necessary to be a successful mathematician, critical thinker, or problem solver. Within public schools, students each come to math class with varying prior knowledge, experiences, academic levels, needs, interests, English language proficiency, learning styles preferences, cultural backgrounds, social skills, etc. Math teachers need to create multidimensional classroom environments and curriculum that take into consideration students' unique characteristics by valuing a range of skills that are more in line with the semiotic domain of mathematics (Cohen, 1994). If students are able to make connections between their self-identity and their "math student" identity, they will be able to engage in mathematics in a real way and ultimately learn skills that will surpass the semester or year's end. Through effectively implementing complex instruction, teachers can start to address the diverse needs of all students and promote positive math identity through creating an equitable, multidimensional classroom environment where all students are valued (Cohen, 1994). This paper explores each component of complex instruction and implementation strategies through highlighting the historical background and professional literature of complex instruction, as well as applying key ideas specifically to teaching high school mathematics.

Rationale

As a high school math teacher, I will inevitably have students that have different needs. Even in a "homogeneous" academic level class, each student still has a variety of different needs that must be addressed in a classroom setting. It is my job to engage all students as they come into the class. A main role of a teacher is to promote academic growth from where the students start at the beginning of the school year or the new semester. When students are not engaged in the curriculum being taught, they are not learning to their potential and may develop negative views of the subject or domain. According to John Dewey (1997), "[C]urriculum must fall within the scope of ordinary life or everyday experiences, made relevant to students' lives and building on prior knowledge" (p. 73-74). Classroom activities that are applicable to everyday issues and the lives of students holds attention and clearly shows the benefits of working hard, thus promoting intrinsic motivation for learning. Teachers must provide activities in which students are challenged but do not view the activity as undoable, deemed within the students' zone of proximal development (Wells, 2002, p. 57). When teachers set up their classroom so that students are engaged within their individual zone of proximal development, they are able to learn to their full potential. To accommodate the various academic needs of each student, teachers can use cooperative learning strategies, such as complex instruction, to facilitate this deep learning. Simply teaching a mid-level curriculum will not engage all students; it may leave some students bored and others lost.

Students that are not fully engaged in the learning are not taking away from the class all of the information they are supposed to, which may result in lower standardized test scores. With policies in place that make standardized test scores increasingly important, the educational community continues to strive for new ways to increase student scores and ultimately "learn" more. Some teachers implement cooperative learning models to engage all students within their heterogeneous classrooms as a way to take into account academic needs, as well as differences in social skills, interests, learning preferences, etc. Teachers find that students need to be taught how to work effectively together and many teachers find that small group work tends to mirror the stratification that is present in the larger society. That is, some students perceived as "high competency" and others are deemed as having "low competency" (Cohen, 1994). The roles students play within groups are subconscious and typically mirror the stereotypes, biases, and perceptions held by society. Teachers typically find that while in small groups students deemed as "high competency status," do the majority of the talking, handle materials more frequently, and are looked at for approval (Cohen; Tammivaara, 1982). Interestingly "high competency status" is associated with white students of higher socioeconomic status (Cohen; Tammivaara). Students who fall into the "low competency status" role are typically not listened to or taken seriously by the group. They do not actively participate, are physically excluded from the group, and are often students of color and/or of lower socioeconomic status. "Those who do not participate because they are of low status will learn less than they might have if they had interacted more. In addition, those who are of high status will have more access to the interaction and will therefore learn more. It is a case of the 'rich getting richer'" (Cohen, p. 36). Despite many teachers' good intentions, the group work is only perpetuating already held academic disparities between students and therefore must be set up to celebrate the strengths and actively engage all students in equitable learning opportunities.

Also, many students find themselves working hard for a specific grade or finishing an assignment quickly. In traditional classes, where end products and speed are valued, students do not work hard to truly understand the material or realize that it is the process of getting the answer that is the most important and potentially beneficial. Teachers, through feedback, can have a large role in influencing whether students place higher value on the end product or the process. By emphasizing effort over ability, students see themselves as playing an active role in whether they are good at math or not (Blackwell, Trzesniewski, & Dweck, 2007). They can move from a static view of mathematical intelligence or ability to believe they are able to grow with experience, persistence, and hard work. This helps students feel that they have control over their intelligence or math ability and has been shown to promote higher levels of motivation, taking risks, and intrinsic challenge (Blackwell, Trzesniewski, & Dweck).

Students need quality educative experiences as defined by Dewey (1997) that not only engage students' everyday life, but also truly mirror the domain in which they are studying. Students should be challenged to be fully engaged in "doing mathematics." VanDeWalle (2007) defined doing mathematics as "engaging in the science of pattern and order…active thinking about mathematical ideas" (p. 13). Many mathematicians argue that traditional high school math classes are not engaging in the domain of mathematics. Successful math students are simply becoming more proficient at memorization and repeating procedures after seeing a discrete example. Thus few students possess the skills necessary to succeed in traditional math classes. Math students are not engaging in finding patterns, making generalizations given patterns, writing proofs, using precise and technical oral and written communication, justifying their answers, showing their thinking clearly using multiple representations, critical thinking, and/or problem solving. The skills required for being successful in the domain of mathematics is much more broad than those of traditional math classrooms. If teachers set up situations where a larger variety of skills are necessary and reflective of the domain of mathematics, more students will be able to see themselves as successful math students (VanDeWalle). Students also may develop negative concepts about math if they are not appropriately challenged in class or if they perceived themselves as not having the skills or abilities necessary for success in the math classroom.

How can math teachers engage students within their zone of proximal development, attend to issues of status in group work, celebrate multiple abilities, encourage effort over ability, and also provide students with quality educative experiences congruent with the domain of mathematics? With these large tasks placed before teachers and the educational community as a whole, we are all looking for answers. The classroom model of complex instruction was designed to attend to the previous question in its entirety.

Controversies

One of the biggest controversies about the topic of complex instruction revolves around tracking systems. Many schools have implemented different "tracks" for students based on academic ability as shown on test scores, grades in previous classes, and/or teacher recommendations. In high schools these tracks may take on the forms such as Honors programs and remedial classes. At some schools tracks are more fluid and self-selected, while other schools have more rigid tracks that are teacher appointed and/or based heavily on achievement on standardized tests. Some argue that heterogeneous classrooms are detrimental to the academic success of students that would be in "honors" or high-level tracks (Webb, Nemer, & Zuniga, 2002). Critics argue that high-ability students are held back by heterogeneous grouping and are not able to move as quickly or learn as much as they could in a homogeneous ability environment (Webb, Nemer, & Zuniga, 2002). When looking at issues such as tracking, it is important to analyze who tracking systems benefit, who tracking system do not benefit, and issues of equity, privilege, and power within our larger society and also within the history of public education. Students who are deemed low ability and are in the corresponding "track" typically have less interesting curriculum, less qualified teachers, bigger class sizes, and teachers who hold lower academic expectations for achievement (Lucas, 1999, p. 3). Researchers of complex instruction have shown that students of both high and low ability have something to gain from heterogeneous grouping in a complex instruction learning environment (Boaler & Staples, 2008). They have empirically shown that achievement gaps between minority groups nearly disappear, overall student relationships with math improve, higher percentages of students take higher level math, and high-ability students still achieve high levels of academic success in an equitable, heterogeneous class environment (Boaler & Staples). Chapter three further analyzes the work of Boaler and Staples.

Complex instruction relies heavily on conceptual based learning. Some critics argue that conceptual based learning takes too long and that teachers are not able to "cover" as much material as teachers who use more traditional methods of teaching. Also, other critics claim that students who do not directly learn discrete facts and procedures in a more traditional setting will struggle to succeed on standardized tests. In Washington State, there have been continual curriculum switches from IMP or conceptual, inquiry based mathematics to more traditional, "back to basics" curriculum reform. The National Council of Teachers of Mathematics (NCTM) calls for teachers to give students multidimensional, real-life tasks that evoke student interest, mirror the domain of mathematics, and enhance student critical thinking and problem solving skills (Stipek et al., 1998; Henningsen & Stein, 1997). Complex instruction allows students to initially learn concepts in a more constructivist way and then reinforce those key procedures and facts through individual work (Cohen, 1994). Constructivists also argue that unless students learn the "why" behind the facts, students do not retain the information being covered by the teacher (VanDeWalle, 2007).

There are also controversies surrounding nearly all aspects of group work. Some argue that groups should receive rewards and/or compete with other groups within the class to make group work most effective (Slavin, 1996). Some researchers think that groups should be ability grouped (Webb, Nemer & Zuniga, 2002), while others believe heterogeneous academic groupings provide more equitable and still advanced learning environments (Boaler & Staples, 2008). Within the complex instruction model, students do not receive rewards or compete with others (Cohen, 1994). Complex instruction researchers believe that by setting up situations in which students are engaged in relevant, domain specific group worthy tasks students will build intrinsic motivation to learn which manifests, for example, as higher levels of students voluntarily upper level math classes (Boaler & Staples). Chapter three addresses in further detail group work.

Another controversy about complex instruction is how to achieve heterogeneous classes. Some schools have eliminated tracking systems all together (for example all freshman taking the same math class), while other schools allow students to choose their type of class setting. There are also debates on how long math classes should be, with most traditional schools offering math classes that are a year long. In one high school implementing complex instruction school wide, the teachers decided to have math classes on a semester long, but with longer overall class periods. Having classes only one semester long, except algebra, allowed students flexibility in their schedules and gave students the opportunity to take classes over if needed without affecting the number of math classes possibly taken or level reached (Boaler & Staples, 2008). Ultimately teachers implementing complex instruction can work with administration and other teachers in the department decide how best to schedule math classes.

Another argument among students, teachers, and parents with models of teaching that rely heavily on group work may be grading. In the traditional grading system, one assignment is given to all students are graded on the same rubric emphasizing achievement. This model seems to be "equal" for all students. However, in a complex instruction model, different groups may receive different tasks and be graded based on effort and growth instead of meeting a specific level of achievement. Students may also receive group grades or take group tests. In a complex instruction model, students receive similar tasks to complete but are not always the same exact assignment. Some students, teachers, and parents may not see this as a "fair" way to set up a classroom or way to grade students. It is important to look at who the current, traditional grading system benefits and what are students really learning from these assignments, rewards, punishments, and/or feedback. In most traditional classes, compliance, timeliness, and the end product is rewarded instead of effort. Successful complex instruction learning environments possess value in collaboration over competition and promote intrinsic motivation instead of extrinsic. Each teacher implementing complex instruction can determine the criteria for assigning grades based school policy, their own personal experiences, and the needs of the students.

Definitions

A main term that will be used throughout the paper is complex instruction. Complex instruction was originally designed to "facilitate the development of the cognitive, academic, and linguistic functioning of all students in heterogeneous classroom" (Lotan, 1989, p. 1). Researchers at Stanford University modeled this teaching method from organizational sociology theory with original research focusing on helping recently desegregated classrooms become more equitable (Lotan, 1989). Over the last two decades, complex instruction has evolved with educational research to include more recent research on cognitive demand, group worthy tasks, and NCTM reform-oriented teaching strategies. Cognitive demand is deemed the level of reasoning, problem solving, and conceptual understanding needed to complete a task or activity. Low cognitive demand is associated with memorization and repeating procedures without conceptual connections, while high cognitive demand activities require students to synthesize information, make generalizations, justify reasoning processes, explore patterns, estimate solutions, apply previously learned information to new concepts, etc (Henningsen & Stein, 1997, p. 525). Group worthy tasks are defined as being open-ended, requiring complex problem solving, having multiple entry-points and ways to show intellectual competence, being realistic to domain specific tasks, including clear criteria of the end group product, and requiring both interdependence between students and individual work (Lotan, 2003, p. 72). As mentioned previously the NCTM are encouraging teachers to vary the cognitive demand in classrooms, including higher level thinking along with memorization and other "low" cognitive demand tasks. Similarly, the math education reforms encouraged include tactics utilized in complex instruction such as multidimensional classrooms that work with real life tasks similar to those found in the domain of mathematics (Stipek et al., 1998). Chapter three specifically addresses and further explores cognitive demand and group worthy task research.

Complex instruction outlines how a teacher can set up the classroom environment and curriculum in order to successfully challenge all students in a heterogeneous class in an equitable way. Complex instruction utilizes heterogeneous, random grouping of students with the belief that all students have something valuable to bring to the group and that all students are capable of playing all roles in the group (Cohen, 1994; Lotan, 1989). Students are given specific group roles and groups are typically four or five students. Over the course of the school year, each student will have tried on each group role and worked with nearly every other student in the class (Cohen; Lotan). Teachers give students group worthy tasks to complete together, while the teacher implements multiple ability treatments and status assignments. The teacher picks groups to present their finding to the class, learning groups switch after each unit or after about two weeks, students complete group and individual work, and teachers work hard to delegate authority to create interdependence between peers (Cohen; Lotan).

In multiple ability treatments within complex instruction teachers remind students that everyone in the group is valuable and explicitly state, "No one has all the skills required for the task at hand, and everyone has at least one skill to contribute" (Cohen, 1994). The teacher also reads aloud or provides a list of necessary skills for the task before students start working together. This helps students see the value in listening to everyone and feeling that they have something to contribute to the group (Cohen; Lotan, 1989). Teachers also assign competence to "low status students." Low status students are those who for a variety of reasons are expected to not be competent in the subject or task at hand (Cohen; Lotan). Teachers use their high status in the classroom, to bring attention to good work low status students are completing. This may take the form of a teacher looking over a low status students' shoulder and ask the student to explain their ideas to the group or a teacher having a low status student be the "group expert" and teach their group a certain idea, concept or topic (Cohen; Lotan). It is important to note that assigning competence needs to be genuine and drawing attention to important skills relevant to the task or class that they possess. Chapter three addresses this topic further (Cohen; Lotan).

Limitations

This paper is devoted to exploring the ideas behind and effective strategies for successful implementation of complex instruction within high school math classes. An inherent bias within this paper is assuming that complex instruction is an effective teaching model and that it is a worthwhile group work model for high school math classes. There is only one section in Chapter three devoted to explicitly showing student achievement within the complex instruction teaching model. The body of literature on the effectiveness of complex instruction when compared to traditional classroom models is still quite small. The emphasis of this paper is to talk at length about the aspects of complex instruction and how to effectively implement each of these aspects.

Summary

Math teachers have the large task of presenting math in a way that engages students and promotes long term learning that they can directly see and use in their everyday lives. Through implementing complex instruction, teachers can start to address the varying needs of students and promote growth while valuing varied perspectives and skill sets. This paper explores the details of complex instruction and how teachers can effectively implement this multifaceted and involved classroom model. Chapter two explores the history that has lead to the advent of complex instruction and specifically how math education connects to this classroom model. Chapter three reviews literature on the different conceptual aspects of complex instruction, academic and social student achievement within complex instruction, and specific tactics for effectively implementing complex instruction. The final chapter, Chapter four, synthesizes information from the history and literature on complex instruction while giving implications for high school math teachers wanting to take on using complex instruction as a class model.

CHAPTER 2: HISTORICAL BACKGROUND

Introduction

Chapter one oriented why complex instruction is worth exploring as well as set up different controversies that lie within this classroom model. This paper overall investigates the details of complex instructions and effective strategies for implementing this classroom model. Chapter two explores the history of math education and tracking, "status" within the education system, multidimensional classrooms, the semiotic domain of mathematics and identity development, and the advent of complex instruction. When American schools started to take shape, teachers were given the duty to educate the children of the surrounding community. For teachers this meant teaching children with varying ages, developmental levels, needs, interests, and backgrounds. Much has changed from the initial formalizing of schools to public education today. We find that students are now grouped by age and also at times by "ability." Although intentional grouping occurs, teachers still find themselves with an ever-growing diverse group of students with different needs, interests, prior knowledge, cultures, languages, etc. Complex instruction has been one way researchers and teachers are trying to reach every student in heterogeneous, multidimensional classes and allowing them to learn to their full potential while gaining important social skills.

History of Math Education and Tracking

Mathematics (deemed arithmetic) was included in core curriculum from the beginning of schools along with Latin and the two other "R's" (Spring, 2008, p. 255). At the birth of high school in 1821, high school curriculum consisted of advanced sciences, mathematics, English studies, history, and political economy (Spring, p. 97). The community expected high schools to adequately prepare students for their future careers, which led to the practice of differentiating the curriculum based on the "future occupational destination of the student" (Spring, p. 252). Students were selected or chose one of the few explicit "curricular locations" before entering high school (Lucas, 1999, p. 2). Students chose or were selected into curricular tracks that maintained socioeconomic status quo and social biases and were in those tracks for the remainder of their time in public education (Lucas, p. 3; Huerta, 2009, p. 90). That is, poor, minority students were tracked explicitly into curricular programs that would make them good workers, while rich, white, male students were exposed to a rigorous curriculum that would prepare them for further study at a university.

Not only were students tracked via their socioeconomic status, but they were also believed to have unitary intelligence (Lucas, 1999, p. 4). Students were deemed either capable to learn complex material or not. Specific skills or subjects were not looked at separately or assessed differently. Students were enrolled in a specific track that contained all advanced classes, all regular classes, or all basic classes. The students in the corresponding tracks were believed to be "smart" or not smart across all subjects. Thus, success in math classes was paired with success in other classes in the minds of teachers and administrators.

Starting in 1965, the practice of explicit tracking was dismantled in public education (Lucas, 1999, p. 6). This revolution of thinking about American schools shifted during the civil rights movement. Explicit tracking was found to provide different curriculum, quality of teaching, cognitive preparation, and socialization for different students based primarily on race, socioeconomic status, and perceived ability and career path (Lucas, p. 3). This changed how schools organized and scheduled students for classes. The old track labels were applied to specific courses instead of entire academic programs (Lucas, p. 6). For example, schools would offer honors math and English courses and students would have to sign up for each separately. Since these tracks were no longer explicit grouped courses, detracking advocates hoped that this would allow minority students easier access to rigorous classes to prepare students for a variety of careers. By 1981 most high schools no longer formally or explicitly tracked students into curriculum programs (Lucas, p. 1). Despite best efforts of detracking, “it seems that the unremarked revolution transformed and submerged, but did not uproot, a stubborn stratification system” (Lucas, p. 1). Within complex instruction, researchers emphasize that math classes are organized in a way that all students start on the same “track” with all freshman enrolled in algebra, giving all students an “equal” opportunity to learn the curriculum as well as learn from working with each other (Boaler & Staples, 2008). One school implementing complex instruction also arranged math courses differently so that students had more flexibility in their schedules to fit in more math classes by making them a semester long instead of a year long like traditional high school math classes (Boaler & Staples). In contemporary schools, despite detracked schools and heterogeneous classroom environments, issues of status are still at play when students interact with each other (Cohen, 1994).

Status Research

Within the field of sociology, many researchers have attributed inequity in the schools to structural features of such as tracking systems and ability grouping (Vanfossen, Jones, & Spade, 1987; Gamoran, 1987). Due to reforms in education, classrooms are becoming more heterogeneous and many teachers are turning to cooperative learning to create equitable learning environment (Rubenstein & Bright, 2004). Despite best efforts, these small groups develop status orders based on perceived popularity and academic ability that mirror larger stratification of society and impact the interactions of the students (Cohen, 1984). Although structural features are important to how schools and classrooms function, like tracking, there is also a large body of sociological research on the "process of stratification in heterogeneous social settings" that has implications for implementing group work in a heterogeneous classroom setting (Cohen & Lotan, 1995, p. 100). When schools were desegregated and busing programs started, educational research started to focus on issues of status and equity within classrooms. Cohen and other researchers at Stanford delve into assessing the affects of status on classroom participation and learning while searching for strategies to make classrooms more equitable places (Cohen & Lotan). The original intentions of using multiple ability curriculum and expectation training were to create equitable participation of students deemed low ability and also minority students in recently integrated classrooms. The National Institute of Education found that academic ability and racial stereotyping were not the only factors influencing how students participate and interact with others in the classroom. Minority, low ability students that were more popular among peers were more influential and seen as competent (Cohen & Deslonde, 1978). Consistent with expectation theory and self-fulfilling prophecy theory, students responded to multiple ability treatments and expectation training by working with each other in more equitable ways (Cohen & Deslonde; Cohen, 1994).

Multidimensional Classrooms

Psychologists like Carol S. Dweck and Howard Gardner have revolutionized ideas about human intelligence. Prior to Stephen Jay Gould's *The Mismeasure of Man* (1981), human intelligence was viewed as unidimensional and uniform across many different domains and even races. That is, people were deemed either intelligence or not. Dweck has done much research on entity and incremental intelligence theory. People who believe intelligence is fixed and an innate characteristic hold entity views of intelligence, while those who hold incremental views see intelligence as malleable and able to change with experiences and learning (Blackwell, Trzesniewski, & Dweck, 2007). Through researching intelligence theories, Dweck has found that teachers, through feedback, praise, and emphasis on effort over ability, can influence students to hold more incremental ideas of intelligence. This also impacts students' motivation, persistence, willingness to take risks, and ultimately learning (Blackwell, Trzensniewski, & Dweck).

In the 1980's Gardner developed different categories of intelligence linked to areas of the brain such as linguistic, musical, logical-mathematical, naturalistic, spatial/ visual, bodily-kinesthetic, intrapersonal, and interpersonal intelligences (Aborn, 2006). Despite a more inclusive view of intelligence, some educators have used theories such as Gardner's to categorize students and view their students' intelligences as fixed (Christodoulou, 2009). An alternate way to look at the research of Gardner is to view students as "unlimited human beings" with many talents and abilities to share with others (Aborn, p. 83). From a malleable and diverse view of intelligence, researchers of complex instruction have developed multiple ability treatments and multidimensional classroom environments. By having teachers explicitly telling students the different skills necessary for each task, students can start to determine personal areas of strength and improvement. Along with this developing self-awareness, teachers also tell students that no one has all of the skills necessary for the task and that everyone has at least one skill to contribute (Cohen, 1994, p. 126). This allows students to see each other as valuable group members that have at least one area of "intelligence." Along with more clearly seeing different strengths, students also are better able to pinpoint specific areas of improvement.

Multidimensional classrooms, specifically math classes, different drastically from traditional math classes yet mirror closely the domain of mathematics. In traditional math classes students that are good at memorization, following procedures shown, and complying with teacher expectations are highly successful math students (VanDeWalle, 2007). This narrow range of skills only allows a small number of students to be successful math students and does not accurately prepare students to become members of the domain of mathematics (Boaler & Humphreys, 2005; VanDeWalle). Having a variety of skills needed for success makes math more accessible to a greater number of students and also creates a need for interdependence between students. When teachers create environments where a variety of skills are necessary, the class environment is more congruent with the real life and the work "real" mathematicians do within the domain of mathematics (VanDeWalle).

Semiotic Domain of Mathematics and Identity Development

James Paul Gee (2007) wrote *What Video Games Have to Teach Us About Learning and Literacy,* explaining how video games model learning techniques that could be used to engage the learner and facilitate deep situated learning. Gee talked extensively about semiotic domains and identity, and their roles in learning. According to Gee, a semiotic domain is “an area or set of activities where people think, act, and value in certain ways” (p. 19). Semiotic domains are not only areas or sets of activities, but also include a community of other people who actively participate in that area or given activity. Gee went on to explain that actively learning a semiotic domain involves learning to experience the world in a new way, joining a new social or affinity group, and gaining new resources that prepare for more learning within the specific semiotic domain (p. 24). Students can be exposed to new semiotic domains through activities and experiences teachers create in the classroom. By engaging students in authentic and “quality” experiences, they can start to see themselves as part of that semiotic domain (Dewey, 1997).

As students see themselves more and more connected to the new semiotic domain, identity “changing” also starts to occur. Gee (2007) defined three types of identity: virtual identity, real-world identity, and projected identity (p. 48). The virtual identity involves the learner of the new semiotic domain taking on characteristics of the semiotic domain group while in a “virtual” situation (Gee, p. 49). An example is a student taking on the “virtual identity” of a mathematician within in high school math classroom. The real-world identity includes the characteristics and prior knowledge that comes with the learner from other semiotic domains of which they are members (Gee, p. 50). For example, outside the math classroom a student may also identify with (be member of the semiotic domains) soccer players, high school seniors, females, and children of divorced parents. Gee defined the final identity, projected identity, as “to project one’s values and desires onto the virtual character… and seeing the virtual character as one's own project in the making" (p. 50). In the math classroom example, this would be when the student starts to combine the virtual identity of a mathematician and the real-world identities. The projected identity expands the virtual identity beyond the math classroom and the student starts to feel like a member of the new semiotic domain.

Some students may come into the classroom with preconceived notions of themselves (real-world identities) as well as what it takes to become a member of the new semiotic domain (virtual identities). These identities may be very divergent. Inner speech and personal dialogue provide people a way to process their environment and to make sense of their surroundings (Sacks, 2001, p. 39). When faced with differences of power and privilege, students' inner speech may inhibit bridging virtual and real-world identities through negative self-talk. For example, a female student may see herself as "naturally" not good at math because she views her gender not represented in the semiotic domain of mathematics. Thus, she may engage in negative self-talk such as, "I'm a bad math student" or "I've never been good with numbers" that reaffirms the gap between identities. The teacher, knowing the student's real-world identities, can then create experiences for this student that helps her bridge the identities of female and successful math student. Maxine Greene in *Releasing the Imagination* mentioned bridge building and repairing these gaps between identities. Greene (1995) explained, "Learning to look through multiple perspectives, young people may be helped to build bridges among themselves… they may be provoked to heal and to transform" (p. 167). She suggests that despite institutional biases and disparities in power and privilege, students can take on identities they may originally thought were inaccessible. A teacher can aid students in seeing multiple perspectives by exposing them to the many layers of the semiotic domain and multidimensional classrooms. An example of this would be a math teacher inviting students to explore why nearly all historically recognized and published mathematicians are white European men. Students then could research the history of mathematics in non-European cultures and their contributions to the math community. Students also could look at currently published mathematics, seeing the growing number of females and people of color actively contributing to the literature of mathematics. By engaging in activities such as these, female students and students of color will be able to see math from different perspectives and hopefully will start to transform their ideas of who is in the semiotic domain of mathematics. These activities also help transform students' ideas of status and perceived ability of themselves and others.

Once students see themselves able to take on the virtual identity of the semiotic domain, they can actively experiment with virtual identities, and in turn projected identities, engaging in deeper learning. Students connect prior identities and prior knowledge to the content learned to form new identities within the semiotic domain. When teachers tap into prior knowledge, students feel that their previous identities and knowledge are valuable and worth bringing into the classroom. Teachers must elicit the students' prior experiences and "funds of knowledge" to help them make meaningful connections between identities and engage the new information and semiotic domain in a way that facilitates deep learning. Students that feel valued are more likely to share the skills and areas of strength with the rest of the class allowing groups to gain more from each other and the tasks they are working on.

History of Complex Instruction

Given the history surrounding math education, heterogeneous classrooms, status issues, multidimensional classrooms, and mathematical identity development, educational researchers at Stanford University were seeking an instructional model that would attend to these issues and more. When complex instruction was first created, it was called an instructional technology (Lotan, 1989). Researchers Cohen, Deal, Meyer, and Scott (1973) started applying theories from sociology to education and found connections to organizational sociology and school classroom environments. They focused particularly on aspects of what makes an organization run efficiently or productively. Complex instruction was described in terms typical of organizational psychology such as of lateral communication, delegation of authority, non-routine tasks, variability, and learning centers (Lotan). Now complex instruction typically includes words that have commonly arose in other contemporary educational research such as cognitive demand, group worthy tasks, intellectual role taking, and cooperative learning groups (Cohen, 1994).

Over the last 20 years, researchers at Stanford have developed a new twist on cooperative learning called complex instruction "designed to facilitate the development of the cognitive, academic, and linguistic functioning of **all** students in heterogeneous classrooms" (Lotan, 1989, p. 1). The main goal of complex instruction was to give students equal access to what the researchers at Stanford deemed critical classroom features including curricular material, interactions with teachers, and interactions with peers (Lotan). Researchers Cohen, DeAvila, and their associates designated three main features of complex instruction: curriculum that includes multiple ability tasks, organizing the classroom so teachers delegate authority to students, and regular implementation of status treatments (Lotan). Also they developed an English/ Spanish math and science curriculum called Finding Out/ Descubrimiento to implement along side complex instruction classroom model (Lotan). The curriculum included activity cards with non-routine and open-ended tasks outlined for students to complete together in small, heterogeneous groups of four or five students (Lotan). The students would move from task to task with different tasks happening simultaneously within one classroom (Lotan). Students were asked to use a variety of skills to complete the learning activities while developing interpersonal skills through working with others (Lotan). Within teacher implementation, Cohen and DeAvila emphasized peer interactions, delegation of authority, including cooperative norms and students taking on roles, letting all students play a variety of group roles, and treating status problems with multiple ability treatments and assigning competence to low status students (Lotan).

The complex instruction model holds the underlying belief that all students are capable of doing the subject, and have something worthwhile to share to the group and overall community (Cohen, 1994). This belief is played out in the classroom by teachers making sure that every student holds a variety of roles in group work and also informing students about the multiple abilities necessary to complete tasks (Cohen). The emphasis on multiple abilities is also paired with teachers reinforcing that not one person has all the skills necessary to complete the group tasks. This helps students to see fellow students as having something worthwhile to contribute as well as making the groups more interdependent. Researchers of complex instruction claim that this model effectively meets the academic needs of all students, lessens achievement gaps, improves students' relationship with mathematics, and leads to more students taking higher level math classes (Boaler & Staples, 2008; Albert & Jones, 1997). Research on student academic and social achievement within the complex instruction model is further addressed in chapter three.

Summary

Despite a history of students not having equal opportunities within the education system, the educational community is working to make our now growing heterogeneous classes more equitable. Through an account of status, multidimensional classrooms, semiotic domains, and identity development, the advent of complex instruction takes pieces of each to help students achieve both academically and socially in an equitable environment. Chapter three reviews research on each specific facet of complex instruction, student social and academic achievement, and strategies for implementing complex instruction.

CHAPTER 3: CRITICAL REVIEW OF THE LITERATURE

Introduction

This paper investigates the details of complex instructions and effective strategies for implementing this classroom model. Chapter one explored why complex instruction is worth investigating and sets up different controversies that lie within this classroom model. Chapter two highlighted important events that have contributed to the history of complex instruction such as tracking systems, status research, multidimensional classrooms, and semiotic domains and identity development. Chapter three critically reviews the literature of complex instruction. Chapter three specifically addresses complex tasks, conceptual learning, and cognitive demand; heterogeneous group work; status and competency expectations; students' beliefs, self-efficacy, and identity; student academic and social achievement; and implementation of complex instruction.

Complex Tasks, Conceptual Learning, and Cognitive Demand

Complex instruction has many different facets that comprise this teaching model. One of the key aspects is having tasks at the core of the curriculum to challenge students, uphold varied levels of cognitive demand, and require a variety of skills of students. Henningsen and Stein (1997) looked at specific teacher actions and how that influences the cognitive demand of the complex tasks. They found that teachers needed to be very mindful of appropriateness of the task, pressing for justification/ explanation, scaffolding, giving the appropriate amount of time, and providing tasks that build on students' prior knowledge. Stipek et al. (1998) found that keeping cognitive demand high, pressing students for justifications, and positive teacher affect and dispositions are important for creating a safe learning environment where students feel comfortable to deeply learn and take risks. Kazemi and Stipek (2001) analyzed classroom practices that create press for conceptual learning. Khisty and Chval (2002) found that teachers' talk plays an important role in equitable student conceptual and deep learning. Arbaugh and Brown (2005) found that despite teachers learning about levels of cognitive demand, it is hard to make consistent changes in mathematical task choices and takes much reflection on the part of the teacher.

Henningsen and Stein (1997) completed a qualitative study with analyses of specific case studies from previously gathered data. The researchers focused on the implementation of high level, complex tasks. Many tasks are set up to be rich and full of opportunities for students to engage in "doing mathematics." However, through the implementation of the task it can diminish in cognitive demand and simply become a more routine, procedural task. The researchers aimed to identify and explicitly illustrate the classroom factors that shape students' engagement with high-level tasks. Henningsen and Stein helped collect data in a previous study as part of the QUASAR project looking at mathematical tasks and how they impact student learning and thinking.

The participants of the original study were middle school students, primarily minority and of low socioeconomic status. The teachers in the study were part of the QUASAR project and represented teachers from ethnically diverse, urban schools from six different geographic locations. The teachers themselves also represented more ethnic diversity when compared to the national profile. The researchers observed twelve teachers three times a year from 1990 to 1993. During each observation the researchers took field notes and videotaped three consecutive days of lessons. The researchers then used both the field notes and videotape to code the data to provide a description of the physical setting, the chronology of events, and make remarks about mathematical tasks, intellectual environment, classroom discourse, management and assessment practices, and group work. From the original study, Henningsen and Stein (1997) examined in detail 58 of the 144 tasks that were deemed “doing mathematics” and implementing high level mathematical tasks. After analyzing of the 58 lessons involving high level mathematical tasks, the researchers found that 22 of the lessons engaged students in doing mathematics, eigth of the lessons resulted in students focusing on procedures without connections to underlying meaning, 11 of the lessons resulted in unsystematic exploration, and 10 of the lessons resulted in no mathematical focus. The researchers, after categorizing the lessons, examined the factors of cognitive demand maintenance or decline. They found that appropriateness of the task, teacher scaffolding, appropriate amount of time, high-level performance modeled, sustained press for justification, student self-monitoring, and teacher drawing conceptual connections were important to the maintenance or decline of the cognitive demand.

The researches then analyzed each task category (doing mathematics, procedures without connections, unsystematic exploration, or no mathematical focus) and calculated the percentage of lessons that utilized the factors of maintenance or decline of cognitive demand. They found that for cognitive demand maintenance, teachers implemented tasks that were align with students prior knowledge (82%), scaffolding (73%), flexibility and allowing appropriate timing (77%), high level performance was modeled (73%), and the teachers sustained a high press for student justification, explanation, and meaning (77%). These findings were congruent with the literature on maintaining high cognitive demand. The researchers found that for cognitive demand decline into using procedures without conceptual connections teachers gave students too much or too little time to work on the task (75%), challenges within the task became nonproblems (100%), and students resorted to focusing from learning and understanding to getting the correct answer and finishing (75%). For cognitive demand decline into unsystematic exploration, they found that teachers assigned tasks that were inappropriate for the students (64%), too much or too little time to work on the task (55%), and challenges within the task became nonproblems (45%). Lastly, for cognitive demand declining into no mathematical activity the researchers found that tasks were inappropriate for the students (80%), classroom management problems were present (70%), and too much or too little time was given to work on the task (50%). The researchers then found four specific teacher observations that fit in each maintenance or decline category to get a more detailed picture of what is happening in the classroom.

The researchers used multiple coders and checked the intercoder reliability to make sure the scores given were accurate. The researchers came up with the set of factors (e.g. to much time or too little time, appropriateness of task, etc.) that maintained or declined cognitive demand before coding the lessons. Thus some lessons do not fit as easily into the categories as others. The researchers stated the higher percentages of factors may have been found in the decline of cognitive demand into unsystematic exploration case if the categories were different. The specific lessons analyzed at the end of the article show how what they found plays out in the classroom. These case studies are not generalizable to every classroom and are not the whole picture of how decline or maintenance of cognitive demand happens. Also when implementing complex tasks, teachers may run into situations in which their cognitive demand level does not fit into one of the four categories. There were seven lessons in the original group of 58 lessons that utilized high cognitive demand tasks that did not meet the criteria for maintaining cognitive demand or having cognitive demand lower due to procedures without connections, unsystematic exploration, or no mathematical activity. This shows that these categories and implementation tactics are important to keep in mind, but do not necessarily show every possibility of what can happen when implementing high cognitive demand tasks.

Overall this article showed the importance of the teacher's role in keeping the cognitive demand of math tasks high. Teachers cannot simply give students complex tasks and expect them to be developing critical thinking, reasoning, and problem solving skills. Henningsen and Stein (1997) showed that teachers needed to know their students and provide appropriate tasks given their prior experiences and knowledge. Similarly, they showed that to keep cognitive demand high they needed to provide the right amount of time for the students to complete the task. This may require teachers to be more flexible with time and changing time schedules based on student progress. Finally, in both maintenance and decline the researchers showed that teachers can either press students to challenge themselves to justify answers, look for new approaches, etc or give students too much information or lessen the work of the high demand task. This article highlighted holding the balance of scaffolding for students without giving too much away and also pressing students to challenge themselves without leading to unnecessary frustration. Ultimately, high cognitive demand tasks can provoke deep mathematical thinking and the development of skills such as hypothesizing, problem solving, and critical thinking if implemented appropriately by teachers.

Stipek et al. (1998) completed a mixed methods study examining the similarities between motivation research and the practices recommended by mathematics education reformers. They also investigated in math classrooms the associations among teaching practices (including both motivation research and mathematics reform literature), student motivation, and mathematics learning. The researchers also predicted that classroom practices that enhanced mastery orientation would correlate with higher scores on the conceptual math problems on the post-test. Mastery orientation, or learning orientation, is defined as "developing skills, increasing understanding, or achieving mastery…focusing on learning and conceptual understanding instead of getting the right answers" (Stipek et al., 1998, p. 466).

The participants included 24 teachers that taught 624 students from fourth through sixth grade. The majority of students were Latino (358), while the rest of students were White (92), Asian (49), African American (46), or another ethnic identification (26). Out of the 24 teachers all but one were female, they had a range of experience ranging from one year to 20 or more years teaching, and all taught at large, ethnically diverse, urban school. The teachers were put into three groups, two of which consisted of teachers willing to commit to implementing reform-oriented curriculum, and previously taught a unit entitled *Seeing Fractions.* The reform practices are similar to those of complex instruction. The other group expressed no interest in reform methods and used traditional teaching practices. Teachers in the reform group participated in an intensive, yearlong intervention designed to help them implement reform-oriented teaching practices.

At the beginning of the school year, the researchers gave students two pre-assessments. A content test including procedural and conceptual questions about fractions was given to students to assess their beginning knowledge about fractions. The students were also given a survey assessing their attitudes and beliefs about math including perceived ability, mastery orientation, performance orientation, help-seeking, positive emotions, negative emotions, and enjoyment. The research team then videotaped at least two instructional periods for each teacher. The lessons were coded on the following nine dimensions of teacher actions: emphasis on effort, focus learning, understanding, and mastery (encourage use of alternative strategies, have students explain strategies, applying concepts in new contexts), emphasis on performance (grades, praising performance, avoid "hard" problems), encourage autonomy, frequency of social comparisons made by teachers, affect displayed by teacher, enthusiasm, and support of risk taking. Each separate videotaped lesson was given a separate rating and an average rating was calculated from the two. The researchers also had two independent raters for each lesson to control for rater error or biases. In addition to videotaped lessons, researchers also observed the teachers two times in person and took field notes from observations. These field notes were then coded similarly to the videotapes. The researchers compared the two sets of rating to make sure the rating of the teacher was valid. After the fraction unit was over, the researchers then administered a post-test similar to the pre-test except that on the beliefs survey they asked specifically about fractions instead of math in general. Lastly, the teachers completed a report on their grading and feedback practices during the unit.

The researchers found many results from the observations and data collected. They found that teacher learning orientation was positively correlated with positive teacher affect ($r = .68$, $p < .001$) and negatively correlated with differential student treatment ($r = -.56$, $p < .01$). They also found that positive teacher affect was also negatively correlated with differential student treatment ($r = -.64$, $p < .001$). Teachers that had high scores on learning orientation kept high cognitive demand, press for justification, and gave the students choices. They did not emphasize performance and instead focused on effort and persistence. High positive affect meant that teachers were genuinely kind and sensitive to students. They avoided sarcasm and did not allow put down in their class. The teachers expressed that they valued everyone's contributions to class and never threatened to call on students as a punishment or to show that they are paying attention. A teacher high in both learning orientation and positive affect are exhibiting strategies that are encouraged in the complex instruction teaching model. Similarly, teachers that reported giving students substantive written feedback were positively correlated with learning orientation ($r = .45$, $p < .07$) and positive affect ($r = .60$, $p < .01$). Substantive written comments were also correlated with students having a higher self-rating of their ability in fractions ($r = .42$, $p < .09$), more of a mastery orientation ($r = .67$, $p < .05$), and higher rates of positive emotions while working on fractions ($r = .52$, $p <$ 05). There were no significant differences of students' motivation from the beginning of the year until the end of the fraction unit, as measured on the pre and post survey. In a regression analysis, the researchers found that a teacher had positive affect (expressed positive emotions, encouraging students to take risks and seek help) the more students reported that they asked for help when they had questions ($p < .01$), were focused on learning, understanding, and mastering the material ($p < .10$), and experiences positive emotions themselves ($p < .05$) while learning fractions. Also as a result of regression analyses, they found that learning orientation (teachers focusing students' attention on learning, mastery, encouraging effort and autonomy) was a predictor of students experiencing positive emotions ($p < .05$) and enjoyed learning fractions ($p < .05$). Finally, they found that teachers' learning orientation was significantly associated with students gains on conceptually oriented questions on the post-test ($r = .51$, $p < .05$).

There were a few areas of critique within the design of the study. The researchers let the teachers select which curriculum group they were going to be in based on their willingness to teach the reform-oriented method. There was no mention of the achievement differences between the two different groups and if the teachers in one group were significantly different than the others. The researchers also did not identify how many teachers in the reform groups and traditional group were considered high learning orientation and positive affect. There is an assumption that reform-oriented teachers would have higher levels of learning orientation and positive affect, but those statistics were not reported. Also, the researchers found that there were no significant correlations between procedural learning and teachers' level learning orientation or positive affect. This was contrary to other research within the body of literature and what the researchers hypothesized. It seemed strange to the researchers that there were correlations between teachers' high levels of learning orientation and conceptual learning, but not procedural. The researchers explained that motivation (something that was not largely changed from beginning to end of study) may have more to do with procedural gains than teacher actions. Correlations were run between motivation and achievement on the pre and post-tests. They found that nearly all of the subsections of motivation were correlated with success on procedural questions on the post-test and none of the motivation subsections were correlated with conceptual questions. Finally, a number of main statistics the researchers were reporting were not significant at a .05 or higher level. The results of those findings may not be completely generalizable due to the significance levels.

This study emphasized the importance of teacher actions and dispositions and how that affects the students profoundly. Teachers that exhibited high learning orientations and positive affects impacted their students' enjoyment of the topic, conceptual learning, positive emotions, asking for help when they need it, and focus on learning and understanding (instead of getting a good grade or finishing quickly). The main teacher actions and dispositions involved in learning orientation and positive affect include keeping the cognitive demand of the tasks high, pressing students to give explanations for their answers, using wrong answers as learning opportunities, giving plenty of time for students to work, requiring students to revise problems or try it from a different approach, avoiding sarcasm, not allowing put downs in class, and valuing all student contributions. When a teacher conducts a class with these characteristics students are able to really dive deep into the material and learn the concepts behind the mathematics instead of memorizing a procedure. Teacher actions with high learning orientation and positive affect also parallel the classroom environment teachers using complex instruction are trying to facilitate. This becomes an environment were students feel safe to take risk to learn.

Kazemi and Stipek (2001) completed a case study that focused on how conceptual thinking and understanding are promoted. Researchers examined the different sociomathematical norms of math classes that heightened conceptual understanding.

Sociomathematical norms are specific social norms that are congruent to the domain of mathematics. For example, sociomathematical norms may include requiring students to justify answers with specific criteria. The researchers highlighted interactions in four classrooms and examined the classroom talk that is conducive to press for conceptual learning.

The participants in this study were four teachers teaching grades four and five. The lessons outlined in detail were all the teachers teaching the same lesson on the addition of fractions. The teachers were part of a larger study on reform mathematics in a project called the Integrated Mathematics Assessment. The teachers were from large, ethnically diverse school in urban California and students were primarily from low-income families. The four teachers all had experience implementing math reform curriculum. Ms. Carter and Ms. Martin taught at the same school, both had their master's degree, and over 20 years experience teaching. Ms. Andrew and Ms. Reed taught at two different schools. Ms. Andrew had a bachelor's degree and two years of teaching experience while Ms. Reed had a master's degree and 17 years of teaching experience. The lessons observed were videotaped and later coded originally by motivational strategies for improving student engagement. Kazemi and Stipek (2001) picked teachers for this study based on their "press for learning" and "positive affect" scores. Press for learning was determined from scores on teachers engaging students in mathematical thinking, emphasizing effort, focusing on learning for understanding, supporting student autonomy, and deemphasizing performance. Positive affect was found from scores on the social environment in the classroom including the teacher displaying a positive demeanor towards students, enthusiasm, interest in mathematics, and a nonthreatening environment where mistakes are viewed as learning opportunities. The researchers analyzed both whole class discussions and small group discussions, while looking at specific student and teacher dialogue. Ms. Carter and Ms. Martin were found to have a high press for conceptual learning, while Ms. Reed and Ms. Andrews did not.

The researchers found a number of similarities between Ms. Carter and Ms. Martin. Both teachers repeatedly asked students to explain their thinking. They explicitly asked students to link problem solving strategies to mathematical reasons, emphasizing sociomathematical norms. With the teachers asking students "why," students were pushed to make conceptual connections to procedures. Ms. Carter and Ms. Martin also asked students for multiple representations and ways to solve a problem. This included verbal, pictorial, and symbolic representations that are congruent with sociomathematical norms for justification. When students gave answers and explanations for their answers, Ms. Carter also had other students explain why they agreed or disagreed with this answer. Ms. Carter had her students also think of ways they could prove their answer was correct. This started conversations about what is proof and what counts as proving something, which is a very important sociomathematical norm. Teachers with low press allowed students to merely summarize their procedures and tell the correct answer. Students were not asked why they performed the steps they did. The teachers with low press also had students, by show of hands, who understood and agreed with the answer given. Students were not asked to tell why they agreed or explain the process in different words. The teachers also showed a lowering of expectations and were mainly concerned with students appearing on task. Finally, student-to-student interactions were missing in the low press classrooms. Students were interacting or responding to the teacher, but students were not commenting on each other's work. Thus conversations were with the teacher and not with other students.

The researchers carefully designed this case study. Kazemi and Stipek (2001) purposely chose the four teachers highlighted in this case study, two with high press and two with low press. The researchers were mainly focusing on specific classroom talk and teacher interactions that affected conceptual learning. Two coders were used in rating each videotaped lesson and an average score was used to categorize the lessons. The same lesson, addition of fractions, was also taught to show the different implementation techniques and teacher choices that affect conceptual learning.

Complex instruction is designed to engage students in conceptual learning. Through the work of Kazemi and Stipek (2001) it has become clearer the ways teachers can reinforce sociomathematical norms and interaction with students to promote conceptual learning. Teachers should work to always ask students to justify their answers, engage in conversations with other students, represent their answers and thinking in multiple ways, and include showing proof for answers. Students will make these practices a habit if continually reinforced from the beginning of the year by the teacher.

Khisty and Chval (2002) analyzed teacher talk and the important role it plays in student learning. The researchers wanted to show clear depictions of an example and non-example of teacher talk promoting equitable, conceptual learning opportunities. They were also showing more explicitly the reform NCTM is calling for, more communication in math classroom (NCTM, 1989; NCTM, 2000). The researchers also drew ideas from Discourse study and were looking for examples of teachers engaging students in "speaking mathematically."

The teachers examined in this study were Ms. Martinez and Ms. Tapia. Ms. Martinez taught fifth grade, with approximately 28 students. The academic ability levels varied, with the majority of students at least one grade level behind in math and some students qualifying for special education. Through standardized test score improvement, she was shown to successfully help students make growth from the level they started. She was also part of a bilingual education program, but her lessons were taught in English. Students in her class were working collaboratively and practicing oral and written communication, justifying answers, learning for understanding, and thinking independently. She strongly believed in students constructing their own ideas and knowledge through her scaffolding and questioning strategies. Finally her lessons were typically longer than other math teachers in the school due to lengthy discussions and amount of time she allowed students to work on arriving at solutions. Ms. Tapia was a highly respected and competent teacher that was recommended by a principal to observe. Ms. Tapia had only been teaching for a few years and at the time of the study taught second grade. She had participated in a NCTM reform project and she typically set up her classes so that she gave a mini-lesson to the whole class and then allowed the children to work with others or work independently. Her lessons were taught in Spanish and in the observed lessons were connecting math and literacy.

The researchers observed Ms. Martinez extensively through videotaping and audiotaping. They observed Ms. Martinez for the whole school year, physically observing her and taking field notes two or more times a week. While researchers took field notes, they also audiotaped the lesson for later reference. They videotaped her class at least two days a week for a four month period. Many different researchers have taken interest in Ms. Martinez and Khisty and Chval (2002) pooled all the various audio and video footage from other researchers for her analysis.

They found a number of interesting results from studying Ms. Martinez. First they found that new concepts were introduced in a number of attention grabbing and interesting ways. The new concepts were contextualized with "word problems, drawings and models, peer discussions, and even physical activity" (Khisty & Chval, 2002, p. 7). Also Ms. Martinez modeled the use and pronunciation of new mathematical terms. She continually repeated the new vocabulary and even had student practice saying the words. She provided pictures, definitions, and allowed students to think of examples or non-examples of the new word. She also helped students understand the connections between new vocabulary and concepts they are learning about and already know. When introducing new vocabulary she accessed students' funds of knowledge. In one example she introduced the word quadrilateral and capitalized on the students' prior knowledge of Spanish to help come up with a definition of the new word. She was very mindful that students needed to know certain vocabulary words to be successful on standardized tests and high-level math classes. Ms. Martinez engaged in bridging the language for students and helping them make connections between their ideas and the precise vocabulary through questioning and asking them to explain their reasoning. She also accepted all student responses in group discussions, but typically rephrased what they said in more precise terms or complete way. She then reinforced this by asking students more questions on the same topic, looking to get more precise answers. Ms. Martinez also very strategically gradually released the responsibility of dialogue in class discussions. At the beginning of the school year she talked most of the time, but by the end of the school year students engaged in talking to each other. She spent much time in her class modeling and helping students practice academic math talk. She reinforced the sociomathematical norms and academic math talk so much throughout the year that at the end of the year students knew what and how they were expected to communicate.

In Ms. Tapia's class the observers quickly saw differences in student understanding of topics. Students in her class seemed disinterested and did not know the deeper concepts behind the procedures they were mimicking. Also, Ms. Tapia rarely engaged the whole class in discussions. She posed few questions in her lessons and typically they did not require mathematical thinking, simply memorization. The researchers also noted that she did not model mathematical talk to the class so that students were not made aware of the precise academic math language.

There are a number of strengths and weaknesses within this study. The researchers did not include specific standardized test scores to show improvement in Ms. Martinez's class or deficiency conceptual understanding in Ms. Tapia's class in efforts to use triangulation to make their cases stronger. The researchers did provide specific dialogue with each analysis to show in detail how the teachers were conducting their classes.

From this study, Khristy and Chval (2002) showed the importance of the teacher in modeling sociomathematical norms and regular use of mathematical talk. Teachers, like Ms. Martinez, can use students' fund of knowledge, gradual release of dialogue control, and bridging language to help students learn deeper and in a more equitable way. By teaching students how to communicate within the domain of mathematics and express their conceptual understandings, they have a better chance of showing what they truly "know" on standardized tests. Also by having students continually justify answers, they develop a sense of what a complete explanation is in mathematics. Within complex instruction classroom, teachers facilitate whole class discussions typically when introducing a concept and also at the end of a group task. Teachers need to successful model precise math language and teach students how to communicate in mathematics so that student discussions during group work is of high quality and students can practice this new form of communication.

Arbaugh and Brown (2005) completed a qualitative study in which they analyzed the affects of teacher education and professional development on levels of cognitive demand and high-level mathematical tasks implemented in their classrooms. They were looking at the choices of mathematical tasks used in the classroom before and after learning about levels of cognitive demand. Due to reform movement, more teachers have been receiving profession development about concepts such as cognitive demand. The researchers were interested in looking at the impact of these workshops and in-service education have on the actual tasks being used in the classroom. The official research question of the study was, "How did learning to examine classroom tasks critically influence the teachers' thinking about mathematical tasks as well as their choice of tasks used in the classroom?" (Arbaugh & Brown, 2005).

The participants in this study were geometry teachers at Ericson Valley High School participating in a Toyota Time Grant to support professional development. The teachers identified that they were interested in focusing the development on pedagogy and curriculum. The teachers were eager for professional development to help them implement reform-oriented teaching strategies in their classrooms. Only seven of the nine original teachers from Ericson Valley High School were included in the written report of the study. The teachers' experience in teaching ranged from one year to 32 years. Two of the teachers used reform oriented geometry textbooks, while the other five used more traditional geometry textbooks. The teachers met about every two weeks from October 1999 until March 2000 for a total of ten times. For the first two meetings, teachers learned about levels of cognitive demand and completed activities to solidify the concepts. In the following meetings, teachers practiced assessing the cognitive demand of tasks and applied that knowledge to assessing tasks they used in their own classes. The teachers also read journal articles and research on cognitive demand and discussed ways to implement what they were learning to their classes specifically. Teachers were interviewed both at the beginning of the professional development meetings and when the meetings had concluded. The interviews were audiotaped and the researchers also took notes during the interviews. As part of the interview, teachers were asked to sort math tasks by their cognitive demand levels. The teachers also gave the researchers copies of all the tasks given to students during the first weeks and last weeks of the study. The researchers then coded the interviews based on the work of Strauss and Corbin (1998) to conclude what the teachers learned about levels of cognitive demand criteria, including their accuracy with the task sorting activity. Tasks given to students by the teachers in the study were also analyzed and categorized based on their cognitive demand.

After comparing the tasks teachers used in their classroom at the beginning and the end of the study, the researchers found little change in the levels of cognitive demand. Thus researcher found no significant change for the group of teachers in the characteristics of the tasks they assigned. The group of teachers overall increased the number of tasks they assigned their students (390 before study, 540 after study) and respectively higher numbers of high-level tasks (91 before study, 130 after study). Although there were a higher overall number of high-level tasks implemented, the percentage of high-level tasks used did not change (23% before study, 23% after study). Despite the overall small improvement made in use of high-level tasks, some teachers individually had large improvements. Brian, the math department head at Ericson Valley High School during the study, had the largest increase of high level task use (32% before study, 100% after). These results may be due to the classroom unit schedules and when the tasks were submitted. This teachers was giving a test and review (with low cognitive demand) during the "before study" task analysis and was teaching a new concept (all high cognitive demand) at the "after study" task analysis. It was also found that teachers using the traditional textbooks used more supplemental and enrichment materials during the "after study" task analysis.

The researchers ran into some issues of attrition with the geometry teachers in this study. Two of the teachers were not included in the analysis, lowering the sample down to seven teachers. Also one of the seven teachers participating in the study did not have data collected during the "after study" analysis. To make the finding of this research more generalizable, further research is needed with a variety of populations and larger sample sizes. The researchers also do not look at how student achievement is impacted or teacher questioning or interaction during the implementation of the tasks. As Henningsen and Stein (1997) point out, just because student are given high-level tasks does not mean they are engaging them in a high-level way.

Arbaugh and Brown (2005) do have important implications for implementing complex instruction. Teachers are to be choosing and implementing high-level tasks within the complex instruction model. It is important that teachers have support while implementing new teaching technique and strategies. Even teachers that attended 10 professional development workshops on cognitive demand and were committed to reform-oriented math methods still had a hard time implementing and choosing high cognitive demand tasks. It could be beneficial for teachers going through reform together to use each other as resources to determine the level of cognitive demand of tasks used in the classroom regularly. Similarly, teachers could observe each other to determine whether high cognitive demand tasks are actually eliciting high levels of conceptual thinking in the students. Specifically for complex instruction, it is important for teachers to give students tasks that are high level tasks and that are deemed group worthy (Lotan, 2003). If teachers are not actually implementing high-level tasks, this diminishes the need for interdependence and collaboration. It is vital that complex tasks are given to students within the complex instruction model.

Through analyzing literature on complex tasks, conceptual learning, and cognitive demand we see that these concepts are very much interconnected. Henningsen and Stein (1997) found that teachers needed to be very mindful of appropriateness of the task, pressing for justification and explanation, scaffolding, giving the appropriate amount of time, and providing tasks that build on students' prior knowledge. Stipek et al. (1998) found that teacher positive affect and learning orientation can impact student conceptual learning as well as enjoyment of the topic, positive emotions, asking for help, and a focus on learning and understanding (instead of getting a good grade or finishing quickly).

Kazemi and Stipek (2001) found that teachers can reinforce sociomathematical norms through their interactions with students such as emphasizing justification of answer, engaging in whole class conversations, requiring students to represent answers in multiple ways, and asking for student to show proof for their answers. Khisty and Chval (2002) found that teachers' talk plays an important role in equitable student conceptual and deep learning. Arbaugh and Brown (2005) found that despite teachers learning about levels of cognitive demand, it is hard to make consistent changes in mathematical task choices and takes much reflection on the part of the teacher. Within the complex instruction model, the strategies found through the research outlined above are important to keep the classroom a place where deep learning is taking place to the full potential.

Heterogeneous Group Work

Another key facet of complex instruction is regularly implementing heterogeneous cooperative group work. Schullery and Schullery (2006) found that there are both advantages and disadvantages to heterogeneous and homogeneous groups. These researchers specifically looked at personality differences, not differences in ability. Webb, Nemer, and Zuniga (2002) looked at how grouping affects high ability students' achievement. Boaler, Wiliam, and Brown (2000) researched student reactions and achievement changes as they moved from heterogeneous, mixed ability groups to ability grouping in math classes. They found that students overall responded negatively to the change in grouping criteria. Cheng, Lam, and Chan (2008) found that the quality of heterogeneous group processes were that main determinant in students' learning efficacy and achievement. Mitchell et al. (2004) studied the effects of student-selected groupings versus teacher-selected. They found that friendship and status strongly influences partner choices and that students feel obligated to choose friends despite who they feel they should work with.

Schullery and Schullery (2006) completed a quantitative study in which they were determining the advantages and disadvantages of participation in heterogeneous group work. They were interested in personal growth, collaboration, promoting discussions, and improved group work skills. They chose to focus on heterogeneous or homogeneous groups based on personality tests not academic ability.

The participants were 16 high school business classes, 13 junior-level Business Communication classes and three sophomore-level Organizational Communication classes (394 students total; 238 males and 156 females). There were approximately two percent and 10% of students that were African American and international (majority of Asian decent) respectively. Students in these classes rarely worked in collaborative groups. When working with others, it was typically brief and included in-class activities only. Students in the study were in groups ranging from two to six students (average 3.8 students). Personality tests were given to each student at the beginning of the study and included a shorted version of the Jungian-Meyers-Brigges Type Indicator and Infante and Rancer’s 20-item argumentativeness communication trait assessment. Heterogeneous groups were students that were opposites as determined by the personality tests. Homogeneous groups were students that were similar or identical on the personality tests. The researchers tried to evenly distribute minority students and balance for gender differences. Researchers did not report how many homogeneous groups and heterogeneous groups were each included in the study and if there were any groups that were neither homogeneous nor heterogeneous. Schullery and Schullery (2006) also measured students' perceptions about group work before the study and after using a self-assessment that focused strengths and weaknesses and on the final assessment, areas of improvement as pertaining to their role in group work. The researchers also measured students' perceptions about their individual skill improvement before and after the study. The researchers included measurements of grades on projects, teacher's perception of group functioning, individual satisfaction with group, and group and individual's total improvement.

The researchers ran Spearman's rank-order correlation analysis and multiple regression analysis to come to their results. The researchers found that there were mixed results for both heterogeneous and homogeneous groups. In highly heterogeneous groups, students reported greater improvements in speaking up or arguing a point, organizing and presenting thoughts (females), and overall individual skill improvement (females) and group skill improvement. In highly homogeneous groups, students reported greater improvement in shyness (males) and earned higher grades. There was no correlation between the teachers' rating of group functioning and type of grouping (heterogeneous or homogeneous). Students in heterogeneous groups reported improved group communicative behaviors. The researchers concluded that students in heterogeneous groups made large improvement in a few areas (mostly pertaining to group communication) instead of improving a small amount in many areas.

The researchers have a number of areas of concern within the design of this study. The sample population of this study was students from business classes in a high school. Business classes tend to be electives and there were low levels of ethnic diversity (and females). The researchers did not assess academic ability of the students and used two personality tests to determine groups as "heterogeneous" or "homogeneous." Within the literature of student grouping, when referring to heterogeneous or homogeneous, researchers are typically meaning mixed or similar ability grouping (Webb, Nermer, & Zuniga, 2002; Cohen, 1994). Within this classroom, the range of ability was not assessed. Also, students were not used to working in group settings and did not have class time to complete the entire project. Students were not given specific roles within the group or taught how to equitably work together. Thus student who reported on the personality tests to be more outspoken and argumentative may have dominated group work time leading some students to not participating equally. Also the students were self-reporting their perceptions of group work and improvements, which can easily be effected by social desirability. Finally, the researchers wording on their assessments were very negatively worded when dealing with group work such as "speaking up or arguing a point," "public speaking anxiety," "impatient with slow progress," and "resent slackers and distrust of other." This type of wording and seemingly lack of counter-balancing could skew the view of group work in general of the students, especially since the participants were not used to working with others.

Despite the many methodological areas of question, there are still some interesting aspects of group work for teachers to keep in mind. One of the finding involved females working better and having more improvements within heterogeneous groups. As a teacher implementing any type of group work, a continual analysis of student achievement and reactions to the in class environment needs to be made so that particular groups are not being held back or left out. Similarly, this study highlights the importance of teaching students how to work together and address issues of status. This study also had students work on a project mostly outside of class. Teachers need to know their students and determine who would be successful on a group project outside of class and who would not. Some students may have after school jobs, extracurricular activities, or other obligations that may inhibit their participation in an out of school group project. Also, teachers do not have as many opportunities to assign status to "low competence" students when they are working together outside of class. Thus teachers impacting the equity of group work using complex instruction techniques are more limited.

Webb, Nemer, and Zuniga (2002) completed a quantitative study in which they were looking at high ability student performance in both heterogeneous and homogeneous groups. They were focusing on co-construction of task solutions, helping behaviors, socioemotional processes occurring in groups, and student contributions to group discussions. They did not include low ability homogeneous groups in their analysis, only heterogeneous groups (including at least one high ability student) or high ability homogeneous groups.

The participants in this study were 81 high ability students from two schools. The high ability students were primarily White (89%) with small numbers of African American, Latino, and Asian students (2%, 2%, and 6% respectively). The students in this study were all given three pretests including a vocabulary, verbal reasoning, and nonverbal reasoning test. They all had similar three-week units on electricity and electric circuits. At the end of the unit, students were given a hands-on test and a paper-and-pencil test to be completed individually. The students then were given the same end of unit hands on test to complete with a collaborative group one month later without any interventions. The following day, students completed individually the same end of the unit paper-and-pencil test. Some high ability students were groups homogeneously, some were grouped with mid-level ability students, and others were grouped with low-ability students. All groups were heterogeneous based on gender and ethnicity. The researchers videotaped the group interactions and coded the group work based on right answers given, the correctness of the justification of the answer, high ability students giving help, high ability students asking for help, high ability students giving correct and complete answers, and the frequency of negative socioemotional behavior.

The researchers found that high ability students in the homogeneous group out-performed high ability students in the heterogeneous groups. There was no significant difference between the performances of high ability students between the different heterogeneous groups (low, low-medium, and medium-high). The variability of performance was higher for the high ability students in the heterogeneous groups when compared to the high ability students in the homogeneous group. High ability students' performance on the after unit test was positively correlated with both giving and receiving help. Negative socioemotional behavior predicted high ability students less frequently completed work fully and less help high ability students received. High ability students in heterogeneous groups that had high performance on the group task, had groups that exhibited more behavior that positively related to performance scores such as more accurate justifications, more volunteered correct suggestions, higher-level help, more often asked for more information during an incomplete explanation, and showed little negative socioemotional behaviors.

There are a variety of areas that were not considered when the researchers set up this study. The tasks given to the groups were originally designed to be completed by a single person. Thus these tasks do not require collaboration and are not deemed group worthy according to Lotan's task requirements (2003). The students had already seen the task and knew how they scored on the task, creating self-expectations for competency. Also the hands on task had definite right and wrong answers, focusing on procedural knowledge instead of conceptual. The researchers also do not explain how the teachers taught the original electricity units. The researchers also remarked that there were no teacher interventions during the group work time. That is, teachers did not help students work together equitably, give students roles, or attend to status. The researchers did have two coders for each videotaping. However, only certain groups were videotaped and the researchers did not explain how they chose specific groups to record. Finally, the sample size of the groups was fairly small and the number of students analyzed through videotaping is even smaller.

Overall, the results of this study are not highly generalizable to random, heterogenous groups working on complex, uncertain, open-ended tasks. This study does show that when students are trying to remember detailed, procedural knowledge, a group of students who all effectively memorized the information do better than a group of students that had mixed levels of memorizing the information. In math classes, students do at times need to remember facts, skills, and procedures. However, the bulk of mathematics done in classrooms need to reflect the domain of mathematics, which includes hypothesizing, justifying answers, representing information in multiple ways, connecting concepts with procedures, etc (Kazemi & Stipek, 2001; Lotan, 2003; Resnick & Zurasky, 2006). Also, complex instruction includes tasks that are multidimensional and cannot be completed by one person alone. The tasks in this study were originally made for one person to be able to memorize information and complete on their own. This article does not show that high ability students would be hindered by working in heterogeneous groups in a complex instruction setting. Webb, Nemer, and Zuniga (2002) inadvertently emphasize the need for complex group worthy tasks, group roles, and status treatments to make group work beneficial for both low and high ability students.

Boaler, Wiliam, and Brown (2000) completed a longitudinal mixed method study to understand how ability grouping students in high school affects their views of math class and academic achievement. In light of much research done in the US on the achievement of high ability students, they wanted to show real student voices from the high level groups as well as the low level groups to show a full picture of what is going on in the classroom, not just a standardized test score.

The participants in this study were students in six schools in the Greater London Area. These schools represented a wide range of learning environments and contexts. The schools were all regarded as having satisfactory or good standing in mathematics due to standardized test scores. Some of the school populations were inclusive of a wide range of ethnical and cultural backgrounds, while other schools were mainly white or Asian. All six schools in the study taught mixed-ability classes when the students were younger. Starting after seventh grade, schools started to categorize students into different groups with four of the schools in the study having ability grouping by ninth grade. Two of the schools maintained mixed ability and these student responses were viewed as a comparison for the other four ability grouped schools. The study was completed at the end of ninth grade so that the four schools of ability grouped students had all been in the "set" for at least a year. The researchers observed classes, gave students questionnaires, and also interviewed students. They collected over 120 hours of observation in math classes within the school, averaging about 10 hours of observation per school. The researchers administered questionnaires about attitudes and beliefs about math at the end of both eighth and ninth grade. The questionnaire included open closed items with approximately 843 students taking it both at the end of eighth grade and ninth grade. The researchers also interviews 72 pairs of students at the end of ninth grade as well. The interviews were of girls and boy from the top, middle, and lower sets. Each interview was about 30 minutes long and was facilitated using 18 predetermined prompts.

The researchers found a number of interesting results. The found that of the 48 students interviewed from the ability groups schools, 40 of them either wanted to return to mixed ability classes or to change the "set" they were currently in. Students in the high set described their classes as high pressure, high expectations, and moving at a hurried pace. One-third of the students in the high set remarked about how the pace is too fast. Every one in eight girls interviewed wanted to move to the next lower set, while six out of eight of the boys interviewed were unhappy with their placement. The students commented that the classes emphasize speed, memorization, and following procedures without conceptual explanations. Students in the high set also frequently said that they teachers used their high set status as a reprimand when students asked questions or did not understand. Examples students gave, and observers actually heard, were "You should be able to, you're in the top set" and "You are the set one class, you shouldn't be finding this difficult." Teachers also frequently made comments about not having much time to cover all the material and hurry or work more quickly. When asked if they enjoy math, 43% of top set students answered either "never" or "not very often" as compared to 36%

in other sets and 32% in mixed ability classes. Also, 68% of high set students answered that memorization was more important than thinking. Finally, despite media claiming that girls are outnumbering boys in advanced classes, the researchers found that boys still outnumbered girls in earning an “A” in the high set math class. Within the low set students, the researchers found a very strong message, low expectations and limited opportunity. The students commented on the frequent change of teachers, with one school changing the low set teacher three times in nine months. The students also complained about having very low level work such as copying already solved problems off the board. Students commented that they are not challenged and if they do finish the work early, they sit and wait for the others to finish. From the questionnaires, the researchers found that 27% of students in the low set thought the work was too easy. The students revealed their frustration and annoyance at the level of work given daily. They frequently talked about math being boring. While the schools that use ability grouping told the researchers that the groups are flexible, students to change ability groups must test into the higher classes. With students not receiving the same curriculum, resources, and learning experiences, the researchers commented that students will likely not be able to compete with students in higher sets, especially students deemed in the low set. The students were very aware of their set status and teachers in the low set classrooms even made comments such as, “You’re the bottom group, you’re not going to learn anything.” The researchers noticed that teacher changed the pedagogy when they viewed their classes as similar in ability. Teachers in the schools with ability grouping resorted to book teaching, lecturing, and “teaching from the chalkboard.” Teachers in the mixed ability groups still taught students at varying paces and allowed students more freedom in the pace of their work. The teachers of ability grouped classes seemed to teach to a middle line expectation and did not even cater to different learning styles or other differences. Students were seemingly viewed as identical since they were all in the same ability class.

There were a number of strengths and weaknesses within this study. The researchers administered the interviews and the questionnaires. They assured students that their responses were going to be kept anonymous and confidential. The teachers selected students to be interviewed who were aware of the focus of the research and more relaxed and glad to talk. The teacher's choice of students may have skewed the responses of the students. Teachers could have chosen students that they believed were the best in class or would more likely talk highly of them. The researchers also had two schools provide as a comparison group. Many of the claims they are making may be explained that as student get older teachers change how they interact and teach the students. With the comparison group, the researchers were able to show that at least in this group that was not a viable explanation. To make further generalizations, a larger sample size with more heterogeneous groups would need to be considered. The researchers did not include standardized test scores, grades, or achievement levels to supplement their detailed class descriptions. The researchers did use observations, student interviews, and previous research to make sense of the findings. Using this form of triangulation made the results more reliable and was more reflective of the school instead of a couple students' experiences.

From this study, Boaler, Wiliam, and Brown (2002) have shown that ability grouping can have very negative impacts on a wide variety of students. While some subjects are still mixed-ability, due to the cumulative nature of math classes, many schools have ability grouped math classes. As this article shows many students in the lower groups are not getting equitable educational experiences when compared to the high level students. Also, teachers in the high level classes seem to be missing the point of learning to understand and have deep conceptual *and* procedural knowledge. This shows that schools need to carefully implement ability grouping or mixed ability classes as they impact student beliefs about the subject, teaching pedagogy, and overall enjoyment of the subject.

Cheng, Lam, and Chan (2008) studied high ability and low ability students working on academic tasks in small, heterogeneous groups. They were interested in achievement levels as well as self-efficacy, collective efficacy, and group processes. The researchers hypothesized that group learning would be more effective when the groups interacted with each other on a more equitable and cohesive basis. They speculated that the quality of group processes and student achievement would both be positive predictors of efficacy. Finally, the researchers hypothesized that the interaction between student achievement and group process would affect efficacy. The researchers also took into consideration group size and gender composition of the groups studied.

The participants were 1,921 students (49.9% females, 50.1% males) in seventh, eighth, and ninth grades. Students were from eight secondary schools in Hong Kong each from a different school district. The students varied in socioeconomic backgrounds and academic standards. All the students in the study were enrolled in classes that were project-based learning as part of a new reform in Hong Kong schools. The tasks students worked on in small groups were open-ended, ill-structured problems of diverse and multidisciplinary topics. Teachers had a number of ways to assign students to groups, with some teachers randomly assigning groups, students choosing their own groups, or students choosing a partner and then joining other pairs to form a group. All groups were ability heterogeneous and had not chosen based on ability. Of all the groups formed in the study, there were all male groups (29.7%), all female groups (28.1%) and mixed-gender groups (42.2%). Group sizes ranged from three and seven students with the average group size five or six students. Teachers attended training workshops before implementing the project-based learning. Teachers varied in their levels of supporting group process and the groups varied in the quality of group functioning.

Researchers administered questionnaires to students after they had been working with their groups for an extended period of time. The questionnaires measured group processes and efficacy. The questionnaires had students rank their agreement with statements on a six-point Likert scale. Group process was measured by statement pertaining to positive interdepence, individual accountability, equal participation, and social skills. Efficacy was broken into self-efficacy and collective efficacy. Student achievement was also measured from a mid-term examination administered in the school. Group heterogeneity, gender composition, and group size were measured and factored into results. Cheng, Lam, and Chan (2008) used hierarchical linear modeling to perform multi-level analyses. The researchers found a positive correlation between self and collective efficacy. They also found that higher achieving students had lower collective efficacy than self-efficacy and lower achieving students had higher collective efficacy than self-efficacy. Also, groups that had higher quality group processes had higher collective efficacy. The researchers also found that both high and low achievers had higher collective efficacy than self-efficacy when group processes were of higher quality. The researchers also found that gender composition, heterogeneity, and size did not influence collective or self-efficacy.

There were a number of strengths and weaknesses within the study. The sample size of this study is very large and inclusive of many schools, teachers, and school districts. However, within the sample there is very little ethnic or cultural diversity. Also, the researchers focused on differences between collective and self-efficacy after substantial time working in groups. There was no pretest given to students to measure change in efficacy due to group work. The hypotheses of the researchers were supported and the results were congruent to previous research. The researchers also rated group processes by student reports and not third party observations. Similarly, efficacy was measured by student self-reports.

There is a main conclusion from this study. The researchers explained that teachers should not put students together in groups and expect them to work well with each other. There are a number of aspects, positive interdependence, individual accountability, equal participation, and good social skills, that teachers need to help students develop as they work within group situations. These group processes also minimize social loafing, which is a negative factor of implementing group work. Within complex instruction, teachers need to build students group work skills by giving them opportunities to practice working with others and also take on specific group roles to help make work more equitably distributed among students.

Mitchell et al. (2004) completed a mixed methods study analyzing the effects of letting students choose their own groups to work with in high school science classes. The researchers were interested in best grouping practices for high school students and also the social and psychological implications for students to choose their own groups. The researchers predicted that students choosing their own group members would have a negative effect on later attitudes about group choice. They also speculated that students would choose friends when given the opportunity.

The participants in the study were 139 students (54 females, 85 males) that were all taught by the same science teacher. Students were high school sophomores and juniors from a small school in Montreal, Canada. The students had varying levels of achievement levels, socioeconomic status, cultures, and ethnicities. The teacher taught five ability-grouped classes of science including two low-achieving classes, two normal-achieving classes, and one honors class. The study was conducted from January to February and ran for a total of six weeks. Students at the beginning of the study completed the "Classroom Life Scale" that measures attitudes about group work including attitudes on cooperativeness, feelings of alienation, academic self-esteem, academic support, goal and resource interdependence, external motivation, grading practices, cohesion, etc. As part of the initial survey, students were also asked about preferences of choosing group members. Within the low-achieving and normal-achieving classes, the teacher selected groups for one class and had self-selecting groups for the other. The honors class, since there was only one class, experienced both teacher-selected and self-selected groups within the six weeks. During the study, the classes were broken into two labs, each running for three weeks. Students were in one group of three for the first three weeks and then switched groups for the final three weeks of the study. In the teacher-selected condition, the teacher selected different groups for each of the two labs. For the self-selected condition, students had to switch at least one of their three group members. The honor class had teacher-selected groups for three weeks and student-selected groups the remainder of the study. Within the group work, the teacher assigned rotating group roles, assigned each group member specific tasks, and gave both individual and group grades. The students completed the same "Classroom Life Scale" measure after the study was completed. The researchers also conducted a focus group, with the variety of students present in the study represented. Mitchell et al. (2004) wanted to check the results with the students to further have them explain how different groupings affected them.

The researchers had a number of interesting results. They found through a repeated measures multivariate analysis that students in the self-selected and the teacher-selected/self-selected groups had negative changes in preferences in choosing group members. They found that the low-achieving students had a more dramatic decline in preference in choosing group members and female students overall had a sharper decline in preferences in choosing group members. The change in attitudes for teacher-selected groups and self-selected groups did not change overall preferences for cooperative learning. Attitudes towards cooperative learning were actually higher in all groups in the posttest. In the focus group, researchers found a much more complicated explanation for the results. Students liked feeling in control of their environment and choosing their own groups. From a developmental standpoint, high school students are seeking autonomy and wanting to do "adult" activities and behaviors. Students also admitted that in self-selected groups there was a tendency to socialize rather than get work done. Focus group students also talked about it being harder to tell people they do not know well to get to work than it is friends. Students also remarked about the difficulty in deciding who to work with and who will make the best partners. Female and male students also exhibited gender stereotyping such as male student remarking they are more assertive and can get to work, while female students talked about the fear of hurting others feelings. Students in the low-achieving class also talked about how friends may not be effective group members and that they wanted to learn from someone who knew more then them, namely the teacher. The researchers pointed to these comments as indicators that tracking negatively impacts low-achieving students and place them at a disadvantage. Further they explained that a benefit of working within a group is learning from a "more capable peer" and that aspect of group work can be lost in tracked classrooms.

The researchers within their mixed methods analysis worked to control for variance. The researcher implemented pre and post tests with comparison groups in the low-achieving and normal-achieving groups. Researchers told students that their results on the "Classroom Life Scale" were kept confidential, minimizing social desirability to please the teacher. The teacher did not administer the focus group and the students were told that their identity would be kept confidential and their responses would be anonymous. The researchers also performed member checks with the students by conducting the focus group to help them further make sense of the results and data collected. Since the researchers only performed this study with one teacher in a single school, the results may not be generalizable to all self-selecting and teacher-selecting group situations.

The researchers did have a number of interesting conclusions from this study. This study showed the students from a developmental perspective and highlighted the need for allowing students to have control of their learning environment. The researchers also suggested that teachers should train students how to work in small groups effectively, emphasizing group roles, group work norms, increased responsibility, and developing facilitation skills. The main suggestions the researchers made are in line with complex instruction teaching methods. Mitchell et al. (2004) also suggested that students should work together for longer period of time to make more connections and really get to know each other. Finally, the researchers emphasized the need for students to feel responsible for their learning environment. When teachers are implementing complex instruction they can give groups choices by offering different activities, giving options when possible, and occasionally allowing them to choose their own groups.

The literature within this section points to the necessity of following the complex instruction model of teaching students how to work together, giving students group roles, releasing some control to students, and having mixed ability groups when working on complex tasks. Schullery and Schullery (2006) found that when groups are not taught how to work together and are not giving in class time to complete the project, personality differences can make the group work less effective academic. Webb, Nemer, and Zuniga (2002) looked at how grouping affects high ability students' achievement. They found that when a group is working on a simple memorization task and are not taught how to work equitably together as a group, homogeneous groups may be more successful. Boaler, Wiliam, and Brown (2000) researched student reactions and achievement changes as they moved from heterogeneous, mixed ability groups to ability grouping in math classes. They found that teachers treated ability grouped students differently and used different teaching strategies then teachers in the mixed ability groups. Students were also less likely to enjoy math and more often felt the material they were learning was either too fast or too slow. Cheng, Lam, and Chan (2008) found that the quality of heterogeneous group processes were that main determinant in students' learning efficacy and achievement. Mitchell et al. (2004) studied the effects of student-selected groupings versus teacher-selected. They found that friendship and status strongly influences partner choices and that students feel obligated to choose friends despite who they feel they should work with.

Status and Competency Expectations

Issues of status and competency expectations are vital facets of complex instruction and were the driving factors for the advent of complex instruction. Cohen and Lotan (1995) showed the importance of frequently utilizing status treatments like multiple ability treatments and assigning competence to low status students. Tammivaara (1982) studied reading ability status and the affects of multiple ability treatments. Ross and Cousins (1994) unintentionally showed that status is very much at work in cooperative learning situations where status is not addressed. Cohen and Anthony (1982) showed that rates talking and working together are determined by status and not academic ability. Cohen and Roper (1972) were interested in creating equal status group dynamics between White and African American males. They applied status treatments and assigning competence to find a more equitable working situation.

Cohen and Lotan (1995) completed a quantitative study in which they tested the affects status treatments had on student participation in mixed status groups. Two of their hypotheses focused on the individual students and the third hypothesis looked at the classroom as a whole. The first hypothesis speculated that the rate of participation of low-status students in mixed status groups will be positively related to the frequency of the use of status treatments by the teacher. The second hypothesis proposed that high-status student participation will not change due to the status treatment frequency of the teacher. The final hypothesis was that at the classroom level, the teacher's used of status treatments will negatively relate to the effect status plays on classroom interaction and participation. Status treatments are defined to include both multiple ability treatments and assigning competence to low-status students. Multiple ability treatments were given before a task and included a teacher reminding the students that the tasks required a variety of skills, listing the specific skills required, and saying, "None of us has all these abilities; each one of us has some of these abilities." Assigning competence in this study, similar to complex instruction, was the act of a teacher using his or her high-status by positively bringing attention to a student of low-status. A teacher would observe a low-status student in the group and call attention to something he or she did that was outstanding. The teacher's remark was specific, related to the task, and connected to important aspects of the task. Also, teachers asked low-status students to explain and show their work to the other group members if they were doing it correctly, leave the group, and check back in a couple minutes to ensure the sharing took place.

The participants consisted of 13 classrooms of second through sixth grade students. The schools were located in the San Francisco Bay area. A large proportion of the students in the study were from low-income and language minority backgrounds. The classrooms at these schools were segregated, including two classes that were composed of Southeast Asian immigrants with limited proficiency in English, nine classes were predominantly Latino students, and the remaining two classes were a mixture of low socioeconomic status White, Asian, and Latino students. Each class still maintained academic heterogeneity, including a range from students performing at grade level to student that did not have the ability to read and write in any language. The thirteen teachers that participated in the study were given a two week training at Stanford University in the summer of 1984. During the training teachers learned about complex instruction and the Finding Out/ Descubrimiento (FO/D) curriculum for math and science. Some of the teachers received extra training in assigning competence to low-status students with the hope that the strategy would be used more often in the study.

The researchers first measured the perceived status of the students in the study. The researchers asked each child to circle the names of those in their class that were "best at math and science" and then to circle those that were their "best friends." They then used this information to assign students a costatus score ranging from two to ten. From these scores the researchers targeted all the students that scored two or three and nine or 10 (low and high status students respectively) and then made random selections from the students that scored from four to eight. They targeted approximately 14 students from each class. Cohen and Lotan (1995) then sent researchers into the classroom to observe the target students working in mixed-status group situations. The researchers observed each child at least six times. During each observation, the researchers charted the target students' dialogue in six 30-second intervals. They categorized the dialogue during the observation into task-related talk and non-task related talk. Task-related talk included procedural talk about group roles, cooperation, or acting as facilitator. The interobserver reliability on the target children measure was 92.93%. There were high rates of students missing school or changing schools, so some targeted students were not able to have a complete set of observations conducted. In addition to measuring student interactions, the teachers were also observed. During 10-minute observation periods, the researchers made note of either the teacher talking about multiple abilities and/or assigning competence to a low-status student. There were 285 observations collected with each teacher observed 17 times or more. The researchers combined both status treatment types to be used as one treatment variable in testing the hypotheses. The researchers controlled for the overall level of interaction promoted by the teacher and extreme heterogeneous academic skills.

In the analysis, the researchers used the information gathered about teacher implementation of status treatments as the independent variables and the information gathered about student interactions as the dependent variables. They found that on average, high status students participated at a significantly higher rate than low status students ($t = 1.81$, $p < .05$). They also found that the rate at which low status students talked or behaved like a facilitator had a positive effect on his or her participation rate, but was not statistically significant. Through a regression analysis, the researchers found that status treatments had no effect on participation rates of high status students. There was also a significant correlation between the use of the two status treatments ($r = .78$, $p < .001$). The researchers also found that status treatments were a statistically significant negative predictor of the correlations observed between status and interaction in the classroom. That is, the more a teacher implemented status treatments in the classroom, the less status and interaction in the classroom were correlated.

The above results supported the hypotheses of the researchers. Although many of the results were not significant, the researchers accurately predicted the relationship and the direction of the relationship. One major limitation of this study, as Cohen and Lotan (1995) mentioned, was the low rate of teachers implementing status treatments despite the additional training. Some teachers had a number of 10-minute observations in which they did not implement any status treatments. The researchers explained that status treatments may not happen at high frequency, but need to be thoughtfully worded and given at appropriate times to be most effective. Thus just measuring the frequency of status treatments may not be sufficient in creating an accurate picture of the teachers' use of status treatments. The researchers did however control for overall classroom interaction levels and extreme heterogeneity to better focus on their hypotheses. This study was also conducted in an elementary school setting and may not be completely generalizable to middle or high school populations due to the more complex nature of classroom participation and status with age. The researchers included in their study the effects of status treatments on high status students. Critics of heterogeneous classrooms or mixed-ability grouping have expressed concern about high status students in situations in which status treatments or the like are occurring. Cohen and Lotan (1995) were able to show that in elementary schools, high status student participation was not affected by status treatments and that teachers felt free to assign competence to both high and low status students.

This study highlighted the importance of utilizing status treatments within a classroom. Comments that may seem simple to a teacher can have profound effects on the learning environment of the students. Multiple ability treatments can easily be written into lesson plans and given before complex task group projects. Assigning competence to students, as shown by Cohen and Lotan (1995), is more challenging, especially if working with high school populations. As a teacher it is important to make sure that assigning competence is happening on a semi-frequent basis during mixed ability group work and also that comments are genuine and relevance to the task.

Tammivaara (1982) completed a quantitative study in which she looked at how multidimensional tasks and multiple ability treatments affect participation of high and low status group members. She hypothesized that high and low status group members will have different participation levels in the condition where reading ability and task criterion are not dissociated. Groups members will have more similar participation in the group setting in the condition where the task is described as needing many skills and reading ability is dissociated from the task.

The participants in this study were 144 males in either fifth or sixth grade. The researchers controlled for ethnicity (white), age (10-12 year old), height (within six inches), and socio-economic status (middle-class) to isolate reading ability as determinant of status. The participants completed a reading ability test, the reading scores from the Comprehensive Test of Basic Skills, and also self-assessed their reading ability in relation to their classmates. This information was combined into one “reading ability” score. Participants were then stratified by reading ability and randomly chosen into one of the two conditions. They were then placed in four person groups consisting of two high and two low reading ability groups. The participants did not know other group members and had no idea of each other’s status before the experiment. All groups started the experiment by researchers publicly assigning status. The participants in a room with others completed a “reading assessment” and were verbally told how they scored in front of the group members (scores given reflected the previously determined high or low ability from the two measures outlined above). If participants were in the control group they were given two tasks to work on together with no reference to reading ability or the skills needed for the tasks. The treatment group was told before the first task that reading ability is not need for the task and were told explicit skills that would be necessary (none of which were related to reading ability). The participants then after the intervention completed the two tasks. The researchers videotaped both the control and treatment groups during the second task. After the tasks were completed, each participant was interview separately and asked questions about the group members' leadership, helpfulness of advice, creativity, and good judgment.

The researchers coded each tape twice, once recording interaction rates and once recording influence rates. The coders made note of suggestions, evaluations, questions to the group, and attempts to get the group's attention as initiation. From the analysis of the video codings and the end of study interviews, Tammivarra (1982) found a number of interesting results. She found that initiation and influence within the group were correlated. She also found that in the treatment group there was a significantly smaller difference in the number of initiated acts within the groups working on the tasks when compared to the control group. She also found in the control group high status subjects held the top rank (16 out of 18 groups), while in the treatment condition top rank was equally held by both low and high status subjects. Finally, from the video the coders determined rank order of the person with the most leadership quailities in the group. The coders for the control condition determined high status subjects were the leaders 88% of the time, while for the treatment group both low and high status subjects were evenly determined the leaders. When Tammaviaara (1982) compared the results of the videotape to the interviews, she found very similar reports across the two sources.

The researchers controlled this experiment to isolate the variables of concern. The researchers controlled for age, height, socioeconomic status, gender, and ethnicity. Due to the controlled nature of this study, there is not much generalizability to larger populations of students with different status factors interacting along with reading ability. Similarly, reading ability is connected with status in elementary grades but has not been shown to hold through middle and high school. They were very concerned with testing group participation and perceptions based on status treatments. The findings of Tammaviaara (1982) are consistent with the body of literature about status and expectation theory. Also, the researchers coding the videotaped group activity did not know the purpose of the study and the interobserver reliability was tested.

From this study, the impacts of multidimensional tasks and multiple ability treatments are profound and very necessary to equalize status. Even when status was artificially imposed on the students, they still responded to multiple ability treatments by working more equitably together. In complex instruction, this shows the importance of emphasizing the different skills necessary to complete the task. Also this shows how much status can affect students taking on leadership roles. With a seemingly simple multiple ability treatment, students were much more likely to take on leadership actions as observed by the coders and other group members. When taking on leadership roles, students show that they feel more confident in their skills and competent. According to Stipek et al. (1998), ability perceptions are central to learning and affect emotional experiences and enjoyment (study outlined in detail previously). Students who thought of themselves as more competent were more focused on learning and mastery, and reported more positive emotions, and greater enjoyment.

Ross and Cousins (1994) completed a quantitative study that looked at the extent to which intentions to seek and give help correlated with observed behavior in classroom situations. After examining the body of literature, they found that most researchers focused on children's attitudes toward helpfulness instead of their actual behavior. They were interested in investigating whether intentions and behavior toward giving and seeking help were correlated.

The participants were 96 ninth and tenth grade students from Ontario, Canada (58% female, 42% male). In addition, 96 seventh and ninth grade students from the same school district participated in group interviews to service as a pilot test group for the self-report survey about giving and seeking help. The teachers in this study were given a two hour in-service on how to teach the topic of correlational reasoning using either cooperative learning or a computer pair approach. The teachers were randomly chosen for one of the two teaching techniques. Teachers then administered self-report surveys to their students about giving and seeking peer help as well as a pre-test on correlational reasoning. Teachers then implemented the teaching technique they were taught at the inservice, while student groups were audiotaped. The audiotapes were later transcribed and coded by the researchers. The cooperative learning technique was modeled on Student Teams Achievement Divisions (STAD). This form of cooperative learning involves putting students in heterogeneous groups of four and giving the students group and individual tasks to complete. The groups were intentionally composed of one high ability student, one low ability students, and two average students per group and adjusted to balance the number of male and female students per group. Rewards for interdependence were given, consisting of recognition awards and bonus points for group performance. Interdependence was deemed based on the sum individual test scores and average gains of the group. The computer pair teaching model consisted of pairing two students with a computer to complete similar individual and group tasks as in the STAD model. The pair was intentionally heterogeneous with regard to both sex and academic ability. This model evoke what the researchers deemed "resource interdependence" since the pair was sharing a computer.

After analyzing the audiotapes using Text Analysis Package software, the researchers found a number of interesting results. They found that both seeking and giving help were significantly higher in the cooperative teaching model than the computer pair model (F (6, 102) = 2.85, $p < .01$; F (6, 102) = 3.66, $p < .01$ respectively). In the cooperative learning model, students who were more aware of the costs of seeking help (deemed from the survey) were less likely to ask for help ($r = -.29$, $p < .05$). In the computer pair model, students who expressed a greater willingness to seek help were observed to do so less frequently than other members ($r = -.27$, $p < .05$). Similarly, students who were most likely to need help (deemed from the pre-test) did not seek help. They also in the cooperative learning condition found that students of lower ability were more likely to make requests for explanations that were ignored ($r = .28$, $N = 45$, $p < .05$). In addition the researchers found that in the computer pair condition students who reported a greater willingness to help others were less likely to acknowledge the contributions of other group members ($r = -.27$, $p < .05$). Overall, the researchers found that intentions of giving and seeking help did not accurately predict observed behavior during group work.

In efforts to make this study more valid, the researchers took great care in designing the study. The researcher tested the internal reliabilities of the scales on their help giving and seeking survey items. Also, they led focus interviews to indicate survey items that were confusing or misinterpreted. The pilot testing showed that the focus group understood the items as the researchers intended. The researchers wanted to be certain that the survey accurately measured intentions of giving and seeking help in group work situations. The researchers also chose to only record group interactions one time for 20 minutes. Although 20 minutes is a fairly long period to record, group interactions change as the group works together. The researchers did not control for when in the group process this recording was taken. The groups worked together for the whole unit and were recorded at some point in that time. In the discussion, the researchers noted the group dynamics change over the course of several weeks. Finally, despite the careful pilot testing for survey understanding, due to the nature of the content of the survey social desirability may have come into play. Students may inflate their self-reporting of helping others, and depending on how they view themselves as a student either inflate or diminish their reporting of seek help. Social desirability may be one factor that contributed to the disconnection between reported behavior and observed behavior.

The group work observed in this study did not mindfully attend to issues of status between students within the STAD model. From the results, status seemed very much at work within these group dynamics. The researchers found that students of lower ability, and quite possibly lower status students, were more frequently ignored in the group. Cohen (1994) talked at length about low status students being left out of conversations, ignored, and physically removed from the task (p. 35-36). This is one instance that shows status is at work within these group dynamics. Similarly, students who reported willingness to give help, and possibly higher status students, were less likely to listen to their group members. This finding paired with the previous shows on both ends that status is affecting the interactions of the students in this group. On top of these results, the researchers found that the students of lower ability and probably lower status were less likely to request help. Thus the students who needed help the most felt uncomfortable asking for it. This shows that both in small group work and in pairs students encounter status issues that affect their interactions and ultimately learning. When issues of status are not addressed in the cooperative learning model, issues of status infect the group interactions and do not provide the equitable educational experience intended.

Cohen and Anthony (1982) conducted a quantitative study looking at how social status affects the interaction and participation of students during group work. The main hypothesis of this study was that the amount of working and talking together would be related to the status of the student, while holding constant pre-test scores on a measure reflecting the curriculum content. They also predicted that children of the same academic ability who were of different status characteristics would have different levels of talking and working together, and ultimately different learning opportunities in group situations.

The participants were 307 children from nine bilingual second through fourth grade classroom from five school districts in the San Jose area. The children were largely from Latino background, with a small number of students of White, African American, and Asian backgrounds. Most of the students' parents were of working or lower white-collar working class and students had varying levels of language proficiency in English and Spanish. The teachers in this study were all members of the Bilingual Consortium and were given an aid to work with them for the duration of the study. The teachers used the math and science curriculum Finding Out/ Descrubrimiento (FO/D) developed by Edward DeAvila.

The participants were given three tests before and after the curriculum treatment. One test was a content-referenced test used to measure the learning outcomes of the FO/D curriculum. Also, the California Test of Basic Skills was given in the fall and spring (as with all schools in California State). Finally, a test developed by DeAvila and Duncan designed to measure English proficiency was given to the students as well. Before the teachers implemented the curriculum the researchers also used a sociometric instrument to determine each student's status in the class. The instrument asked students to choose classmates that were for example "best at math and science," "had most trouble with reading," etc. There were both English and Spanish versions of this instrument to be sure that each child was able to accurately give their input. The Cohen and Anthony (1982) research team then visited classrooms one day a week during the implementation of the curriculum for 15 weeks. They used a small group scoring system previously developed in past experiments and focused their observations to a group of target children. They chose target children based on two categories to include children with varying levels of English and Spanish proficiency and also students that were selected by the teacher to have difficulty in math and science areas. For each target child, the researchers made note of the frequency of task-related talk, working alone, working together with other students, and off-task behavior. The reliability of these observations were tested and found to have over 90% agreement. To better analyze their original question, Cohen and Anthony (1982) combined the information gather on talking and working together to create one index which they called Rate of Talking and Working Together.

Given the data collected, the researchers had a number of interesting results. The researchers found that the different measures of status on the sociometric instrument were all significantly correlated. Thus high reading ability, high math and science ability, best friends, and good at sports all showed significant positive intercorrelations. On the other hand, low reading ability showed significant negative correlates with all of the high status measures (including high reading ability, high math and science ability, best friends, and good at sports). The researchers found that all of the status variables were significantly correlated with the rate of talking and working together (high reading ability $r = .22$, $p < .05$; low reading ability $r = -.21$, $p < .05$; high math and science ability $r = .24$, $p < .01$; best friends $r = .21$, $p < .05$; and good at sports $r = .20$, $p < .05$). The researchers then looked at just the attractiveness and math and science status characteristics to form a costatus score for the target students. They then used a regression analysis to test how the pre-test scores and the costatus scores relate to rates of talking and working together. They found that the costatus had a statistically significant beta weight (F (6, 97) = 5.28, $p < .01$) but the pre-test did not. Thus students that were ranked of higher status talked and worked together at a higher rate, but students that had a higher pre-test score did not talk and work together at any consistent rate. The researcher also found that scores on the curriculum content post-test was correlated with talking and working together ($r = .22$, $p < .05$). Thus high rates of talking and working together were correlated with high curriculum content post-test scores.

Due to the set up of the groups, the students were not interdependent but individually responsible to finish the corresponding worksheet for the learning center. Although interdependency was not a factor in the groups, status still played a role in who talked and worked together. When applying these results to other populations it is important to note that this was a bilingual classroom, with both teachers and aids fluent in both English and Spanish. The researchers in designing the study took great care to make sure task cards, directions, and surveys were provided in both English and Spanish. Proficiency in English may be more or less of a status determinant in other classroom situations. Similarly, reading ability may not be as strongly related to status in middle or high school. Other factors may play bigger roles such as popularity, membership of clubs, grades, socioeconomic status, gender, ethnicity, etc. The researchers in this study did not attempt to measure status on sex, socioeconomic, or racial characteristics. The researchers stated that after looking at choices students made for “good at math and science,” gender did not seem to play into the decision and was not a factor of status in this class. Another point of concern is that researchers found some teachers inadvertently reinforced status by assigning “good readers” to each learning center to help others. Due to the limited training on status, teachers may have been reinforcing status with comments and body language. With the similarities between children’s ranking of others reading ability and the teacher’s ranking, it is important for teachers to be mindful of status while implementing group work and engaging the classroom in general.

From the results there are a number of ideas teachers can take away. High status or low status was uniformly given across multiple domains including academics, attractiveness, and athletic ability in this study. This type of uniform status may become more complicated as students progress through middle school and high school and definitions of popularity change. The regression analysis pointed to the fact that ability did not necessarily account for high rates of talking and working together. However, status was a predictor of rates of talking and working together. In group situations where status is not addressed, students are not learning from each others prior knowledge and experiences. They are simply listening to students who have high status while they explore and learn from the given task.

Cohen and Roper (1972) completed a quantitative study in which they were studying the affect of status treatments on group dynamics of White and African Americans males. The researchers were replicating previous studies done to show white dominance in group work situations. Using previous studies as the comparison group, the researchers tested three treatments of status for African American males. The researchers hypothesized that racial imbalance in group interactions and influence is less likely to occur if both African American and White group participants had expectations treated when compared to only black expectations treated. Similarly, they also hypothesized that racial imbalance in group interactions and influence is less likely to occur if relevance of the training task is specifically stated. That is, if African American males are praised explicitly for “self-confident” behavior and “good explaining” during their teaching time and told that those skills are necessary for being a good team member, then the White and African American males would work together more equitably.

The participants were 57 four-man groups with half White and the other half African Americans. There were 19 groups in each of the three conditions set up in the study with participants randomly assigned to groups. In Treatment A, African American participants were taught how to build a transistor radio and were reinforced by being shown a video of them learning to build the radio. The African American participants then taught research staff members how to build the radio and then again were reinforced by being shown a video of them teaching others how to build the radio. Then the African American participants completed a group task that was done in a group of four other men, two African American and two White. In Treatment B, the researchers added having White students present during the reinforcement of building the radio. Also, White participants that would be in the group with the African American participants became the subjects to teach how to build the radio. The White participants again were present for the reinforcement of teaching others to build the radio. The group of four then went on to complete the group task. For Treatment C, there was only one change to the procedure of Treatment B. The researchers between having the African American participants teach how to build the radio, were taken aside and told that they did very well teaching and explicitly told them that those skills (specifically listed self-confident verbal behavior, explaining ideas well, and speaking up) would be useful in the next group task. Each participant were also interviewed using a standard questionnaire, including a question specifically asking them to rank order the group members (including themselves) as to who best contributed the best ideas, gave the group guidance, and was overall the leader. While the groups completed the group task, the researchers videotaped their interactions and later coded the verbal actions. The researchers noted performance outputs, action opportunities, positive evaluations, and negative evaluations. The researchers also tallied group members making the response (initiator) or receiving a response (recipient) as shown important in previous studies.

Cohen and Roper (1972) found support for one of their hypothesis and mixed results for the second hypothesis. From the videotaped group tasks, they assigned each person a participation number and top ranked each group. In Treatment A, they found that White participants were more likely to hold the top rank position then African Americans. In Treatment B, the rank order patterns were not determined by race, with some groups having White group members top ranks for participation while other groups had African American group members top ranked. They also found no correlation between race and lowest participation rank as well. In Treatment C, the researchers did not see the change they were expecting. White and African American group members still had equal chances of being top ranked for group participation, however 13 of the 19 groups had African American group members holding the lowest rank for participation. They also looked specifically at initiation rates between group members. In previous studies they have found this to be an important indicator of group inequity. The researchers found that the difference in initiation decreased in both Treatment B (2% difference in activity) and Treatment C (1% difference in activity). The change in the most active African American group member in Treatments A and C are also statistically different ($p < .025$). Finally the researchers found that in Treatment A group members chose Whites more often as the leader (African American, 28; Whites, 43), Treatment B group members chose nearly equal numbers of African Americans and Whites as leaders (33 and 32 respectively), and in Treatment C African Americans were more often chose as the leader (African American, 45; White, 26).

The researchers had two independent observations of each videotaped group activity. They were also very purposeful in having at least one White and one African

American code the videotapes to help minimize biases in the data collected from the group tasks. Also observers were unaware of the hypotheses and the differences between the treatment groups. The researchers randomly assigned White and African American men to groups, controlling for racial make up of each group. The researchers did not explain how the participants were selected from the overall population. Also, the age, socioeconomic status, and education levels of the participants were not included as well as the geographic location of this study was also not disclosed.

Cohen and Roper (1972) show through this study that teachers implementing complex instruction can overcome strongly held societal biases and stereotypes at least within group situations. In complex instruction teachers are to assign competence publicly instead of privately as in the study. However, even private assigning competence via specific and relevant skills still had a positive affect on leadership skills and minimizing differences between group members of different status. This study showed that even with in the increase of participation from the most active African American in the group, the most active White in the group did not have a diminished level of activity. This shows that status treatments were not negatively affecting the participation of the "high status" group members.

Within complex instruction attending to issues of status and competency expectations is very important. If students are not working together in an equitable way, learning opportunities for all students are being lost and diminished despite the tasks given. Cohen and Lotan (1995) showed the importance of frequently utilizing status treatments like multiple ability treatments and assigning competence to low status students. Tammivaara (1982) found that status treatments had profound affects even on artificially imposed status situations. Ross and Cousins (1994) unintentionally showed that status is very much at work in cooperative learning situations where status is not addressed. Cohen and Anthony (1982) showed that rates talking and working together are determined by status and not academic ability as well as higher rates of talking and working together is correlated with more successful learning. Cohen and Roper (1972) were interested in creating equal status group dynamics between White and African American males. They applied status treatments and assigning competence to find a more equitable working situation.

Student Beliefs, Self-Efficacy, and Identity

Despite the importance of the classroom curriculum, grouping strategies, and attending to issues of status, teachers within the complex instruction model also need to be aware of student beliefs, self-efficacy, and identity development. Boaler, Wiliam, and Zevenbergen (2000) interviewed high school math students and found they viewed math as rigid, inflexible, and not relevant to their developing identities. Woods, Kurtz-Costes, and Rowley (2004) found that students from upper elementary and middle school all believe that rich students do well in academics, while mixed views were held about music and sports. Rodriguez, Jones, Pang, and Park (2004) analyzed an outreach program that promotes academic achievement and identity development among culturally diverse high school students. Blackwell, Trzeniewski, and Dweck (2007) analyzed how students' beliefs about intelligence affect their academic achievement in math course through the transition from elementary to middle school. Schunk (1982) showed that effort attributional feedback for past achievements lead to greater math skill development and higher perceptions of self-efficacy. Schommer-Aikins, Duell, and Hutter (2005) explored how epistemological and mathematical problem-solving beliefs influence academic performance. Crosnoe, Riegle-Crumb, and Muller (2007) looked at how perceived intelligence changed due to gender, learning disability diagnosis, and failing classes.

Boaler, Wiliam, and Zevenbergen (2000) completed a qualitative study looking at the connections between identity development and high school mathematics. Looking at math classrooms from the perspective of Lave and Wenger and a community of practice, teachers show students what it means to be a member of the community of mathematics and students decide, either consciously or unconsciously, whether or not they want to become part of that community. The researchers were curious how math classes were viewed through the eyes of the students, connections students made between their identity and math classes, and implications for student enrollment in higher level math classes.

The participants in this study came from two major sources. The first main group of students interviewed was 48 Advanced Placement calculus students from six northern California public schools. The second group were 72 students from six secondary schools within the United Kingdom. The researchers did not report demographic information about the students. The students were each interviewed individually with one other researcher and at the beginning of the interview to describe math lessons and classes. The interviewers then asked questions and engaged in conversations with the students about the different aspect of the lesson described. Through the conversation students discussed their views and beliefs about math classrooms, teacher actions, likelihood of taking more math classes, and their overall impressions of math.

The researchers found a number of interesting results after analyzing the interview responses. Students from both countries gave very similar accounts of math lessons, which included teachers lecturing, teacher showing example problems to the whole class, and students doing practice problems for homework. Students from two of the US schools and two of the UK schools reported that they were encouraged to work on homework problems collaboratively. Students from the other four US and UK schools reported that their math classes were individual environments with little opportunity to discuss or work with others. In the individualistic classrooms, the teachers instructed students to follow the procedures outlined in the textbooks, while in the other classroom the teachers instructed student to work in small groups and discuss the different questions. Students in more individualistic classrooms in both the US and UK described math as "procedural, absolute, memorization, concrete, always a right answer, and rule-bound." This view of math was not dependent on gender, math achievement, or confidence levels. The students in classes that incorporated group work saw math as a "field of inquiry, discussion, and exploration." The students in more collaborative classes also reported more conceptual understanding and striving for understanding "why" not just "how" along with a greater enjoyment of math.

Most students in the US reported disliking math, even though they were successful math learners (AP students). Students reported that math was not reflective of what they were interested in or wanted to spend more time doing. Ultimately, math was not seen as congruent with their identity. Most of the students described themselves as a "creative, social, and/or emotional" person and more into history, English, and visual arts. Being a "math person" was not associated with how they saw themselves. Students in the UK had similar reports as to US students. Students in the UK seemed to have more intense opinions. The students that reported disliking math did so with intensity. The students who did like math did not identify as a "math person" or with mathematics. Students that liked math tended to like it because they were good at it or they needed it for further education or employment.

The researchers had some strength and weaknesses in their study. The researchers did not explicitly say if member check were performed to make sure that their assumptions about identity and being a "math person" were completely correct. Also, the researchers did not include the demographic information about the participants. Similarly, the participants were all students in advanced math classes. The researchers did compare students' view on math from different countries.

This study is very telling of the current state of mathematics in many schools. Students do not see math as relevant to their daily lives and in turn do not incorporate being a math learner into their identity. Students that were able to bring more of themselves into the classroom, students that worked in small groups, enjoyed math more. Math educators need to take a hard look at what high school math classes are communicating to students about the nature of mathematics. To fully engage in mathematics, students need to see mathematics as it truly is, as creative, social, and emotional subject with room for uncertainty, discussions, and collaboration. To be a successful math learner or mathematician, students need to develop more skills than repeating procedures, following rules, and getting the right answer. In implementing complex instruction, the teacher needs to include tasks that are relevant to the students lives as well as multidimensional, uncertain, and open ended. Students need to be encouraged bringing their identities, skills, and funds of knowledge into the classroom so they can start to see themselves as a "math person."

Woods, Kurtz-Costes, and Rowley (2004) completed a qualitative study examining students' beliefs about "rich" and "poor" and their abilities in academics, sports, and music. The researchers were interesting in how stereotypes about poor people held in society have affected the beliefs of school children. The researchers hypothesized that as children get older, they hold more complex views being "rich" or "poor" and being good in academic, sports, or music. The researchers were also interested in testing whether in group biases held (high-income students gave "rich" higher ratings while low-income gave "poor" higher ratings) or if low-income students would view "poor" more negatively because of their experiences and awareness of hardships that face people in poverty.

The participants were 438 African American and European American students from North Carolina (182 boys and 256 girls). The students were from 10 different schools in two school districts, one urban and the other rural. The students were approximately half African American and European American and nearly evenly split between the three grades tested, fourth, sixth, and eighth. In the two school districts used in this study 32 and 36.5% of students received free or reduced lunch. Researchers also sent home a parent/ guardian questionnaire that included asking about yearly income. Only 172 of the parents returned the questionnaires, which limited the size of the sample for analyses involving parental income. The self-report questionnaires were administered at the children's school and parents were mailed their self-report questionnaires. The student questionnaires assessed beliefs about competence in academics, sports, and music for the social groups rich and poor.

The researchers completed multivariate analyses of variance and univariate analyses of variance to interpret the data collected using social class stereotypes for sports, academics, and music as dependent variables. The researchers found that younger students favored rich in all three domains, academics, sports, and music. While, older students held views that favored rich in academics, poor in sports, and egalitarian views in music. White students favored the poor in sports stereotypes, where African American students viewed both rich and poor equally good at sports. Both Whites and African Americans held similar beliefs about music ability. Low income students favored rich students to be good at sports, while high income students favored poor students to be good at sports. It is important to note that family income and race were highly correlated in this study (most low income families were African American and high income families White) and it is not possible to determine if income, race, or the interaction are causing the "family income" results.

The researchers had a number of strengths and weakness in their study. The researchers counter-balanced the questionnaires given to students and have three different sequences of items used to help eliminate response biases. The researchers also made sure to include both urban and rural students to get a more full picture of students from a variety of schools. The researchers decided to combine questionnaire scores on math, science, reading, writing, school grades, and smartness into one "academics" score. Since academics had an overall similar view from all students across grades and background that rich students were better at academics, it would have added more complexity to the results to see if any differences arise when separated into the subjects. Students of ethnicities other than African American and European American were excluded from this study.

This study has interesting implications for classrooms implementing group work. This study shows how stereotypes about socioeconomic status can affect how students perceive the ability and competence in other students and even themselves. With most students reporting that rich students are good at academics, teachers implementing group work need to emphasize that students of all backgrounds are equally capable to completing tasks and effort over “natural” ability. This study also highlights why some students do not feel that school and certain subjects are not for them. With such strongly held stereotypes about people who are poor, students of poverty may not identify with school and academics. Teachers need to work hard to show counter examples of stereotypes and work to not perpetuate those stereotypes in group work by using multiple ability treatments and assigning competence.

From the work of Woods, Kurtz-Costes, and Rowley (2004), students’ beliefs are highly influenced by stereotypes that are held by American society as a whole. These stereotypes not only affect people of poverty but also students belonging to minority groups. Rodriguez, Jones, Pang, and Park (2004) completed a case study focusing on an outreach program that work to promote academic achievement and identity development among diverse high school students. The researchers were interested on how scores on math and science problem solving tests changed after the summer intervention program as well as hearing what students specifically are taking away from the summer experiences. The researchers also highlighted in the case study what the outreach program does to promote academic achievement, increase percentages of students admitted to college, and identity building.

The participants of the study were 193 adolescents in four cohorts enrolled in the Science Enrichment Program between the years 1998 and 2001. The students enrolled within the outreach program were tenth grade students with high levels of cultural diversity. To become enrolled in the program, students must going into tenth grade and either a first generation college bound and/or from a low income family. Of the 193 students enrolled in the program, students were from a variety of geographic locations within the United States and of Mexican American/Latino (45.75%), Native Hawaiian/Pacific Islander (17%), African American (21%), Native American (12.25%), or Other (4%) cultural and ethnic backgrounds. Forty percent of the students in the Science Enrichment Program have a first language other than English. The outreach program allowed approximately 50 students in each year and was a yearlong program starting with a six-week summer intervention. During the summer, students all traveled to a university, lived in the dorms, attended various math and science classes, and engaged in field trips. The classes students took were hands on, culturally relevant, and taught by certified teachers. The students also had a study hall time each night to complete work for class and to collaborate with other students. Experienced resident assistances also lived in the dorms with the students to help facilitate a productive and healthy living environment for the students while there for the summer. Most of the summer intervention teachers were Latino and 18 of the 22 staff were members of underrepresented populations served by the program.

The researchers, to test academic achievement of students within this summer intervention, administered the Test of Integrative Process Skills (TIPS) to all students before and after the summer program. The TIPS assessed five academic skills within math and science including identifying variables, identifying hypotheses, identifying operational definitions, designing investigations, and graphing and interpreting data. Rodriguez et al. (2004) also administered group interviews with the first three cohorts of the study. The group interviews were voluntary and the groups consisted of eight to ten students.

The researchers found some interesting results from the pre and post test as well as the interviews. There was a documented increase from pre to post test across the cohorts. Not only were students test scores higher, but through student interviews the researcher found that students reported feeling more competent in math and science. They felt better prepared to complete coursework that emphasized problem solving. Within the focus group, students remarked that they felt that everyone was treated equally and more fair then typically at their high schools. This learning environment encouraged biculturalism and strived to affirm students cultural systems and values. Also due to the high levels of diversity, students felt that they learned about other cultures and their own in addition to math and science. Many students remarked feeling empowered, increased confidence levels, and higher levels of participation in the learning experiences within the six-weeks. Students also liked that the classes were focused on improvement and personal growth. This lead to motivation to learn and development of identity as a “math person” or a “science person.”

There are some areas of strength and weakness in this study. The researchers did not mention if the students were given the same exact pre and post-tests. The increase in test results may be due to learning that occurred from taking the test. The researchers also did not test students at the end of the year program to see if the students retained the information learned over the summer. The researchers also did not track if students had improved success at their respective high schools or entered college. The researchers showed a more complete picture of what the students gained from the summer intervention by analyzing both test scores and focus group interviews.

Although this study did not explicitly show long term growth in the students, there are still some interesting conclusions that can be drawn. The teachers within the intervention summer program were dedicated to valuing biculturalism and implementing culturally relevant curriculum and pedagogy. Teachers implementing complex instruction in math classes especially need to show students that their whole self is valued in the classroom and their experiences are important to bring to their group work. Tapping into students cultural funds of knowledge helps to empower the student and show other students different perspectives. To help students see themselves as success math students, the identity they bring to the class must be valued. This study also shows that students have increased motivation to learn new material when working for personal improvement and growth. Motivation is also improved by a focus on effort and process instead of ability and end product (Blackwell, Trzeniewski, & Dweck, 2007).

Blackwell, Trzeniewski, and Dweck (2007) completed a quantitative study focusing on assessing the intelligence beliefs held by students, how those beliefs impact their academic achievement, and if intelligence beliefs can be manipulated. The research questions for this study were, “Are students’ theories related to their achievement trajectory,” “Why is intelligence theory related to grades,” and “Does teaching an incremental theory provide an added benefit over a similar academic intervention?” The researchers were very interested in how student math achievement is affected by the transition from elementary to junior high school.

The participants in this study were 373 students (198 female and 175 male) in a public school in New York City. The sample of data was collected on four entering seventh-grade classes ranging in size from 67 students to 114 students. Due to the small sample size, the four seventh grade cohorts were collapsed together for data analysis. The participants represented a variety of ethnicities (55% African America, 27% South Asian, 15% Latino, and 3% East Asian and European American), achievement levels, and levels of socioeconomic status (53% of students eligible for free lunch). As an overall group they were deemed high-achieving, with the average sixth-grade math test scores in the 75th percentile nationally. In each cohort, all the students had the same math teacher, same math curriculum, and the math classes were heterogeneous. The researchers collected standardized test scores from the math achievement task student completed at the end of sixth grade to serve as a prior math achievement baseline. The researchers then collected math class grades throughout seventh and eighth grade to assess math achievement throughout junior high school. The researchers also gave students a questionnaire about motivational variables at the beginning of junior high school. The questionnaire had students mark agreement with statements on a six-point Likert scale including theories of intelligence, learning goals, effort beliefs, and helpless responses to failure.

Blackwell, Trzeniewski, and Dweck (2007) found a number of results from the data collected. They found that incremental theory of intelligence was positively correlated with effort beliefs, learning goals, low helpless attributions, and positive strategies. The researchers found that at the beginning of seventh grade there was not a significant correlation between prior math test scores (in sixth grade) and theory of intelligence or other motivational variables. As students moved through seventh and eighth grade, intelligence theory beliefs were significantly correlated and a predictor of math achievement. That is, students that held incremental beliefs about intelligence at the beginning of junior high school earned higher math grades at the end of eighth grade (controlled for prior math achievement). They also found that incremental intelligence theory was related to motivational constructs and positive motivational beliefs.

There were a number of strengths and weaknesses within the design of the study. The researchers had a relatively small sample size from each cohort and were only tested students at one school with one teacher. These results would need to be replicated in multiple schools to show the results are generalizable to all students in junior high schools. This study also was conducted with students that attended elementary school until sixth grade. The transition into new schools happens earlier in some districts, having sixth grade in the middle school. This documented shift may occur earlier or later in adolescents due to changing schools. Also, the sample of students included in the study was deemed high achieving. The researchers did control in their analyses for prior math achievement and the variability of different teachers grading scales (all students had the same math teachers).

This study highlighted the importance of holding an incremental theory of intelligence for learning math successfully. Students going through a, sometimes, challenging transition were significantly more successful when they believed they were in control of their ability to gain "intelligence." Students that held an incremental theory of intelligence were also found to set learning goals, believe high effort pays off, and did not feel helpless after experiencing a failure. Although this study did not show that teachers can change students' beliefs about intelligence theory, emphasizing the components of incremental theory could at least encourage the students that do hold those beliefs. Another study performed by Blackwell, Trzesniewski, and Dweck (2007) however showed that after learning about the malleability of intelligence, students deemed low achieving endorsed more incremental views of intelligence.

Schunk (1982) took the idea of giving student feedback based on effort even further. He completed a quantitative study analyzing which type of effort attributional feedback is best, feedback on past accomplishments or future accomplishments. He also included a group with no feedback to show that effort attributional feedback is effective in improving skill development and self-efficacy. He hypothesized that effort attributional feedback for past achievement would promote children's achievement and perceptions of self-efficacy. He also predicted that effort attributional feedback for future achievement would not significantly promote achievement or self-efficacy. An example of effort attributional feedback for past achievement from the study is "You have been working hard," while future achievement feedback is "You need to work hard."

The participants in this study were 40 children (26 males, 14 females) from ages seven to ten. The children were of predominantly middle class background. The researcher did not report demographic information about the participants. The children included in this study were identified as lacking skills in subtraction. The participants completed an initial pre assessment of arithmetic skills and self-efficacy. The subtraction skill test included 25 problems and a researcher timed the completing of the "easy" problems and the "harder" problems. Low-difficulty problems involved less steps and less borrowing (in subtraction terms). The self-efficacy test was administered directly after the subtraction skill test and was a measure used in previous self-efficacy research. The students were shown a subtraction problem for two seconds and then asked if they would be able to solve that type of problem. The children were then randomly selected into one of the four treatment groups. The four groups were past attributional, future attributional, monitoring, and training control. All of the students were given three consecutive school days to work through a subtraction packet that gave examples and problems to improve student skills in arithmetic. The packets had lower difficulty problems at the beginning and became progressively harder. For the past attributional treatment, the researchers administering the packets would monitor the progress of students every eight minutes by asking what page they are on. The research would then follow that child's response with, "You've been working hard." The researcher then left. For the future attributional treatment, the researchers monitored the children every eight minutes by asking what page they were on and then following the response with, "You need to work hard." Within the monitoring condition, the research would ask every eight minutes about the child's progress and then leave without saying anything else. Finally in the training control group, children were allowed to work on the packets without any interruptions. For the post assessment, the researcher administered a self-efficacy measure, a subtraction skill test, and then another self-efficacy measure. The post assessments were similar to the pre assessments.

Schunk (1982) found that his hypothesis was supported by the data collected. He found that all groups except the monitoring groups showed significant improvement from pre to post subtraction skill test. The past attribution effort feedback group had significantly greater subtraction skill improvement than the other condition groups. The past attributional group also showed a significant increase in self-efficacy. He also found that children in the past attributional group completed 81% of the subtraction packet on average. Children in the monitoring, future attribution, and training control had significantly lower completing percentages when compared to the past attributional group (58%, 50%, and 46% complete respectively). The persistence measure (timing of the skill test) was positively correlated with the self-efficacy measure and accurate problem solving. All in all, the past effort atttributional feedback resulted in higher academic achievement and higher self-efficacy. These findings are consistent with findings from the literature.

There are a number of strength and weaknesses in the study. Children were given the self-efficacy measure alone, without other children present. Children were given a practice question on the self-efficacy measure to make sure children understood the scale and format of the measure. The children were also told to be honest and to mark how they really felt. The researchers also pilot tested the subtraction packets to make sure that students could complete the packet with diligent work over the course of three sessions of 40 minutes. The researcher staggered the times students were let into the room to work on the subtraction packets. This allowed for students to work at their own paces and not compare their progress to that of other students. The research used a pre and post test to measure achievement and also included two different control groups within the study. Also, the researcher administering the post assessments was blind to the experimental group the child was assigned. A limitation of this study is the generalizability of this study. The researcher did not include demographic information, school size, or students older than 10 years old. Schunk (1982) stated that children as they get older, theorized until 12, their views of ability and effort change and become more complicated. Younger children tend to explain success with effort, while older children attribute success or achievement more to their personal ability instead of effort. Thus the findings of this study many not be completely generalizable for students in middle and high school populations.

Similar to the suggestions from Blackwell, Trzesniewski, and Dweck (2007), Schunk (1982) gave a very clear and controlled study that showed the direct affects of the very simple past effort attributional feedback, "You have been working hard." When giving feedback to students the wording can be very critical. Teachers need to be mindful of the way they word feedback and purposely emphasize past achievement when giving direct feedback to students. Self-efficacy is important in a subject like math that can be challenging and/or frustrating. Students use their past experiences to judge their own capabilities. Teachers can draw attention to times when students worked hard and expended effort to gain success or academic achievement. This allows students to vividly connect past achievement with effort, which in turn has been shown to increase self-efficacy and ultimately future achievement (Schunk). Cheng, Lam, and Chan (2008) found that when group processes were of high quality, both high and low achievers had higher collective efficacy than self-efficacy. This efficacy adds to motivation and ultimately performance as a group (study previously outlined in detail). The researchers also point to the importance of giving effort-based feedback for tasks that are intermediate to difficult. As students get older their ideas about effort and ability become complicated. When students view a task as challenging, they will expect to have to work hard. When a task is “easy” students feel that they should have the ability and not have to expend a lot of effort. When implementing complex instruction, giving students complex tasks to work on is crucial to the group functioning, but also continuing the positive affects of giving effort based feedback. As teachers give feedback, Schunk acknowledged, they should consider the difficulty level of the task and also the ability of the student before giving feedback.

Schommer-Aikins, Duell, and Hutter (2005) completed a study examining how epistemological beliefs interact with problem solving performance and overall academic achievement. Epistemological beliefs were defined as “beliefs about the nature of knowledge and learning” (Schommer-Aikins, Duell, & Hutter, p. 289). The researchers hypothesized that general epistemological beliefs would affect students’ domain-specific mathematical beliefs, including problem solving beliefs, and in turn influence academic achievement and mathematical ability. They hypothesized that specifically that the less students believed in quick and fixed learning, the more likely students would believe that problem solving requires effort, confidence, and understanding math and in turn believe mathematics is useful.

The participants in this study were 1,269 middle school students from two different Midwest middle schools. Students were approximately equal in gender (587 boys, 657 girls) and grade (619 seventh grade, 644 eighth grade). Most of the students in the sample were white (86% European American, 5% African American, 5% Latino, 3% Asian American, 3% Native American) and middle class (23% of students receiving reduced school lunch prices). The researchers had the teachers administer questionnaires assessing epistemological beliefs and problem solving beliefs along with the standardized

Kansas State Assessment Instrument given in January. The epistemological beliefs were assessed by a reformed high school and college assessment. The four aspects assessed were beliefs about the structure of knowledge, stability of knowledge, speed of learning, and ability to learn. The reformed questionnaires had fewer items and were more simply worded. The students rated agreement with statements on a five-point Likert scale. The researchers also measured beliefs about mathematical problem solving using the Indiana Mathematics Beliefs Scale and the Usefulness of Mathematics Scale. The first scale measured five beliefs including ability to solving long problems, multi-step word problems, understand concepts, and effort can increase math ability. The second scale measured the usefulness of mathematics in the students' everyday lives. This questionnaire was also measuring students' agreements with statements on a five-point Likert scale. The researchers measured academic performance by collecting standardized test scores on mathematical problem solving and reading. Also grade point averages were used in the analysis of data.

With all of the data collected, the researchers had a number of results. The researchers used a exploratory factor analysis to analyze epistemological and problem solving beliefs. The researchers found that quick/fixed learning and studying aimlessly were positively correlated, while they were negative correlated with effortful math, useful math, math confidence and understanding math concepts. They found that the less students believed in quick/fixed learning, the more likely they were to believe that mathematical problem solving is effortful and useful. Also they found the belief in studying aimlessly was negatively correlated with the belief that mathematical problem solving is effortful and useful. The researchers found that the more students believe in fixed/quick learning, the less they believed math was useful and the worse they were at problem solving. Similarly, quick/fixed learning beliefs negatively impacted beliefs about the usefulness of math and the three measures of academic performance. They also found that students believed math was useful had better grade point averages and also had higher performance on mathematical problem solving measure. Schommer-Aikins, Duell, and Hutter (2005) also found that reading comprehension scores predicted math problem solving performance and grade point averages.

There were a number of strengths and weaknesses within the design of this study. Researchers reverse scored and randomly ordered items on the questionnaires to control for testing biases. From previous research, the problem solving beliefs measure had validity shown by predicting students' math scores and success in math courses. The researchers had the teachers conduct the assessments over a two-week period to avoid student fatigue. The researchers also used the two strongest epistemological beliefs found from the questionnaires, quick/fixed learning and studying aimlessly, in the analyses to strength the reliability of the results. The researchers also acknowledged that since they reformed the questionnaires from those used with older students, there needs to be more testing on the validity of the measure used in testing epistemological and problem solving beliefs in middle school students. Finally, the sample of middle school students in this study did not include high levels ethnic, lingual, or cultural diversity. Epistemological beliefs of a more diverse group of students may yield different results.

From the results, there are a number of conclusions that are important for teachers to know and consider when conducting a math class. The researchers found that younger students were more likely to hold epistemological beliefs about studying aimlessly, which suggests that younger student do not believe learning is strategic and is more related to luck or chance events. Teachers in middle school and even high school students can counter this belief by explaining to students how to study and emphasizing that learning is about effort and not luck. They also concluded that students' general beliefs about the nature of learning impact their thinking in math specifically and contribute to their overall academic success. The researchers also suggested that middle school teachers with students that believe in quick/fixed learning should tell students that math tasks are challenging and time-consuming for everyone. Also, students should be given encouragement and help to alleviate unnecessary frustration that would lead to students giving up. Also, students that do not believe that math is relevant to their lives will expend less effort and have lower motivation to spend the time necessary to learn and understand math concepts. Teachers need to help students see how math relates to their daily lives through relevant math tasks on topics that interest students. Within complex instruction, teachers can design group worthy tasks that are applicable and interesting to students to that they will want to spend time and effort to complete the task. Teachers emphasizing effort over ability is also a feature of complex instruction that may broaden students' beliefs about the nature of learning and being a math learner.

Crosnoe, Riegle-Crumb, and Muller (2007) completed a longitudinal study focusing on how self-perceptions about intelligence are affected by learning disability diagnosis, failing classes, and gender. The researchers were also interested in how nonacademic performance, family structure, and peer groups influenced perceptions of intelligence. The researchers went on to find a path model for explaining why some students take higher levels of math and science (classes that are seen as hard, scary, challenging, etc.) and others do not. The researchers predicted that girls tie more of their self-concept into the role of being a "good student" and thus failure or negative feedback will have a greater impact on their perceived intelligence.

The participants in this study were 90,118 students from 80 high schools and 32 middle schools with a nationally representative sample of seventh through twelfth grade students in 1995. Participants were selected with a stratified samples design, taking into consideration school region, urbanicity, racial composition, and size. The data were derived from the leading national survey of American adolescents (National Longitudinal Study of Adolescent Health or Add Health) and the Adolescent Health and Academic Achievement Study (AHAA). All students in the study were given pencil-and-paper questionnaires, a subgroup evenly selected across high schools participated in in-home interviews, and school administrators and parents also contributed information about schools and family interactions. Since this was a longitudinal study, there were three different "Waves" of student data collection. Each new wave of data collection involved less students due to graduations of senior classes involved in the study, with approximately 77% of the original Wave I participants in Wave III. Each student had a number of categories of data collection including course taking, self-perceptions, external academic markers, social context, and adolescent characteristics. The researchers followed how many math and science classes students enrolled in during high school. Wave I and II students completed self-perception measures about their intelligence relative to other students their age. This self-perception measure focused on intelligence change between the two sampling times. During Wave I, researchers noted which students had failed one or more classes and those that had not. Also parents reported whether or not the student had at some point been diagnosed with a learning disability. These two factors, class failure and learning disability diagnosis, contributed to the external academic markers. For social context, researchers recorded highest level of parents' education attainment in the family, academic achievement, educational aspirations, peer group orientation, and enrollment in math and science classes. Researchers also measured student truancy, delinquency in the past year, athletic status, extracurricular participation, closeness of relationship with parents, close relationships with others at school, and overall happiness with school for the adolescent characteristics measures.

The researchers found that their pathway models were supported by the data. During the analysis, Crosnoe, Riegle-Crumb, and Muller (2007) categorized students into the following categories: both academic markers (failed class and learning disability diagnosis), learning disability only (no failed classes), failure only (no learning disability diagnosis), and no academic markers (no class failure or learning disability diagnosis). They first found that the majority of students had not been diagnosed with a learning disability and that girls were much less likely than boys to be diagnosed with a learning disability or fail one or more classes. Also white students were much more likely to be diagnosed with a learning disability than minority students. Gender was not a predictor of perceived intelligence change between Wave I and Wave II data collections, although females overall took significantly more high-level math courses by their senior year. Girls that had failed one or more classes were found to have a significant decrease in their perceived intelligence between Waves I and II when compared to girls in the "no academic markers" category. High levels of Wave I perceived intelligence and an increase in perceived intelligence were predictors of girls taking high-level math classes by their senior year in high school. The pathway model constructed for girls is that failing one or more classes yields a decrease in perceived intelligence, which leads to a decrease in the number of high-level math courses taken. For boys, those who had been diagnosed with a learning disability, regardless of academic success, had a decrease in perceived intelligence from Wave I to II when compared to boys with "no academic markers." The two groups that included being diagnosed with a learning disability (failing classes and learning disability; no classes failed and learning disability diagnosis) were significantly different from the two groups that did not include being diagnosed with a learning disability (failing classes; and no classes failed and no learning disability) in perceived intelligence change. Boys who had failed a class and were diagnosed with a learning disability had the largest decrease in perceived intelligence, while boys who were diagnosed with a learning disability but had not failed any classes represented the second largest decrease (over students that had failed classes without a learning disability diagnosis). The pathway model constructed for boys is that diagnosis of a learning disability, with or without failing classes, yields a decrease in perceived intelligence, which leads to a decrease in the number of high-level math courses taken. The researchers also found the perceived intelligence of girls, and not boys, was affected by social contexts such as peer group and family.

There are a number of strengths and weakness within the design of this study. The researchers mentioned that there was no information about the types of learning disability that were being diagnosed in this sample. Similarly, it was not documented the numerical differences of the diagnosed learning disabilities of boys versus girls. The researchers suggest that learning disorders may be more highly diagnosed in boys due to behavioral issues. Also, they did not factor into the results when the learning disability was diagnosed and how that diagnosis changed the child's education (pullout program, mainstreamed, special education classroom, etc). The researchers suggest that pullout programs and special education programs may not have introduced students to the same content as students in the mainstream classroom, in turn affecting their performance in certain classes such as math. The study also had a small population of students in general diagnosed with learning disability, especially girls with learning disabilities. Within the learning disability "academic marker" there are many possibilities of confounding or lurking variables affecting perceived intelligence. The researchers also did not measure academic achievement on standardized test score or analyze grades earned in classes (potentially interacting with perceived intelligence and probability to continue taking high level courses). The researchers also only reported perceived intelligence change between Wave I and II due to lack of resources and participants graduating. To strengthen the pathway model claims, the researchers could show continued decrease, increase, or stability of perceived intelligence by measuring it multiple times (more than two) over a number of years.

This study highlights the importance of properly diagnosing a learning disability and knowing that this label may forever impact the way this student perceives him or herself. The researchers pointed out that girls are more impacted by negative social feedback, such as failing a class or receiving negative feedback from parents, peers, or teachers, than boys. As a teacher implementing complex instruction it is important to make accommodations for students with learning disabilities and continually utilizing status treatments to help boost student perceptions of their intelligence. Students that do not perceive themselves as intelligent will likely not participate equally in a group work situation. It is important that all students see themselves as able to contribute something meaningful to the group and as intelligence and competent. Also, it is important for teachers to give extra support and encouragement to students who have failed classes or been diagnosed with a learning disability to challenge themselves by taking high-level math classes. As the researchers have pointed out, "success" in our society is largely based on post-secondary education that requires students to complete high-level math in high school as a prerequisite.

How students view themselves and others connects directly to status treatments as well as group work within a complex instruction classroom. Boaler, Wiliam, and Zevenbergen (2000) interviewed high school math students and found they viewed math as rigid, inflexible, and not relevant to their developing identities. Woods, Kurtz-Costes, and Rowley (2004) found that students from upper elementary and middle school typically believe that rich students do well in academics, while mixed views were held about music and sports. Rodriguez, Jones, Pang, and Park (2004) analyzed an outreach program that promotes academic achievement and identity development among culturally diverse high school students. Blackwell, Trzeniewski, and Dweck (2007) analyzed how students' beliefs about intelligence affect their academic achievement in math course through the transition from elementary to middle school. Schunk (1982) showed that effort attributional feedback for past achievements lead to greater math skill development and higher perceptions of self-efficacy. Schommer-Aikins, Duell, and Hutter (2005) explored how epistemological and mathematical problem-solving beliefs influence academic performance. Crosnoe, Riegle-Crumb, and Muller (2007) looked at how perceived intelligence changed due to gender, learning disability diagnosis, and failing classes.

Student Academic and Social Achievement

From the studies previously analyzed, there have been a number of interesting findings related to student academic and social achievement within a complex instruction classroom. Lotan, Cohen, and Holthuis (1994) found that the social studies classes the implemented complex instruction had significantly better post assessment scores than the comparison classrooms. Cohen, Lotan, and Leecher (1989) showed that through complex instruction students show less variation in score in post test, thus achievement gaps between students of different status levels lowered. Also they found connections to students talking/working together on uncertain tasks allow from greater achievement on conceptual problems.

Within the body of literature on complex instruction, there are studies that focus on the academic and social achievement of students within this model of teaching. Bianchini, Holthuis, and Nielsen (1995) found within science classrooms students had academic gains in high school science classes utilizing complex instruction teaching techniques. Boaler and Staples (2008) compared high school students in traditional and complex instruction math classes. They tracked standardized test scores, the number of students taking upper division math courses, as well as achievement differences between minority students. Albert and Jones (1997) found that students had a number of social gains from being within a complex instruction classroom.

Bianchini, Holthiuis, and Nielsen (1995) completed a mixed methods study exploring student learning in a complex instruction science classrooms. The researchers first assessed if the classrooms met complex instruction standards for students talking and working together and engagement. They looked at the relationship between the quality of group work and student achievement. They also measured the student learning of scientific concepts and skills. Complex instruction is designed to stimulate higher-order thinking and the researchers wanted to measure the extent to which this occurred in the classrooms. The researchers also wanted to measure the equity of the group work and determine if gender or previous academic ability influenced learning. They hypothesized that gender and previous academic ability would not influence achievement gains due to the equitable nature of complex instruction classrooms. They also analyzed the curriculum used in this study.

The participants were 260 sixth and eighth grade students in 13 middle school science classrooms in the 1992-1993 school year and 80 sixth grade students in 10 middle school science classrooms in the 1993-1994 school year. Four experienced complex instruction teachers taught the science classes in the study. The schools were in the San Francisco Bay Area and were academically, linguistically, and ethnically diverse schools. To test implementation of complex instruction researchers observed each class a minimum of 10 times and assessed the group work with the Whole Class Instrument. The instrument assessed percentages of students talking and working together and student disengagement. The researchers also assessed student conceptual and factual knowledge while looking at their ability to apply and synthesize the information learned in class. The researchers gave students a pre test at the beginning of the unit and a post test at the end of each science unit with questions ranging in cognitive demand from recall to synthesis. During the two school years of the study, some classes were able to complete three units, while other classes only completed one or two. The researchers scored the tests using rubrics and multiple scorers, testing the inter-scorer reliability. Bianchini, Holthuis, and Nielsen (1995) also noted students' gender and reading achievement scores from the most recent standardized test (typically the CTBS).

The researchers had a number of interesting results. They found that all of the classes met the complex instruction standard for percentage of students talking and working together (more than 35%) and all but one class met the student disengagement standard (less than 14%). With the group work meeting this standard, it was deemed "good" group work. The researchers found that all the science content post tests were significantly higher than the unit pre tests. Similarly, the students scored significantly higher on the post test high order questions when compared to the pre test high order questions. The researchers also analyzed the open-ended questions to assess student misconceptions, ability to draw scientific diagrams, and ability to construct reasoned arguments. They found a number of students confusing consequences with the cause, lack of understanding the concept of experimental control, and mixed success on diagram/ drawing questions. The researchers also found that reading scores on the standardized test were significantly correlated with pre and post test scores on all of the unit tests. Students that had higher reading achievement also had higher gain scores on two of the unit tests, but not on the other two. In sixth grade, girls scored significantly higher than boys on two of the unit tests, while the gain scores for girls and boys on the other units were not significantly different. In eighth grade, boys had significantly higher gain scores on one unit test and an overall higher average score on all the unit tests.

There were a number of strengths and weaknesses in the study. Researchers achieved a 90% reliability with the Whole Class Instrument measure. The researchers were trying to make success on the science unit tests not dependent on reading ability, however their analysis shows that students with higher reading ability had higher unit test scores. The researchers did add pictures and diagrams to the unit tests in hope of triggering other abilities students may possess. The length of the units was different and they found that students had the highest gain scores on the shortest unit. The researchers also did not achieve inter-scorer reliability with the students' picture drawing and diagramming unit test questions. The researchers due to lack of resources did not have a control group.

There are conclusions we can draw about the achievement of students while implementing complex instruction. This method did show that students learn on a variety of different level cognitive demand such as recalling facts, conceptual knowledge, and synthesizing information. The researchers also concluded that open-ended questions and diagram/pictorial questions are hard to score but show a great deal about the misconceptions and thought processes of the students. Also, they found mixed achievement of girls and boys within the complex instruction model, showing that the gender gap was lessened within the teaching model. The researchers also highlighted the challenge in creating curriculum and corresponding tests that focus on the specific knowledge in the content area and not solely on reading ability. This study did not show how the learning of students in complex instruction compared to learning in traditional classrooms. Boaler and Staples (2008) compared traditional instruction and complex instruction to analyze academic and social gains students make in complex instruction math classroom environments.

Boaler and Staples (2008) completed a five year longitudinal, mixed methods study looking at differences in student achievement and achievement gaps among ethnic groups at one school employing reform methods to teaching math, Railside, when compared to two other schools employing more traditional teaching methods, Greendale and Hilltop. The main goal of the study was to add to the literature a detailed view of what equitable and successful teaching looks like in a high school math setting. They also included quantitative data about student success and achievement to show that the teaching practices of complex instruction were indeed worth investigating.

The participants of this study were about 700 students from three different California high schools. The three schools were of similar size and all had dedicated math teachers. The schools were however in different locations and had different student demographics. The main treatment group of this study was Railside High School. This school was in an urban neighborhood and had a diverse student population with students from a variety of ethnic and cultural backgrounds (40% Latino/a, 20% African American, 20% White, 20% Asian/ Pacific Islander). Railside had a high percentage of English Language Learners (30%) and students qualifying for free and reduced lunches (30%). The teachers at this school implemented a reform-oriented curriculum in math classes that focused on conceptual problems and utilized group work and complex instruction strategies. Greendale High School was a school close to the coast and very little ethnic or cultural diversity (90% White, 10% Latino/a). This school had no reported English Language Learners and has a fairly low rate of free and reduced lunches (10%). Hilltop High School was in a rural area with some student ethnic and cultural diversity (60%

White, 40% Latino/a). This school had a population of ELL students (20%) and moderate level of free and reduced lunches (20%). Both Greendale and Hilltop teachers implemented a more traditional curriculum including demonstration by teacher and short practice problems. These two schools did give students the option of taking IMP math that is more similar to reform-oriented math. Each school had a very small number of students participating in IMP math and were not included in the study (although their test scores are factored in for standardized state tests).

The three schools were selected purposively. These three schools allowed the researchers to compare schools with different levels of diversity and different teaching models. The Railside students were not allowed to choose their math placement. All freshman were enrolled in heterogeneous, reform-oriented algebra classes. In Greendale and Hilltop students were allowed to chose their level of math when entering high school or test into a math class. The teachers at Railside worked very closely together and designed the curriculum used in the math classes. The main comparison groups were approximately 300 students from Hilltop and Greendale in the traditional curriculum approaches and approximately 300 from Railside in the reform curriculum and teaching methods. Class sizes were similar between the different schools.

The researchers gathered both quantitative and qualitative data from the schools. The data were intended to give a clear picture of the teaching approaches and the classroom interactions, as well as students' views on math and student achievement. The teaching approaches and class interactions were videotaped for approximately 600 hours of lessons. The researchers from this data wrote "thick descriptions" about each teaching and learning occurring in the class. The researchers also analyzed the time spent in each lesson into categories of teaching talking, teacher questioning whole class, students working alone, and students working together. From these videotapes researchers noted the amount of time spent on each problem. The teachers' questions asked during lessons were also divided into categories such as probing, extending, and orienting. Video coding was only performed for Year 1 of the study due to lack to resources. To assess student beliefs and relationships with math the researchers interviewed at least 60 students each year. The students chosen included both high and low achieving students and students of different cultural and ethnic groups. The researchers also administered questionnaires to all the students in the study Years 1, 2 and 3 assessing experiences in math classes and their perceptions about mathematics and learning. Lastly, the researchers also analyzed student understanding of math content through a content-aligned test and open-ended project. The researchers administered a pre-assessment when students started Year 1, at the beginning of high school, to gain a baseline of where the students were at on math skills supposed to be mastered at the end of middle school. At the end of Year 1 and beginning of Year 2 the researchers gave an assessment on algebraic topics. At the end of Year 2, the assessment evaluated algebra and geometry knowledge and skills. At the end of Year 3, the assessment included algebra, geometry, and some more advanced algebra topics. The researchers wrote the content tests, but had teachers from all three schools review the assessments to make sure they were aligned with the content being taught in the corresponding courses. The open-ended projects were longer and more applied problems that were given to students to work out in groups. These projects were administered at the end of Years 1, 2, and 3. The researchers also collected data on the students from the state administered tests such as the CAT6 and the

California Standards Test of algebra.

From the numerous sources of data, the researchers found many conclusions. They found that at Greendale and Hilltop, in the traditional teaching method, students spent the majority of time in class listening to teacher lectures (21%), being questioned from the teacher in a whole class situation (15%) and working individually on practice problems (48%). Of the time when teachers were questioning approximately 97% of the questions were procedural. At Railside the students spent the majority of their time working with other students (72%) with minimal time spent listening to teachers' lectures (4%) and whole class questioning (9%). When teacher questioning was analyzed they found that 62% of the time it was procedural, 17% conceptual, and 15% probing. These detailed analyses were derived from Year 1 of the study. The classroom environments did not significantly change over the course of the study.

There were also interesting results from student achievement data analyses. They found that Railside students at the beginning of Year 1 scored significantly lower on the content assessment than Greendale and Hilltop ($t = -9.14$, $p < .001$, $n = 658$). At the end of Year 1, Railside students scored on average slightly lower on the content test than Greendale and Hilltop, but not significantly lower ($t = -2.04$, $p = .04$, $n = 637$). At the end of Year 2, Railside students scored significantly higher than Greendale and Hilltop on the content assessment ($t = -8.30$, $p < .001$, $n = 512$). By the end of Year 3, Railside students had a higher average score on the content assessment, but not significantly higher than Greendale or Hilltop ($t = -1.75$, $p = 0.08$, $n = 420$). The researchers explained that the curriculum for more advanced classes were not as developed as that used in algebra courses. In addition to having positive achievement results, Railside also had a higher number of students enrolled in advanced math classes like pre-calculus and calculus when compared to Greendale and Hilltop (Railside, 41%; Greendale and Hilltop, 27%).

The researchers also analyzed the achievement gap between students of different ethnic or cultural backgrounds. They found that at the beginning of Year 1 at Railside, Asian, Filipino, and White students each were significantly outperforming Latino and African American students ($p < .001$). In the course of just one school year, there were no longer significant differences between the achievement of White and Latino students, or Filipino students and Latino and African American students. The only significant difference that remained was between White and African American students and Asian students and African American and Latino students. At the end of Year 2, the only achievement difference that remained was between Asian students and Latino and African American students. There were no gender differences found throughout the study.

The Railside students not only had high achievement, but they also reported enjoying math more than students at Greendale and Hilltop. At the end of each year of the study, students at Railside consistently reported having more positive experiences in math classes. At the end of Year 2, significantly more Railside students reported "enjoying math class" (71%) compared to students from the traditional approach (46%). Students are Railside also reported enjoying math in school all or most of the time (54% of students) while a significantly smaller number responded as positively from Greendale and Hilltop. There were also significant results of Railside students reporting intentions to continue to pursue math courses (100% of Railside students; 67% Greendale/Hilltop students) and planning a future in mathematics (39% of Railside students; 5% of Greendale/Hilltop students).

Finally, when compared on California standardized tests there are mixed results. The Railside students scored at higher levels on the standards test for algebra than Greendale and Hilltop. Fifty percent of students at Railside were at or above basic level, while 30% at Greendale and 40% at Hilltop. On the CAT6, Railside students scored lower than Hilltop and Greendale. Boaler and Staples (2008) explained that some of the lowering in test scores may be due to cultural and linguistic barriers present in the state test. Correlations were found between language arts scores and math scores across the whole state of California, suggesting that the math sections are inadvertently testing language as much as math skills and knowledge. This argument, according to the researchers, cannot be made in the other direction since the language arts tests contain no math.

Given the wealth of information from this study, this study was very carefully implemented and the researchers were very mindful of the design of the study. They used both qualitative and quantitative data to create a more complete picture of what is happening at both the reform-oriented school and the traditional schools. They also wanted to look at the effects of complex instruction and reform teaching practices over time. The researchers carefully selected schools there were of similar size and all the schools taught the same sequence of courses. To make sure that data collected were valid, the researchers performed member checks, had a minimum of two people separately analyze the data, and the researchers also reported results to the research team for review. The researchers did not continue observing and collected as detailed data as the years of the study went on. There was a decline in the significance of the achievement between Railside and the traditional schools. With more detailed data collected for Year 3, the researchers may have been better able to explain what was going on that year at Railside. Also, the researchers stated that curriculum was not as developed for higher math courses. All in all, the researchers were very careful with the design of this study to ensure that they were able to create accurate images of what was going on in these different schools.

From this very large-scale study, there are many implications for better teaching math in a high school setting equitably. The researchers first emphasized that it was not only the curriculum changes that the teachers made at Railside that positively impacted the students, but also how they interacted with students and the entire set up of math education at the whole school. To have this type of impact it took a whole math department's effort and dedication to the reform. On a classroom level, this article also highlighted giving students more responsibility over their education, implementing complex instruction strategies, being mindful of questioning techniques, keeping the cognitive demand high, giving students more conceptual based, longer problems to complete, and emphasizing effort over ability.

Albert and Jones (1997) completed an "interpretive collaborative" case study analyzing the implementation of complex instruction. They were specifically looking at the collaboration between two teachers both implementing complex instruction for the first time while connecting complex instruction methods to science standards. Through the case study the researchers found that collegial relationships are very important especially when implementing new teaching models.

The teachers highlighted in this case study were two Central Kentucky teachers that decided they wanted to change their teaching practice to involve students more in the learning process. Casey and Bennett were both primary teachers that were currently working in a school with non-graded classes with students ages eight to 10. The students taught were 40 children (23 boys and 17 girls) within a school with 13% minority. They acted as both teacher and researcher in this study. The data were collected over a nine-month period during the September through June school year. They volunteered to participate after completing a one-week intensive Complex Instruction Professional Development Institute and were seeking a different way to approach cooperative learning. Both teachers, Casey and Bennett, have taught for 20 and 15 years respectively and taught together for the past four years. The teachers collected data and documented the project in their field notes while the researchers Albert and Jones (1997) supported the teachers with classroom visits (formally eight times each for 90 minutes), conferences, and planned discussions. Casey and Bennett worked closely together throughout the school year, with their classrooms only separated by a collapsible wall. The researchers analyzed observational data, discussion notes, and reflective journals in attempt to determine the teachers' understanding of complex instruction and the relationship to science teaching standards. The teachers also had continual professional development by participating in a yearlong program that involved class observations with feedback about the implementation of complex instruction.

From the journals there were a number of important findings. The teachers remarked that they felt overwhelmed preparing materials and activities. Both Casey and Bennett worked hard to get everything ready for the school year during the summer.

Casey remarked that all of the laminating, cutting, coding, planning, making badges, etc took "an incredible amount of time" (Albert & Jones, 1997, p. 286). The teachers also spent time at the very beginning of the year to orient students to the roles, tasks, and concepts that were going to be learned in their groups. Both teachers were disappointed in their beginning attempts to implement complex instruction. Casey remarked, "The first lesson was horrible. The only behaviors I observed were the usual arguing and compromising" (Albert & Jones, 1997, p. 286). Despite their first attempt, the following activities went much better. The teachers both noticed that students were starting to work together differently and even started communicating differently. Students would use phrases like "Let's make a plan" or "No one is done until everyone is done." Students exhibited self-regulation and used democratic processes for decision-making. They also noticed that changing roles helped different students take leadership roles (wider variety of students taking leadership roles, not just older students). The teachers noted that their classrooms had become a more caring environment throughout the school year. Groups were self-regulating and holding sophisticated discussions. Both teachers after participating in this study decided that they would never teach science or any other subject as they had in the past. They felt that within the complex instruction model, they had created a community of science learners and fostered collegial relationships in a way that had not been achieved before. The researchers also found that complex instruction mirrors many of science standards including children engaging in inquiry, collaboration, and building a community of learners and collegial relationships. The activities were also intellectually rigorous and framed in ways that encouraged students to construct meaning for themselves.

In light of classroom observations, the teachers also made observations of changed behavior in individual students. Bennett remarked that she saw status changes happen in her classroom. One of her students had been very shy and reluctant to participate in the beginning. Throughout the year, this student took on leadership roles and actively participated in group activities. She also overheard students complimenting one another's abilities during wrap-ups. Bennett also observed students asking each other questions, instead of telling group members they were wrong. Overall both teachers observed a heightened ability to express ideas, concerns, and productively communicate with each other in general.

The researchers specifically chose these teachers for this case study because they are able to work so closely together. Unlike other research studies, the teachers were not given aids to help them implement complex instruction. This creates a more realistic view of what it would be like for a "real" teacher to implement a new teaching model. The researchers also used triangulation to draw conclusions from this study. They conferred the teacher notes and journals, formal observation notes, as well as the conferences and discussions between teachers and researchers to ensure the validity of the conclusions. The teachers also taught children that were heterogeneous in academic ability and development, especially since some children varied in age by two years. It is important to note that all teachers may not have the same experiences when implementing complex instruction for the first time. The case study is not widely generalizable, but does give a more detailed and clear picture of what exactly was going on in the classroom. The academic achievement of the students was not measured in tests scores or other measures. Similarly there is no record of a quantifiable change in the students' behavior, only account from the teachers' journals and reflections.

This case study showed that complex instruction can have an impact on even young children's social skills. Other quantitative studies show that complex instruction can have significant effects on testing skills and math class retention rates, however this study shows more personally how it changes the children's interactions with each other. This case study is consistent with other research on the effects of complex instruction disrupting status and making group interactions more equitable (research studies previously explored). This study showed a very candid view of teachers' struggling to implement a new teaching model. The conclusions about implementing complex instruction are addressed in the following section of this chapter.

Complex instruction has been shown through different researchers to be an effective model of teaching to evoke both academic and social gains with heterogeneous students. Bianchini, Holthuis, and Nielsen (1995) found within science classrooms students had academic gains in high school science classes utilizing complex instruction teaching techniques. Boaler and Staples (2008) compared high school students in traditional and complex instruction math classes. They tracked standardized test scores, the number of students taking upper division math courses, as well as achievement differences between minority students. Albert and Jones (1997) found that students had a number of social gains from being within a complex instruction classroom such as more effective communication skills and playing different roles within group work settings.

Implementation of Complex Instruction

The research on the implementation of complex instruction reveals the importance of the role of the teacher as an active participant within the classroom. Lotan (1989) found that there were relationships between teacher conceptual understanding of complex instruction and teacher facilitation, percentages of students talking and working together, and the amount of time teachers talked about cooperation and role taking. Lotan and Cohen (1994) found that teacher direct supervision of student groups negatively impacts the proportion of students talking and working together. Similarly, Cohen, Lotan, and Leechor (1989) found that the more direct supervision a teacher uses the lower the rate of student lateral communication. Scarloss (2002) found that group assessments can capture a wider picture of the collective group knowledge than essay and multiple choice tests individually taken. Lastly, Perrenet and Terwel (1996) found that when teachers do not implement complex instruction with high fidelity, students do not show as high of gains in academics and social relations. Teacher implementation of complex instruction is pivotal in creating an environment of risk taking, taking on group roles, communicating with peers, sharing talents with the group, and ultimately learning.

Lotan (1989) completed a quantitative study in which they researched how complex instruction theory was being implemented in the classroom. The main hypotheses of this research were that teacher conceptual understanding of complex instruction would positively relate to the frequency of non-routine behaviors in instruction, productivity in the classroom as measured in completion of curricular units, percentage of students talking and working together, teacher talks about cooperation and roles, and transferability of skills learned in certain problem solving situations to other, similar problem solving situations. Non-routine behaviors are defined as instructional practices by the teacher that are highly differentiated for groups of students or individuals, student tasks are varied and open-ended, and activities and materials are nonstandard. They also predicted that teacher conceptual understanding of complex instruction will negatively relate to teacher facilitating during group work time.

The participants consisted of 390 second, third, fourth and fifth grade students from the San Jose area in the 1982-83 school year. The majority of the students were Latinos, with white, Asian, and African American students present as well. Most of the students were deemed from working to lower middle class. The researchers chose 15 classrooms in 10 schools from three different school districts to implement complex instruction and the Finding Out/ Descubrimiento curriculum (FO/D). Teachers received a two-week training on how to implement this method of teaching including the theoretical framework and rationale as well as time to practice teaching using this method and curriculum.

The researchers regularly observed the treatment teachers measuring the teachers' use of routine and non-routine teaching behaviors, counting the number of students in each learning group, and tracking student participation into different categories. Teachers were given specific feedback about their implementation in January 1983 based on observations. Also, teachers had additional support from Stanford trainers and attended monthly meetings about their use and implementation of complex instruction and FO/D. The researchers administered a questionnaire in June 1983 to the treatment teachers that consisted of both open-ended and forced choice questions to assess the teachers conceptual understanding of complex instruction and FO/D.

The researchers found a number of interesting results, having direct implications to their hypotheses. They found that there was a negative relationship found between teacher conceptual understanding and teacher facilitation during group work ($r = -.62$, $p <$

.01). That is, teachers who had better conceptual understanding of complex instruction facilitated less during group work. The researchers also found there to be a significant relationship between teacher understanding and the percentage of students talking and working together ($r = .50$, $p < .05$) and time the teacher talked about cooperation and role taking ($r = .47$, $p < .05$). Through these findings, Lotan (1989) concluded that teacher conceptual understanding of complex instruction is an important factor in the implementation of complex instruction.

There were a number of areas that may complicate the results of this study. The researchers mentioned that the second grade students in the treatment group struggled reading the task and role cards and thus did not complete as many units as older children in the study. There was also a low number of participants in the study overall. The researchers expressed a lack of being able to accurately measure the conceptual knowledge of the teachers about complex instruction and FO/D. Despite the critique, Lotan (1989) did highlight the importance of teachers having strong conceptual understanding of teaching methods and curriculum they are employing. This also connects to the need of teachers having support of other teachers, administration, and ongoing professional development to help continue education about the theoretical ideas behind teaching methods and curriculum choices, specifically complex instruction. Additionally, Albert and Jones (1997) found that teachers needed continuous professional development and collegial support to effectively implement complex instruction.

Lotan, Cohen, and Holthuis (1994) completed a quantitative study in which they researched a number of aspects of complex instruction through four hypotheses. The first hypothesis was that given an ill-structured problem and a group worthy task, group productivity would depend on student interaction. The second hypothesis was that direct supervision would negatively relate to the proportion of student talking and working together on the group task. The third hypothesis was the number of different activities in simultaneous operation would be negatively correlated with direct supervision. The fourth and final hypothesis was that size of the small group would be directly related to the percentage of students observed talking and working together in the group. Each hypothesis was directly addressing aspects of complex instruction that the teacher can control such as size of group and how often direct supervision is implemented.

The participants of this study were 42 middle school classrooms in the Bay Area in the 1991-1992 school year. The classes were chosen from five different school districts. The students in the study were racially and ethnically heterogeneous and representative of California's demographics at that time. The researchers observed teachers in math, social studies, and human biology classes. The classes chosen for the study were also tested for academic range, and the researchers purposefully chose heterogeneous classes at some schools and a heterogeneous and homogeneous comparison groups at other schools. Teachers participating the study received two weeks of training workshops on complex instruction. Both complex instruction and comparison social studies classes received a pre and post assessment to measure gains in knowledge throughout the school year. The other subjects were not able to complete both the pre and post test due to lack of resources. Students in the complex instruction classrooms participated in cooperative learning skill-building exercises before working in groups to teach students how to work together more effectively. The observers in the complex instruction classrooms measured the number of students per group, the amount of direct supervision given by the teacher, the number of different tasks simultaneously going on at once, and the student interactions.

The research team had a number of interesting findings pertaining to the implementation of complex instruction. Three of the four hypotheses were supported. Thus they found average amount of interaction per classroom (students talking and working together) was positively correlated with average achievement gains per classroom ($r = .50$, $p < .01$). They found that direct supervision by the teacher negatively correlated to students talking and working together on the task ($r = -.52$, $p < .001$). They also found that the number of different activities operating simultaneously was negatively correlated with teacher direct supervision ($r = -.40$, $p < .01$). Hypothesis four speculated that the smaller the group size the higher percentage of students working and talking together. Lotan, Cohan, and Holthuis (1994) actually found a significant relationship between group size and interaction in the opposite direction than originally thought. After considering the classroom dynamics, the researchers decided that groups with only two or three members may not have the variety of intellectual resources to pull from to complete the complex task. Also, there were no groups that were smaller than 2 or larger than 5 in the study. The researchers suggest that there may be an upper and lower limit to group size given these findings, keeping groups to four or five members.

The researchers overall were very mindful when designing the study. The researchers used a large number of schools, student populations, and controlled for using untracked classes with heterogeneous populations. The researchers gave the same test as a pre and post assessment to both the treatment group and comparison group. The teachers were also consulted as to how to modify the test in future research. The researchers measured the percentage of student gain to control for prior knowledge of students, taken from the pre assessment. The assessment were comprised of a combination of conceptual and factual questions. The results found confirmed a number of other studies looking at complex instruction implementation. The researchers were not able to conclude very firmly any differences between the complex instruction group and the comparison group due to the lack of giving only the social studies classes both the pre and post assessment. This also was not a main focus of the study. To further highlight the impact of complex instruction on student learning, a larger group of pre and post assessment administered among all the subjects would be needed.

The researchers in this study were focusing on teacher actions and behaviors that were conducive to student learning or impeding student learning in a classroom that uses complex instruction. According to this research teachers should avoid direct supervision of group work, group students in numbers of four or five, have different groups working on different activities if possible, and also faithfully implement complex instruction. Teachers that have a harder time straying from direct supervision of group work can use tactics such as having many different activities going on at one time so that he or she simply cannot directly supervise all the different groups. This research emphasized the importance of teacher interactions and forethought that must be put into each complex task and grouping.

Cohen, Lotan, and Leechor (1989) completed a mixed design study in which they were examining teacher behavior and how that effected student interaction and ultimately achievement. The research team had three main hypotheses and one corollary, all of which assumed that the tasks given to students were open ended and deemed “uncertain.”

The uncertainty of a task is defined as classroom tasks that are different from routine tasks and do not have a standardized procedure for completion. The first hypothesis speculated that when there are different tasks going on simultaneously within the same classroom, that teachers are less likely to use direct supervision (and in turn delegate more authority). The second hypothesis was that the more a teacher used direct supervision, the lower the rate of lateral communication would be among students. The third hypothesis was that the rate at which students talk and work together would be correlated with gains on achievement tests, especially in areas of conceptual learning and problem solving. The corollary from hypothesis three was that the rate of students talking and working together would be positively related to a reduction in the variation of the distribution of the post assessment scores compared to the pre assessment score distribution.

The participants consisted of students in second through sixth grade. The children were generally from low socioeconmic status family backgrounds and many students had limited proficiency in English, while others had limited proficiency in both English and their first language (Spanish within this sample). The researchers collected data from 15 classrooms in 10 different schools in 1982-83 school year and 13 classrooms from five schools in 1984-85. Each teacher involved in the study was paired with an assistant to work together with in the classroom during the study. Each teacher and assistant pair completed a two-week training in complex instruction. The teachers used the Finding Out/ Descubrimiento curriculum (FO/ D). The researchers administered pre assessments before the treatments started called Math Concepts and Applications.

The researchers observed the teachers using complex instruction and collected data in 10-minute observation intervals. During the 1982-83 and 1984-85 school years, teachers were observed approximately 10 times and 20 times respectively. Researchers also recorded how many students were at each learning center and the number of different learning centers occurring simultaneously. In 1982-83 the researchers categorized teacher interaction with students as either facilitating or not facilitating. In 1984-85 the researchers added more categories for teacher interactions with students to better capture what was happening in the classroom. The researchers also noted student disengagement and the amount of students talking and working together. Lateral communication between students was measured in 1984-85 by observing 10 target students in the classroom. The researchers also administered a variation of the Math Concepts and Applications test as a post assessment after the completion of the school year. Finally, the researchers also collected standardized test scores of the students that participated in the study.

The researchers found that all of their hypotheses were supported. The ratio of learning centers to students was negatively related to facilitation by the teacher in 1982-83 ($r = -.50$, $p < .05$) and direct supervision in 1984-85 ($r = -.55$, $p < .05$). That is, when there are more learning centers and different activities going on in the classroom, the teachers simply cannot directly facilitate or directly supervise the group work. They also found that direct supervision was negatively related to the percentage of students talking and working together in 1982-83 ($r = -.43$, $p < .05$) and the target children talking and working together in 1984-85 ($r = -.49$, $p < .05$). The measure of communication among students showed a significant and positive correlation to the average gains on the test of Math Concepts and Application in both the 1982-83 school year ($r = .72$, $p < .01$) and the

1984-85 school year ($r = .52, p < .05$). Lateral communication between students was positively correlated with a gain in scores on the math computation section of the assessment in 1982-83 ($r = .61, p < .05$), but not for the 1984-85 school year.

Despite having their hypotheses supported there were some holes in the research design and procedures that were left unexplained. The researchers did not categorize teacher interaction and/or questioning into levels of thinking it provokes. For example, the cognitive demand of questions asked were not analyzed. To not specifically categorize the type of questions asked can provide misleading data about teacher interactions since not all questions promote the same type of thinking. Throughout the studies there was little variation in the number of students per learning center (typically 5) and there was also little variation in the number of different learning centers per classroom. However, the researchers used this variation to show support for a hypothesis about teacher facilitation and direct supervision. Given their large pool of data, only two sets of classrooms were used to test the hypotheses. The researchers did not specify how these classes were chosen or why not all of the data were used to test the hypotheses. The researchers changed the operational definition of students talking and working together from year 1982-83 and 1984-85. The researchers also do not address if the same teachers or different teachers were used in each separate school year or why data was not collected in the school year 1983-84. The researchers also did not explain how they chose schools, teachers, or target students for the study as well. The researchers did conceded that alternative hypotheses may explain the gains in achievement due to the small sample size, although they refuted that this could have been caused by the "Hawthorn effect."

Although this study has some areas of question, the researchers findings are similar to a number of other studies on the teachers' role in setting up the classroom and interacting with students during complex instruction tasks. The researchers concluded that students interacting with each other and the material is related to learning. They also followed this conclusion with, "Students talking and working together is not a panacea for improving learning in the classroom" (Cohen, Lotan, & Leecher, 1989). They also stated that learning tasks given to groups must be thoughtfully designed, contain elements of uncertainty, and be truly groupworthy. The researchers found that teachers need to encourage students to work together and delegate authority to students. This research emphasized that teacher interactions with students are vital and to fully implement the complex instruction model, teachers need to rely on class norms and group roles to help manage the classroom instead of falling back on direct supervision.

Scarloss (2002) completed a study in which they developed group performance and product rubrics to assess student group work. When implementing complex instruction, teachers run into problems assessing students. Some fall back on averaging individually taken multiple-choice tests, while others give group tests. The researchers proposed that by scoring the group product or performance, based on both factual and conceptual knowledge, teachers could get a more full picture of the collective group. The research questions were "How would you assess academic performance at the group level" and "Why not just aggregate individual scores" (Scarloss, 2002, p. 3). They also hypothesized that rubrics would need to focus on qualities of the group work instead of the amount of work completed.

The participants were 163 sixth grade students in five different classes. The students were from multiracial, multiethnic, and largely low socioeconomic status backgrounds and lived in the California Central Valley. Many students had limited English proficiency (25%), students averaged the 34^{th} percentile on the standardized reading tests SAT-9, and many were children of immigrant or migrant agricultural workers. The teachers involved in the study all completed a 10-week training in complex instruction and were deemed skilled at implementing complex instruction. Three of the teachers had also received advanced training in complex instruction and four of the five teachers had completed advanced training in curriculum development. The teachers were chosen purposely on the criteria that they had effective classroom management skills, solid content knowledge and understanding of the curriculum, and successful prior implementation of complex instruction. All of the teachers taught at year-round schools and used the same curriculum during the study. Each teacher taught four complex instruction units. Three of the classes in the study used the new evaluation criteria while the other two did not. Each teacher was observed during the last unit, which included five lessons. Observers came into the classroom during lessons one, three, and five. The researchers collected field notes, recordings of the classes, photos of group work and presentations, and group products. Both Scarloss (2002) and an outside researcher scored the group performances. The rubrics developed from the literature included a developmental and a task-specific rubric.

The rubrics were focused to assess specific academic content of the group performance or product as well as reflect intrinsic characteristics of the medium the students presented their knowledge. The rubrics were to be explicit and clear so that they would be equitable assessments. The rubrics could also be given to students to make more explicit what information to include and focus on during the presentation or creating the product, after more unstructured exploration. The rubrics were broken into the categories of concrete content, conceptual content, and presentation conventions. Each scorer on the rubric marked the performance or product on a four-point scale for each category. Scarloss (2002) also had students complete individual performances on essay tests and multiple-choice tests. The analysis she found that the rubric scores were correlated with scores on the essay test and the multiple-choice tests, although the two individual measures were not correlated. This showed that group performance and product scores encompasses both the individual scores showing that “a group is more than its parts” (Scarloss, 2002, p. 20). The researchers controlled for reading scores and found that group performance predicted essay performance.

There were a number of strengths and weaknesses in this study. The author had another person also score the group performance and products to reduce bias and she tested the inter-scorer reliability of the rubrics were tested. Scarloss (2002) accounted for variance by controlling for reading ability as the literature suggests. The researchers had a small sample size of only five social studies classes. Also, they used the rubric scoring with the last unit only. It would have added to the findings to show change rubrics being used to score group performances and products throughout the different units for comparison.

Using rubrics allowed students to show their range of intelligent abilities, which is more consistent with the goals of complex instruction. Assessing the group in part also shows students that the collective knowledge of the group is important and this could build more interdependence and motivation to work well together. When implementing complex instruction, thoughtful assessment is important to truly reflect what the students are learning and doing in their groups. While it seems important to have students be individually accountable for knowledge, using rubrics similarly outlined in this study could help give a more full and representative picture of what is going on in the classroom. In addition to using rubrics, Stipek et al. (1998) found that substantive written feedback from teachers (instead of a letter grade or check mark) gave students heightened feelings of competence and overall positive attitudes about the class.

Perrenet and Terwel (1996) completed a qualitative case study in which they researched the degree of fidelity that teachers implemented complex instruction after receiving training and support from complex instruction coaches. The main question they were researching was, “Has complex instruction been implemented in the classroom process and in teacher guidance?” This study was completed in the Netherlands and had support and materials from Stanford University where complex instruction was developed.

Two teachers were chosen to participate in this study. For the formative evaluation, one class comprised of 10 boys and 13 girls were observed. The teacher of this class implemented complex instruction in both a math and “mother-tongue” class. For the final evaluation, two classes were observed, both with 21 students. One class implemented complex instruction for math class (nine boys and 12 girls) and the other in the mother-tongue class (11 boys and 10 girls). Students lived in the Netherlands and a third of the students were of non-European background, mostly from North Africa (Morocco). The teachers in this study were given complex instruction training before the school year from Hogeschool Midden Nederland with one of the coaches from Stanford to train the other coaches. Teachers were also provided with a personal coach to support their adherence to complex instruction. The curriculum used in the treatment was developed by the Hogeschool Midden Nederland and one unit was based on the curriculum Stanford's complex instruction team use in their research.

The Perrenet and Terwel (1996) research team observed lessons and transcribed teacher-coach guidance sessions. During lesson observations student utterances were audiotaped and the researchers took notes on the main aspects of complex instruction: multiple-ability tasks, status treatments, three stages of lesson (class orientation, group work, and class wrap-up), group work roles, group work rules, heterogeneous group composition, and intercultural education. The researchers collected data from 20 (out of 68) lessons and six (out of 16) teacher-coach guidance sessions. From the data collected, the researchers analyzed the level of compliance with complex instruction and categorized it as low, medium, or high.

From the analysis, the researchers found that the teachers' fidelity of the implementation of complex instruction varied depending on the specific aspect. When looking at the formative and final evaluations, they found the implementation of multiple-ability tasks high, status treatments low, three stages of lesson high, group work roles medium, group work rule medium, heterogeneous group composition high, and intercultural education medium. The researchers overall deemed the implementation of complex instruction in these classrooms not as successful as in the United States. The researchers did note that after specific guidance, the teachers did in fact work to change their behavior in the classroom. The researchers explained the low implementation of the status treatments may be due to the cultural differences between American culture and

Dutch culture. According to the researchers, teachers in the Netherlands are reluctant to praise students, fearing that their peers may reject the students for being the "teacher's favorite." The researchers also state that the teachers may not have fully understood all the aspects of status treatments. They also found that students cooperated more after having some time to practice taking on the roles. The researchers speculated that a longer training period for the teachers would also be helpful (their training period was shorter than what Stanford typically employed for training complex instruction teachers).

There were some areas of this study that were not fully explained. The researchers mentioned that each complex instruction classroom was given a poster of group work rules to put up in the classroom but that none of the teachers used the posters. This shows that there may have been a lack of communication with the teachers and coaches from the beginning of the study. Also, the study quoted specific questionable teacher-student dialogue taken from complex instruction lessons. The dialogue showed a teacher accepting student interactions that were not congruent to their assigned roles in the group, although there was no mention in the article about coaches catching this or having a conversation about this issue with the teacher. The study was also very small, only involving two teachers in one school and there was no mention of how they chose the teachers or the school for the study. This study does not seem to be generalizable to all Dutch teachers. There needs to be more research done in the Netherlands with longer training periods and proactive coaches to make certain that low status treatment implementation is in fact due to a culture difference and not oversight. The researchers did mention that Dutch classrooms are deemed more "feminine-oriented" where praise is rarely given. Thus the status treatments may have seemed strange to the teachers and not as easy for them to implement. Finally, the researchers broke the lessons and teachers into the categories formative evaluation and final evaluation. There is no mention as to what the differences are between these groups and if the teacher from the formative evaluation group was part of the final evaluation group (group consisted of two teachers).

Although this study had some points of concern, the researchers overall had interesting conclusions that apply to the implementation of complex instruction in the United States, as well as other countries. They found that when students are allowed more time to practice taking on the group work roles, they cooperate with each other more. They also highlighted areas of complex instruction that may be harder to implement such as status treatments, group work roles, and group work rules. These three aspects of complex instruction take very specific and ongoing attention by the teacher and takes practice to make those aspects part of the daily classroom environment. Since the teachers in this study had abbreviated training in complex instruction, it seems that they may have less conceptual knowledge of the underlying theory behind the different aspects of this teaching method. Lotan (1989) emphasized that it is important for teachers to have this conceptual knowledge base before trying to implement complex instruction. This may be one reason status treatments in this study were not implemented. This study also pointed out that parts of the method may be inconsistent with some cultural practices and norms. As a teacher, it is important to be aware of the effect particular teaching practices are having on students.

Given the work of Lotan (1989), Lotan, Cohen, and Holthuis (1994), Cohen, Lotan, and Leechor (1989), Scarloss (2002), and Perrenet and Terwel (1996), the importance of the role of the teacher in the implementation of complex instruction becomes clearer. To create a classroom environment that allows students to benefit from the complex instruction model, teachers need have a deep conceptual understanding of complex instruction (Lotan, 1989), delegate authority to students (Lotan & Cohen, 1994; Cohen, Lotan, & Leechor, 1989), minimize direct supervision of group work (Lotan & Cohen, 1994; Cohen, Lotan, & Leechor, 1989), and implement complex instruction with high fidelity (Perrent & Terwel, 1996).

Summary

Chapter three is a review of the literature about complex instruction. The findings of the studies were summarized and critiqued, based on the conclusions provided. The research was reviewed to explore the details of complex instruction, the academic and social achievement of students within this model, and strategies for effectively adopting complex instruction. The research of the Complex Tasks, Conceptual Learning, and Cognitive Demand section indicated that in a complex instruction classroom it is vital that teachers are giving students complex tasks that are conceptual and requires a variety of cognitive demand levels. The research of the Heterogeneous Group Work section indicated that heterogeneous ability groups need to be taught how to work together effectively so that students will get the most out of the collaborative situation. The research of the Status and Competency Expectations indicated multiple ability treatments and status assignments can make group work more equitable. The research of the Students' Beliefs, Self-efficacy, and Identity section indicated that students beliefs about themselves, others, and intelligence can highly impact group work. The research of the Student Academic and Social Achievement section indicated that students have heightened conceptual understandings, enjoyment of math, likelihood of taking upper division math classes, and trying on new roles in group settings. The research of the Implementation of Complex Instruction section indicated that teachers need to have support when adopting a new teaching model like complex instruction as well as give students responsibility for their own learning, follow the processes of complex instruction with fidelity, and make sure to limit group work facilitation. Chapter four outlines a summary of the findings, considers classroom implications, and areas of future research based on the literature presented in chapter three.

CHAPTER 4: CONCLUSION

Introduction

This paper investigates the details of complex instructions and effective strategies for implementing this classroom model. Chapter one examined why complex instruction is worth investigating and sets up different controversies that lie within this classroom model. It explained that many students do not find mathematics in line with their identity and that with modification to tradition teaching methods, teachers can start to create a classroom environment that is more in line with student interests as well as the domain of mathematics. Chapter two identified key areas of history that relate to the development of complex instruction. Throughout the history of math education, practices like tracking have made classrooms mirror society biases and in turn be places of inequity. Through status, multidimensional classroom, and identity development research, the advent of complex instruction was in response to making classroom environments a place where all students are valued, interact with each other in equitable ways, and can have deep learning experiences. Chapter three reviewed the research on complex instruction. The research of each facet of complex instruction, including complex tasks, conceptual learning, and cognitive demand, heterogeneous group work, status and competency expectations, and students' belief, self-efficacy, and identity development, was analyzed and critiqued. Similarly, literature on student academic and social achievement and implementation of complex instruction was also reviewed. The main overall findings of this chapter is that teachers should adhere with high fidelity to the components of complex instruction, students can have substantial social and academic growth within the complex instruction model, and that teachers need continual support when adopting new classroom models. Chapter four summarizes the finding from chapter three, addressed classroom implications from the literature review, and gives suggestions for further research.

Summary of Findings

What is complex instruction? What is student social and academic achievement within a complex instruction model? How does one adopt a complex instruction teaching model? These were the guiding questions for this review of professional literature. One facet of complex instruction analyzed in chapter three was complex tasks, conceptual learning, and cognitive demand. Henningsen and Stein (1997) found that to maintain the cognitive demand of a task, teachers aligned tasks with students' prior knowledge, provided appropriate scaffolding, maintained a high press for justification and explanation, modeled high level performance, and were flexible with time allowed to complete tasks. They also found that classrooms in which cognitive demand declined when teachers lessened the challenges of tasks, gave inappropriate amounts of time for activities (either too much or too little), and classroom management issues disrupt the learning environment (Henningsen & Stein). Stipek et al. (1998) found that teachers who have learning orientation (teachers focusing on students' attention on learning, mastery, encouraging effort and autonomy) also are more likely to have positive affect and not employ differential student treatment. They also found that teachers with learning orientation and positive affect, positively impacted students' focus on learning, understanding, and mastering material, student emotions, enjoyment of the subject, and conceptual learning (Stipek et al.). Kazemi and Stipek (2001) found that teachers press for explanations, emphasizing sociomathematical norms, engaging in whole class discussions, and asking for multiple representations allowed students to express their ideas more clearly in math class as well as engage in more student-to-student dialogue (as opposed to student to teacher). Khisty and Chval (2002) found that it is important for teachers to model sociomathematical norms, engage students in whole class discussions, access students' funds of knowledge, and use bridging language with new math vocabulary. These teacher behaviors allow students to learn the concepts behind procedures while learning how to communicate within the domain of math (Khisty & Chval). Arbaugh and Brown (2005) found that for teachers to effectively change the levels of cognitive demand, they need support from colleagues as well as continuous professional development opportunities.

Another facet of complex instruction examined was heterogeneous group work. Schullery and Schullery (2006) found mixed results when looking at student achievement within heterogeneous and homogeneous groups. They found that groups, hetereogeneous in personality, allowed students greater opportunities to develop communication skills (Schullery & Schullery). Webb, Nemer, and Zuniga (2002) found that high-ability students in homogeneous groups outperformed high-ability students in heterogeneous groups. The researchers and teachers implementing the experiment although did not teach students how to work together or give students complex tasks that required a group to complete. Boaler, Wiliam, and Brown (2002) found that many students in ability-grouped classes were unhappy with their placements and overall did not like math classes as much as students in mixed ability classes. Also, the teachers in ability-grouped classes employed less differentiation strategies and taught overall differently than mixed ability classes (Boaler, Wiliam, & Brown). Cheng, Lam, and Chan (2008) found that when group processes are high quality students believe that they can achievement more as a collective group than individually. Mitchell et al. (2004) found that high school students like to feel in control of their learning environment, including options of choosing their own groups to work with. They also found that students would like, after experiencing both self-selected and teacher-selected groups, teachers to primarily select groups so alleviate pressures to always work with friends (Mitchell et al.).

Status and competency expectations were examined in the literature review as well. Cohen and Lotan (1995) found that students of high status participated in the group activity more than students of low status. They also found that use of status treatments was correlated with more equitable talking and working together between students in small groups (Cohen & Lotan). Tammivaara (1982) found that "status treatments" effectively made the group work more equitable. That is, high and low status group members holding with more equal influence, leadership roles, and top rankings (Tammivaara). Ross and Cousins (1994) found that giving and seeking help were more frequent in the cooperative learning model. They also found that students that were low ability and low status were less likely to ask for help and more frequently ignored (Ross & Cousins). This study highlighted the need for teachers utilizing status treatments to make group dynamics more equitable for both high and low status students. Cohen and Anthony (1982) found that high reading ability, high math and science ability, best friends, and good at sports were all correlated to make a student "high" status. They also found that students were ranks as high status also talked and worked together with others at higher rates and in turn scored higher on the post test (Cohen & Anthony). Cohen and Roper (1972) found "status treatments" resulted in more equitable group work between

African American and White males.

The last specific aspect of complex instruction analyzed was students' beliefs, self-efficacy, and identity. Boaler, Wiliam, and Zevenbergen (2002) found that AP math students did not connect math to their identity or relate this subject as something they were interested in pursuing as a career. Students in traditional math classes reported that they were individual environments, which consist of teachers lecturing and students practicing problem sets. Students that worked in math classes that were more collaborative enjoyed math more (Boaler, Wiliam, & Zevenbergen). Woods, Kurtz-Costes, and Rowley (2004) found that students believed rich people as better at academics, while poor students were better at sports. This study showed how perceptions of others and themselves complexify as students move into middle school.

Rodriguez, Jones, Pang, and Park (2004) found that a summer intervention emphasizing cultural identity, math and science skills, and critical thinking/problem solving helped students feel more competent and capable. When students were in a setting that helped them connect their personal identity with math and science, they felt more able to be successful within those areas. Blackwell, Trzeniewski, and Dweck (2007) found holding incremental theories of intelligence increased effort beliefs, learning goals, low helpless attributions, and positive strategy use in students. They also found that students holding an incremental beliefs about intelligence earned higher grades at the end of eighth grade. Schunk (1982) found giving students past attributional feedback ("You have been working really hard") resulted in higher completion of work as well as higher academic achievement and self-efficacy. Schommer-Aikins, Duell, and Hutter (2005) found that students who believed math was useful had better grade point averages and higher performance on math problem solving exercises. They also found that students that believed learning was fixed (innate intelligence) had lower levels of math confidence, understanding of math concepts, and put less effort into math. Crosnoe, Riegle-Crumb, and Muller (2007) found failing classes and/or being diagnosed with a learning disability have different affects on female and male student learning within math classes. They found that male students that were diagnosed with a learning disability had the most dramatic drops in perceived intelligence, which leads to students taking lower numbers of high-level math classes.

The next body of research analyzed was the student social and academic achievement within a complex instruction model. Bianchini, Holthiuis, and Nielsen (1995) found that students had significantly higher post test after completing a science unit with a complex instruction classroom model. While this study showed students learn in a complex instruction classroom, they did not compare the gains in learning to other classroom models of teaching science. Boaler and Staples (2008) compared two different math departments and their student academic and social gains throughout high school. They found that through math reform classes students achieved equivalent (some years higher) success in math, the achievement gaps was minimized, as well as more students were enrolled in high level math classes (Boaler & Staples, 2008). Also, students reported enjoying math more and foreseeing themselves pursuing math in the future at the school utilizing complex instruction techniques (Boaler & Staples). Albert and Jones (1997) found that students worked more equitably together and took on different group roles in a complex instruction model.

Lastly, the implementation of complex instruction was examined. Lotan (1989) found teachers that had a stronger conceptual understanding of complex instruction implemented it with higher fidelity. Cohen, Lotan, and Leechor (1989) found that the more learning centers and different activities going on at once, the less the teacher can facilitate the group work. They also found that the less teachers directly supervised students working together, the more they worked and talked together. Scarloss (2002) found that assessing group work via rubrics was a helpful way to gauge student individual and group learning. Perrenet and Terwel (1996) found multiple ability treatments and assigning competency were hard aspects to implement with high fidelity and that those aspects of complex instruction may be more relevant for dominant American culture.

Implications for Teaching

From the body of research analyzed in chapter three, there were a number of implications for teaching. Teachers' have a large role in keeping the cognitive demand of math tasks high (Henningsen & Stein, 1997). Teachers cannot simply give students complex tasks and expect students to be developing critical thinking, reasoning, and problem solving skills (Henningsen & Stein). Henningsen and Stein showed that teachers needed to know their students and provide appropriate tasks given students' prior experiences and knowledge. Similarly, to keep cognitive demand high, teachers need to have flexibility to allow for more or less time devoted to tasks and activities as students need (Hennningsen & Stein). It also seems very important for teachers to press students to challenge themselves to justify answers, look for new approaches, and uphold sociomathematical norms (Henningsen & Stein). If teachers give students too much information or shorten the task, the cognitive demand will lessen and students will not gain as much from the activity (Henningsen & Stein). Teachers also need to balance scaffolding for students without giving too much away and pressing students to challenge themselves without leading to unnecessary frustration (Henningsen & Stein). Ultimately, high cognitive demand tasks can provoke deep mathematical thinking and the development of skills such as problem solving and critical thinking, if implemented appropriately by teachers (Henningsen & Stein). Teachers that have high levels of learning orientations and positive affects positively impact their students' enjoyment of the topic, conceptual learning, positive emotions, asking for help when they need it, and focus on learning and understanding (instead of getting a good grade or finishing quickly) (Steipek et al., 1998). To hold high levels of learning orientation and positive affect, teachers need to keep the cognitive demand of the tasks high, press students to give explanations for their answers, use incorrect answers as learning opportunities, give plenty of time for students to work, require students to revise problems or try it from a different approach, avoid sarcasm, do not allow put downs in class, and value all student contributions (Steipek et al.). These strategies and tactics help create a safe environment where students feel comfortable to take risks and ultimately learn (Steipek et al.).

Through the work of Kazemi and Stipek (2001) we see that teachers can reinforce sociomathematical norms and interact with students to promote conceptual learning. Teachers should continually ask students to justify their answers, engage in conversations with other students, represent their answers and thinking in multiple ways, and show proof for answers (Kazemi & Stipek, 2001). Students will make these practices a habit if teachers continually reinforce sociomathematical norms from the beginning of the year. Also when introduced to new concepts, students will be much more engaged if their they are initially hooked by activating previous knowledge or interests (Khristy & Chval, 2002). The new concepts should be contextualized with "word problems, drawings and models, peer discussions, and/or even physical activity" (Khristy & Chval, p. 7). Also teachers need to provide bridging language for their students by using new terms regularly in class and teaching students how to pronounce new mathematical terms (Khristy & Chval). Teachers can do this specifically by continually repeating new vocabulary and having students practice saying the words (Khristy & Chval). Also teachers can provide pictures, definitions, and allow students to think of examples or non-examples of the new word (Khristy & Chval). When introducing new vocabulary, teachers can also bring in students' funds of knowledge by creating specific connections that will spark student knowledge or ask students for connections they can see (Khristy & Chval). Teachers need to remember that students need to know certain vocabulary words to be successful on standardized tests and high-level math classes (Khristy & Chval). To encourage students participating in class discussions, teachers can accept all student responses in group discussions, but rephrase what students say, if need be, in more precise terms or complete ways (Khristy & Chval). Teachers can reinforce this by asking students more questions on the same topic, looking to get more precise answers (Khristy & Chval). Through strategic release of responsibility of classroom dialogue, teachers can help guide students to gradually take over class communication so that by the end of the year students are running class discussions (Khristy & Chval). A model of this is at the beginning of the school year the teacher may talk most of the time, but little by little students spend more time talking and by the end of the school year students engage in discussions with each other (Khristy & Chval). Teachers should spend time modeling and helping students practice academic math talk if they are wanting students to be able to communicate math ideas verbally or through writing (Khristy & Chval). Teachers need to model precise math language and teach students how to communicate in mathematics so that student discussions during group work is of high quality and students can practice this new form of communication (Khristy & Chval). Similar to students not knowing how to work effectively in groups without practice, students do not know how to communicate mathematical ideas unless they practice.

There were interesting implications from the research about ability grouping. When students study for a test or memorize procedural information, high ability students perform better when they are in a similar ability groups (Webb, Nemer, & Zuniga, 2002). Teachers wanting to use complex instruction need to voice resistance to moving to set tracks or ability groups in math classes, due to the variety of negative impacts it can have on students (Boaler, Wiliam, & Brown, 2002). If teachers find themselves teaching ability grouped classes, it is important to remember that these students still have different needs, prior experiences, funds of knowledge, interests, etc (Boaler, Wiliam, & Brown). Also, math teachers need to make sure they do no fall back on regularly teaching from the chalkboard, focusing on memorization, speed, and procedural knowledge that many teachers do without realizing it is ability grouped situations (Boaler, Wiliam, & Brown).

There are a number of aspects about group work that need to be kept in mind while utilizing complex instruction or any type of collaborative group work. Students need time at the beginning of the year to practice taking on the group work roles in a low stakes environment to encourage more cooperation throughout the school year (Perrenet & Terwel, 1996). Also, students should complete extensive community building and get to know each other before working on complicated learning tasks together (Mitchell et al., 2004). Teachers need to make sure that the group work is equitable for all students and that some are not performing better based on gender, ethnicity, cultural background, etc (Schullery & Schullery, 2006). Students need to be taught how to work together and teachers need to regularly and actively address issues of status (Cheng, Lam, & Chen, 2008; Schullery & Schullery). Specifically for complex instruction, it is important for teachers to give students tasks that are high level tasks and that are deemed group worthy (Lotan, 2003). Continual reflection and even employing collaboration among teachers about the tasks given can help ensure that tasks are in fact of high cognitive demand and are truly group worthy (Arbaugh & Brown, 2005). Specifically, teachers need to remember not to give groups tasks that were originally designed for one person to complete on their own (Webb, Nemer, & Zuniga, 2002). This diminishes the learning environment for all the students, low and high ability, in the group (Webb, Nemer, & Zuniga). When implementing group work and complex instruction it is very important to implement with high fidelity not only group worthy tasks, but group roles, and status treatments as well to ensure the group work is beneficial for both low and high ability students (Webb, Nemer, & Zuniga). Teachers need to help students learn how to work in small groups by developing positive interdependence, individual accountability, equal participation, group work norms, and good social skills in order for students to get the most out of group work (Cheng, Lam, and Chen; Mitchell et al., 2004). This can be accomplished by following the guidelines of complex instruction such as group role taking, status treatments, complex task assignment, and allowing students time to discuss and practice role taking. To be developmentally appropriate with high school students' need to take on more adult roles and responsibilities, teachers can give high school students more choices and control of their classroom environment (Mitchell et al.). When implementing complex instruction, teachers should avoid direct supervision of group work, group students in number of four or five, have different groups working on different activities if possible, and also faithfully implement complex instruction (Lotan, Cohan, & Holthuis,1994). Teachers that have a harder time straying from direct supervision of group work can use tactics such as having many different activities going on at one time so that he or she simply cannot directly supervise all the different groups (Lotan, Cohan, & Holthuis). Tasks given to groups must be thoughtfully designed, contain elements of uncertainty, and be truly groupworthy (Cohen, Lotan & Leecher, 1989). By delegating authority to students and encouraging students to work together, teachers can rely on class norms set up at the beginning of the year and grou roles to help manage the classroom instead of falling back on direct supervision (Cohen, Lotan & Leecher).

Status treatment research has pointed directly to the importance of regularly and carefully using status treatments within the classroom (Cohen & Lotan, 1995; Tammviaara, 1982; Cohen & Roper, 1972). Multiple ability treatments can easily be written into lesson plans and given directly with complex task group projects so that the teacher does not forget (Cohen & Lotan). As a teacher, it is important to make sure that assigning competence happens fairly frequently (approximately once every 10 minutes) during mixed ability group work and also that comments are genuine and relevant to the task (Cohen & Lotan; Tammviaara; Cohen & Roper). When students are interacting with each other, teachers need to be continually mindful of status issues when conducting class both within the complex instruction model and when using other models (Ross & Cousins, 1995). Teachers need to be aware that status issues also come into play even when students are working together in pairs (Ross & Cousins). Specific sharing procedures or group steps to follow to make sure that students are continually working together equitably. Status may become more complicated as students progress through middle school and high school as definitions of popularity change, thus status treatments may need to be slightly changed to fit the student population in which a teacher is working (Cohen & Anthony, 1982). Giving students both public and private competence assignments have shown helpful to boost students' self-confidence and participation in group situations (Tammviaara). It is important for teachers to continually remind themselves that status treatments are important and do not negatively affect the participation of the "high status" group members (Tammviaara).

There are a number of actions teachers can take to influence student attitudes, beliefs, and self-concepts in mathematics. Having students practice working together will improve the quality of group interactions and both high and low achievers will have higher collective efficacy and ultimately be more motivated to work together (Cheng, Lam, & Chan, 2008). Teachers need to be mindful that students' hold beliefs about themselves and others similar to stereotypes held in the greater society, especially about socioeconomic status (Boaler, Wiliam, & Zevenbergen, 2000). By working with students to help overcome negative stereotypes about math, teachers can help students increase their interactions with other group member and overall engagement in the math concepts (Boaler, Wiliam, & Zevenbergen). Teachers need to particularly watch for populations found to have declining perceptions of their intelligence, such as boys diagnosed with learning disabilities and girls failing one or more classes, to make sure that they equally participate in group work (Crosnoe, Riegle-Crumb, & Muller, 2007). Also, teachers implementing complex instruction emphasize effort over “natural” ability to help create a more equitable classroom environment (Boaler, Wiliam, & Zevenbergen; Woods, Kurtz-Costes, & Rowley, 2004; Blackwell, Trzeniewski, & Dweck, 2007). Teachers can facilitate opportunities to help students see that they can be good at math, no matter their background or stereotypes they hold about themselves and math achievement (Boaler, Wiliam, & Zevenbergen). A way to promote a positive math identity is by constantly pointing to counter examples of stereotypes and working to not perpetuate those stereotypes in group work by multiple ability treatments and assigning competence (Boaler, Wiliam, & Zevenbergen; Woods, Kurtz-Costes, & Rowley). Teachers can positively impact students’ identity development through the use culturally relevant curriculum and pedagogy (Rodriguez, Jones, Pang, & Park, 2004). To have successful group work and collaboration students need to feel that their whole self is valued and their experiences are important to bring to the group (Rodriguez, Jones, Pang, & Park). Teachers can tap into students’ funds of knowledge to help empower student and show other students different perspectives (Rodriguez, Jones, Pang, and Park). Teachers can increase student motivation to learn new material when they emphasize personal improvement and growth over competition (Rodriguez, Jones, Pang, & Park). High school and middle school teachers also need to recognize the difficulty some students have during transitions from elementary to middle school and also middle to high school (Blackwell, Trzeniewski, & Dweck).

Teachers also need to be mindful of the types of feedback they are giving students. Using rubrics is one way teachers can assess group products or performance instead of having students always take individual test (Scarloss, 2002). Although it is important to have students being assessed both on group work and individual work as well (Scarloss). The rubrics teachers create should be specific and include both developmental and task specific skills (Scarloss). Teachers should encourage the development of incremental theories of intelligence in students by emphasizing the intelligence is not fixed but malleable (Blackwell, Trzeniewski, & Dweck, 2007). Teachers can impact students' views of intelligence theory, incremental or entity intelligence, by the way they give feedback to students (effort vs. ability) (Blackwell, Trzeniewski, & Dweck). Also giving feedback on past achievement in which effort was expended result in the largest gains in motivation, persistence, achievement, and self-efficacy (Schunk, 1982). Teachers need to be very careful of the wording of feedback given to students (Schunk). Teachers need to make sure that when giving negative feedback or failing students to make sure to give explanations in ways that emphasize that students can do better or revise so that it does not impact their intelligence self-perceptions as dramatically (Crosnoe, Riegle-Crumb, & Muller, 2007). Examples of students expending a lot of effect and achieving success need to be highlighted (Schunk). Also, teachers need to make sure that they give effort based feedback on tasks that are either intermediate or difficult, not low difficulty tasks (Schunk). Researchers found that younger students were more likely to hold epistemological beliefs about studying aimlessly, which suggests that younger students do not believe learning is strategic but more related to luck or chance events (Schommer-Aikins, Duell, & Hutter, 2005). Teachers in middle school and even high school students can counter this belief by explaining to students how to study and emphasize that learning is about effort and not luck (Schommer-Aikins, Duell, & Hutter). They also concluded that students' general beliefs about the nature of learning impacts their thinking about math specifically and contributes to their overall academic success (Schommer-Aikins, Duell, & Hutter). Teachers need to help students see how math relates to their daily lives through relevant math tasks on topics that interest students. Within complex instruction, teachers can design group worthy tasks that are applicable and interesting to students so that they will want to spend time and effort to complete the task. Teachers emphasizing effort over ability is also a feature of complex instruction that may broaden students' beliefs about the nature of learning and being a math learner.

It is also important for teachers to appropriately recommend students for learning disability testing and be aware that this labeling follows students and can forever change how their view their intelligence (Crosnoe, Riegle-Crumb, & Muller, 2007). Especially boys are affected by being labeled with learning disabilities (Crosnoe, Riegle-Crumb, & Muller). Also teachers need to make sure that girls are also being recommended for learning disability testing when appropriate, since many girls are not diagnosed with learning disabilities (Crosnoe, Riegle-Crumb, & Muller). Also, it is important to give extra support and encouragement to students who have failed classes or been diagnosed with a learning disability to challenge themselves by taking high-level math classes (Crosnoe, Riegle-Crumb, & Muller).

Teachers can use complex instruction methods to help students learn on a variety of different levels such as recalling facts, conceptual knowledge, and synthesizing information (Bianchini, Holthuis, & Nielsen, 1995). Teachers need to closely align content knowledge in my assessments with math skills and knowledge and not test reading ability inadvertently (Bianchini, Holthius, & Nielsen). Curriculum changes are only one facet of implementing complex instruction (Boaler & Staples, 2008; Bianchini, Holthuis, & Nielsen). Teachers need the support of other teachers and administrators to implement school-wide change (Boaler & Staples). Important aspects of complex instruction are giving students more responsibility over their education, implementing with high fidelity complex instruction strategies, being mindful of questioning techniques, keeping the cognitive demand high, giving students more conceptual based, longer problems to complete, and emphasizing effort over ability (Boaler & Staples). Teachers need to give students more full and explicit feedback to help them feel they are more capable of completing the tasks and have positive attitudes about the class (Steipek et al. 1998). This type of feedback helps students feel competent in certain areas and better able to make a plan as to how to address areas for improvement (Steipek et al.).

It can be very challenging implementing new teaching models, especially without the support of another teacher and a research team (Albert & Jones, 1997). Teachers need to have support from other teachers, administration, and ongoing professional development to help continue education about the theoretical ideas behind teaching methods and curriculum choices, specifically when implementing complex instruction (Lotan, 1989; Albert & Jones). Teachers need to have a strong conceptual understanding of teaching models and curriculum they are employing (Lotan). Teachers may find some aspects of complex instruction harder to implement regularly than others such as status treatments, group work roles, and group work rules/ norms (Perrenet & Terwel, 1996). These three aspects of complex instruction take very specific and ongoing attention by the teacher and take practice to make those aspects part of the daily classroom environment (Perrenet & Terwel). Teachers need more explicit and longer training sessions involving more conceptual knowledge of the underlying theory of complex instruction if it is to be implemented correctly and with high fidelity (Perrenet & Terwel). Some of the practices and strategies may not be consistent with all students' culture and teachers need to be very aware of the effect particular teaching practices have on students, for example assigning competence (Perrenet & Terwel). It is important that teachers have support while implementing new teaching techniques and strategies (Arbaugh & Brown, 2005). It could be beneficial for teachers going through reform together to use each other as resources to determine the level of cognitive demand of tasks used in the classroom regularly (Arbaugh & Brown). Similarly, teachers could observe each other to determine whether high cognitive demand tasks are actually eliciting high levels of conceptual thinking in the students (Arbaugh & Brown).

Suggestions for Further Research

There are still relatively few empirical studies on student academic and social achievement due to the complex instruction teaching model. There is a large body of literature about each area of complex instruction, but there needs to be more research conducted on the affects of teachers implementing this model on student achievement. There are a couple important studies that show impressive lessening of achievement gaps, changes in perceptions about mathematics, and even high rates of students taking advanced math courses (Boaler & Staples, 2008; Albert & Jones, 1997). For more administrators and teachers to take notice of complex instruction, educational researchers need to continue longitudinal studies to see if student achievement shown also reflects increases in standardized math test scores. If researchers can show that teachers using complex instruction have an increase in standardized test scores, more people within the field of education will take on the challenge of implementing this teaching model.

Also, the main study showing the many successes of implementing complex instruction at the high school level (Boaler & Staples, 2008) was done with an entire math department adhering to this teaching model, working collaboratively, changing and rewriting curriculum, detracking all math classes, and rearranging how math classes were scheduled and offered. Realistically, most high school math teachers will be trying out complex instruction without the support of the entire math department. It remains unclear if only a few teachers implementing complex instruction in a school have similar positive impacts on student academic and social achievement. Similarly, the researchers have yet to examine the results of teachers implementing some strategies of complex instruction regularly, but only giving students group worthy, complex tasks about once a month. Again, it is more realistic that teachers will be implementing complex instruction on their own or only with the support of a couple other teachers. They may not have the curriculum, resources, and/or energy to move from task to task with their students. It remains unanswered if teachers working hard to implement a truly "group worthy" task using complex instruction strategies about once a month will have the intended results on student achievement. What would the everyday instruction look like if teachers did not move from task to task using complex instruction tactics? Would students work in groups or alone if the tasks were not considered "group worthy"? Do using these strategies with standards based curriculum reproduce similar results to Railside (Boaler & Staples)? These are questions that are left unanswered by the research analyzed in this paper.

Conclusion

Chapter one explored the rationale of complex instruction as a topic, defined relevant terms within complex instruction, and highlighted different controversies about this classroom model. It explained that many students do not enjoy mathematics or find it in line with their identity. Complex instruction proposes that with modification to tradition teaching methods, teachers can start to create a classroom environment that is more in line with student interests as well as the domain of mathematics. Chapter two examined key areas of history that relate to the development of complex instruction. Throughout the history of math education, practices like tracking have made classrooms mirror society biases and in turn be places of inequity. Through status, multidimensional classroom, and identity development research, the advent of complex instruction was in response to making classroom environments a place where all students are valued, interact with each other in equitable ways, and can have deep learning experiences. Chapter three reviewed the research of complex instruction. The research in chapter three was organized into six themes: Complex Tasks, Conceptual Learning, and Cognitive Demand, Heterogeneous Group Work, Status and Competency Expectations, Students' Beliefs, Self-efficacy, and Identity, Student Social and Academic Achievement, and Implementation of Complex Instruction. These themes were used to answer this paper's guiding question: what is complex instruction and how can one adopt a complex instruction classroom model? The main findings of this chapter is that teachers should adhere with high fidelity to the components of complex instruction, students can have substantial social and academic growth within the complex instruction model, and that teachers need continual support when adopting new classroom models. Chapter four included a summary of the findings, based on the six categories from chapter three, implications for classroom practice, and suggestions for further research. Although it is challenging to adopt new classroom models such as complex instruction in a math classroom, students gain so much from working with each other in equitable and substantial ways. Students deserve learning opportunities in which they are valued, interested in the material, and learn the “why” behind the procedure.

REFERENCES

Aborn, M. (2006). An intelligent use for belief. *Education, 127(1),* 83-85.

Albert, L. R., & Jones, D. L. (1997). Implementing the science teaching standards through complex instruction: A case study of two teacher-researchers. *School Science and Mathematics, 97(6),* 283-291.

Arbaugh, F., & Brown, C. A. (2005). Analyzing mathematical tasks: A catalyst for change? *Journal of Mathematics Teacher Education, 8*, 499-536.

Bianchini, J., Holthuis, N., & Nielsen, K. (1995, April). *Cooperative learning in the untracked middle school science classroom: A student of student achievement.* Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA.

Blackwell, L. S., Trzeniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78(1),* 246-263.

Boaler, J., & Humphreys, C. (2005). *Connecting mathematical ideas: Middle school video cases to support teaching and learning.* Portsmouth, NH: Heinemann.

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside school. *Teachers College Record, 110(3),* 608-645.

Boaler, J., Wiliam, D., & Brown, M. (2000). Students' experiences of ability grouping: Disaffection, polarization, and the construction of failure. *British Educational Research Journal, 25(5)*, 631-648.

Boaler, J., Wiliam, D., & Zevenbergen, R. (2000, March). The construction of identity in secondary mathematics education. Paper presented at the International Mathematics Education and Society Conference. Portugal.

Cheng, R. W., Lam, S., & Chan, J. C. (2008). When high achievers and low achievers work in the same group: The roles of group heterogeneity and process in project-based learning. *British Journal of Educational Psychology, 78*, 205-221.

Christodoulou, J. A. (2009). Applying multiple intelligences. *School Administrator, 66(2),* 22.

Cohen, E. G. (1994). *Designing groupwork: Strategies for the heterogeneous classroom.* New York: Teachers College Press.

Cohen, E. G. (1984). Talking and working together: Status, interaction, and learning. In P. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *Instructional groups in the classroom: Organization and process* (pp. 171-187). Orlando, FL: Academic.

Cohen, E. G., & Anthony, B. (1982, March). *Expectation states theory and classroom learning*. Paper presented at the American Education Research Association. New York.

Cohen, E. G., & Deslonde, J. (1978). *Status equalization project (Grant No. NIE-400-76-0059).* Washington, DC: National Institute of Education.

Cohen, E. G., & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. *American Educational Research Journal, 32(1)*, 99-120.

Cohen, E. G., & Roper, S. S. (1972). Modification of interracial interaction disability: An application of status characteristics theory. *American Sociological Review, 37(6)*, 643-657.

Cohen, E. G., Lotan, R. A., & Leechor, C. (1989). Can classrooms learn? *Sociology of Education, 62(2),* 75-94.

Crosnoe, R., Riegle-Crumb, C., & Muller, C. (2007). Gender, self-perception, and academic problems in high school. *Social Problems, 54(1),* 118-138.

Dewey, J. (1997). *Experience and education*. New York: Simon & Schuster.

Dweck, C. S. (2007). The perils and promise of praise. *Educational Leadership, 65(2),* 34-39.

Gamon, A. (1987). The stratification of high school learning opportunities. *Sociology of Education, 60*, 135-155.

Gee, J. P. (2007). *What video games have to teach us about learning and literacy*. New York: Palgrave MacMillian.

Greene, M. (1995). *Releasing the imagination: Essays on education, the arts, and social change*. San Francisco: Jossey-Bass.

Henningsen, M., & Stein, M K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhabit high-level mathematical thinking and reasoning. *Journal for Research in Mathematical Education, 28(5),* 524-549.

Huerta, G. C. (2009). *Educational foundations: Diverse histories, diverse perspectives.* Boston: Houghton Mifflin.

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classroom. *The Elementary School Journal, 102(1)*, 59-80.

Khisty, L. L., & Chval, K. B. (2002). Pedagogic discourse and equity in mathematics:

When teachers' talk matters. *Mathematics Education Research Journal, 14(3),* 4-18.

Lotan, R. A. (2003). Group-worthy tasks. *Education Leadership, 60(6),* 72-75.

Lotan, R. (1989, March). *Complex instruction: Theory and classroom implementation.* Paper presented at the Annual Meeting of the American Education Research Association. San Francisco, CA.

Lotan, R. A., Cohen, E. G., & Holthuis, N. (1994, April). *Talking and working together: Conditions for learning in complex instruction.* Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA.

Lucas, S. R. (1999). *Tracking inequality: Stratification and mobility in American high schools*. New York: Teachers College Press.

Mitchell, S. N., Reilly, R., Bramwell, F. G., Solnosky, A., & Lilly, F. (2004). Friendship and choosing groupmates: Preferences for teacher-selected vs. student-selected groupings in high school science class. *Journal of Instructional Psychology, 31(1),* 20-32.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston ,VA: Author.

Perrenet, J., & Terwel, J. (1996, April). *Complex instruction in the Netherlands: A case study*. Paper presented at the Annual Meeting of the American Educational Research Association. New York, NY.

Pool, H., & Page, J. A. (1995). *Beyond tracking: Finding success in inclusive schools.* Bloomington, IN: Phi Delta Kapa Educational Foundation.

Resnick, L. B. E., & Zurawsky, C. E. (2006). Do the math: Cognitive demand makes a difference. *American Education Research Association, 4(2),* 1-4.

Rodriguez, J. L., Jones, E. B., Pang, V. O., & Park, C. D. (2004). Promoting academic achievement and identity development among diverse high school students. *High School Journal, 87(3),* 44-53.

Ross, J. A., & Cousins, B. (1994). Intentions to seek and give help, and behavior in cooperative learning groups. *Contemporary Educational Psychology, 19(4),* 476-482.

Rubenstein, R. N., & Bright, G. W. (2004). *Perspectives on the teaching of mathematics.* Reston, VA: The National Council of Teachers of Mathematics.

Sacks, O. (2000). *Seeing voices*. New York: Vintage Books.

Scarloss, B. (April, 2002). *Assessing complex academic performance at the group level.* Paper presentd at the Annual Meeting of the American Educational Research Association. New Orleans, LA.

Schommer-Aikins, M., Duell, O. K., & Hutter, R. (2005). Epistemological beliefs, mathematical problem-solving beliefs, and academic performance of middle school students. *The Elementary School Journal, 105(3),* 289-304.

Schullery, N. M., & Schullery, S. E. (2006). Are heterogeneous or homogeneous groups more beneficial to students? *Journal of Management Education, 30(4)*, 542-556.

Schunk, D. H. (1982). Effects of effort attributional feedback on children's perceived self-efficacy and achievement. *Journal of Educational Psychology, 74(4),* 548-556.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21(1)*, 43-69.

Spring, J. (2008). *The American school: From the Puritans to No Child Left Behind*. Boston: McGraw Hill.

Stipek, D., Salmon, J. M., Givvin, K. B., Kazemi, E., Saxe, G., & MacGyvers, V. L. (1998). The value (and convergence) of practices suggested by motivation research and promoted by mathematics education reformers. *Journal for Research in Mathematics Education, 29(4)*, 465-488.

Tammivaara, J. S. (1982). The effects of task structure on belief's about competence and participation in small groups. *Sociology of Education, 55(4),* 212-222.

Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: Association for Supervision and Curriculum Development.

VanDeWalle, J. A. (2007). *Elementary and middle school mathematics: Teaching developmentally*. Boston: Virginia Commonwealth University.

Vanfossen, B. E., Jones, J. P., & Spade, J. Z. (1987). Curriculum tracking and status maintenance. *Sociology of Education, 60*, 263-279.

Webb, N. M., Nemer, K. M., & Zuniga, S. (2002). Short circuits or superconductors? Effects of group composition on high-achieving students' science assessment performance. *American Educational Research Association, 39(4),* 943-989.

Wells, G. (2002). *Dialogic inquiry in education: Building on the legacy of Vygotsky*, In (Eds.) Glee & P. Smagorinsky. *Vygotskian perspectives on literacy research: Construction meaning through collaborative inquiry* (pp. 51-85). New York: Cambridge University Press.

Woods, T. A., Kurtz-Costes, B., & Rowley, S. J. (2004). The development of stereotypes about the rich and poor: Age, race, and family income differences in beliefs. *Journal of Youth and Adolescence, 34(5),* 437-445.