Accountability Groups to Enhance Language Learning

in a University Intensive English Program

by

Dianna Lippincott

A Dissertation in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

Approved April 2017 by the
Graduate Supervisory Committee:

Kathleen Puckett, Chair
Robert Kleinsasser
Steven Zuiker

ARIZONA STATE UNIVERSITY

May 2017

ABSTRACT

This mixed methods classroom research examined if accountability groups in the lower proficiency levels of a university intensive English program would improve students' language acquisition. Students were assigned partners for the study period with whom they completed assignments inside and outside of class, as well as set goals for use of language in their own context. Based in the ecological perspective and socio-cultural theory, activities reinforced social bonds, scaffolded the learning objectives in a communicative way, modeled the transfer of knowledge to the world outside the classroom, and allowed students to create new affordances in which to practice and use the language. Analysis of qualitative data from interviews, text messages, exit slips, and field notes, as well as quantitative data from student academic records, pre and post tests of curricular objectives, and pre and post attitudinal surveys, showed that students were developing a stronger sense of autonomy in their language learning. They viewed their peers and themselves as knowledgeable others, helping one another to learn vocabulary and structures in each student's zone of proximal development. Learner engagement in the treatment groups, as measured by classroom attendance, increased over a control group, as did overall grade averages in all courses. Students with no previous time in the program showed more improvement than those who had been in the program for at least one session prior. Students also showed increased fluency, as measured by the word count on a constructive task in the pre- and post-test of curricular objectives.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

APPENDIX Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Chapter 1 - Introduction

As a teacher of international students at a large university, I am always learning new things from my students. One class had a WhatsApp group, to which I, as the teacher, was not invited. One of the students had a question about an assignment, and I told him that a classmate had taken a picture of the whiteboard and put it on WhatsApp. He handed me his phone to find the picture that we were referencing. As I was flipping through, I came upon a photo of the previous night's homework. Someone had done the work, taken a picture, and shared the answers with the class. When I asked the student about it, he very sheepishly said that he had not used it. However, when I looked at the homework papers, every student in the class had the exact same right and wrong answers, including misspellings.

I did not confront anyone for the copied work as it was a small homework assignment. I wanted to gain more understanding of what the students were thinking. When I returned the papers, I asked the students, all Arab males between the ages of 18 and 23, why this had happened. They told me that it was important to help their "brothers." Then I questioned them about the value of honesty. They agreed that honesty was very important also, but when forced to choose between honesty and helping a colleague, they would choose to help the colleague.

This was a revelation to me. I knew some of their defining cultural characteristics are tight social connections and being very affable and hospitable. However, I did not realize how deeply these connections ran. Because of the complexity of culture, it is impossible to discuss groups of people without making generalities, but these generalities often provide insight; the accuracy of the generalities depend on the context. That being

said, the societal culture of the Arabian Peninsula is based on familial structure; the family is the cohesive center of life (Barakat, 1993; Hammad, Kysia, Rabah, Hassoun, & Connelly, 1999; Wunderle, 2006). However, family is defined as groups of up to 20 large extended families or qabila (Hammad et al., 1999). This projects to a strong collectivist culture, featuring the importance of the group over the individual, high value on the loyalty to group, group protection of members, and behavior that is guided by group norms (Barakat, 1993; Hammad, et al., 1999; Wunderle, 2006). Along with the focus on family, the history of Arabic language also shapes the culture: the strong oral culture of the language shows that it was not created to transmit facts, but to create a social, group experience (Zaharna, 1995). Because the Arabic culture is so social, friendships develop quickly and bonds as strong as family are created, carrying an obligation to do favors and give help, including helping a classmate give an answer and sharing work with a friend (Flaitz & Eckstein, 2003; Wunderle, 2006).

How can students retain the cultural values of collectivism in an American educational setting while helping each other in a productive way that truly improves language acquisition and not just gives the immediate, short-term answer? A closer look at the situational context of the university will help enlighten this discussion.

Context

America: The Land of Opportunity. One place this moniker is true is in the colleges and universities of the United States. Students come to the United States to earn college degrees and get a start on their future careers. For some, such as students from Brazil and China, their country's system of higher education is high quality, but does not have the capacity to accommodate all the students who wish to earn a degree. For others,

either their governments or the United States fund their scholarly endeavors. Still other students have made an individual decision to live and learn in a country with an innovative, vibrant economy and culture. According to the Institute of International Education (2016), a body of 1,200 member institutions of higher education, 1,043,839 international students attended U.S. colleges and universities, and of these, 133,335 students participated in intensive English programs.

To meet the needs of both students and universities, students must attain high-level mastery of the English language used in academics. The Test of English as a Foreign Language (TOEFL) and International English Language Test (IELTS) are the standard tools used to assess English language proficiency. Unfortunately, some prospective students are not proficient enough to receive the score designated by each university as "passing." It is generally accepted the most efficient way to learn a language is to study in the country where the language is spoken. Thus, Arizona State University, like many of its peer institutions, offers intensive academic English as a Second Language courses prior to entering a degree program for those students who do not have the passing score for admission to the university.

In this intensive English program (IEP), six levels of English language instruction are offered, from beginners who have not studied any English prior to coming to the United States, to students who are conversant in English and need only to enhance their English skills for a university classroom. A student can advance through a level in one session, or 7.5 weeks. In a typical week, a student attends a reading/writing course for nine hours, a listening/speaking class for six hours, a structure/grammar course for three hours, and an elective for three hours, for a total of 21 hours of classroom instruction.

Attendance is very important, and a student can fail a level for missing more than 600 minutes of class time in reading/writing, or 400 minutes in listening/speaking class, or 200 minutes in structure and elective classes. Classes are supplemented with 8 to 20 hours of homework per week and various activities, seminars, and clubs to enhance language learning.

In 2016, ASU's Global Launch served 1,437 individual students, with the average student attending for 23.5 weeks. About half of the students moved on to degree programs. The top two native languages were Arabic (38.5%) and some form of Chinese (32%). More than half of the students (56%) were males who were between the ages of 18-25. The participants/target audience of this research were those students who were in the second level, many of whom had recently arrived in the United States and were adapting to a new academic culture.

These students come from school systems that have teacher-centered classrooms, where foreign languages are taught by lecture and rote memorization (Flaitz & Eckstein, 2003). On the other hand, US language classrooms are oftentimes centered around student activity and discussion to acquire the language. The way that these students were learning English in their home country is not serving them well in what they need to accomplish, in this case, earn a degree in the United States. Because of their prior educational experiences, students are not aware of alternative methods for acquiring English in a fast and efficient way.

As discussed earlier, Arabs are a very tight-knit community with a collectivist culture. Likewise, the second largest population in the English program, Chinese, also come from a collectivist culture that values the importance of and obligation for the

family and community. Loyalty, duty, harmony, and trust are core to family relationships and friendships. Education is important as a demonstration of obedience and love of family and fulfills obligations for improvement of the family and community. Because of the seriousness of this obligation, failure of any form invokes feelings of shame; thus, face-saving is an important tenet. Students don't ask questions or speak in class for fear of losing face or causing the teacher to lose face. This leads to Chinese students forming close-knit social groups who complete assignments by compiling the knowledge of the group, rather than doing assignments individually (Alon & McIntyre, 2005; Fan, 200; Flaitz & Eckstein, 2003).

Both Arabs and Chinese, like many other ex-patriates, tend to associate with others from their own language group. These compatriots provide support for each other, but sometimes in ineffective ways, such as giving complete answers to homework, speaking to each other in their native tongues, and reinforcing some of the collective behaviors that are detrimental to academic achievement in an individualistic society.

Intervention

The Accountability Group Innovation scaffolds students' learning through evidence-based methods that increase language acquisition. Students worked with partners in the classroom on tasks that promote noticing of language, exchange of information, and exchanges that explain language constructs and vocabulary. This pair also functioned as a cooperative base group to give each other support, help, encouragement, and assistance needed to progress academically (Johnson & Johnson, 1999). Students set individual goals, which helped partners to hold each other

accountable in order to create new affordances in their own environment to make English their own.

Research Questions

This mixed methods action research study is designed to answer four questions:

1. To what extent do accountability groups help students succeed in an IEP in terms of curricular objectives, grades, and attendance? This question looks at the widely accepted indicators of success in an educational setting.
2. To what extent do accountability groups help students gain confidence in using English? Confidence in speaking and using a second language greatly increases the rate of acquisition. Students who feel more confident will use the language more, leading to further growth.
3. How do students in an accountability group create experiences where English language is used? Language can not be acquired only in a classroom. Students need to use the language in their own contexts if they want to become more proficient.
4. How do accountability groups influence the students' perceptions of friends and classmates as support systems in language learning? Because the majority of the students are from collectivist cultures, they have strong bonds to similar others. The Accountability Group Innovation is designed to leverage these social connections to enhance language learning.

Chapter 2 - Literature Review

Theoretical Frameworks

When studying the academic success of international students, a number of different theoretical perspectives could undergird the thinking of a researcher. Because of my interest in the relationships among students as a tool for student success and creating a context for learning that extends beyond the classroom, this research is grounded in the theoretical frameworks of an ecological perspective and socio-cultural psychology.

Ecological perspective. Traditionally, learning has been considered a very personal activity that takes place in the mind of the learner. However, proponents of ecological psychology have maintained the learner is in a context, a "person-in-situation" (Barab & Plucker, 2002, p. 169) and emphasize the "coupling" of the two (Barab, Squire & Barnett, 1999). Knowing is not in the head, but in the interaction of learner and context. This explains why people perform similar cognitive tasks differently in different settings. For example, although vendors in markets could complete complex mathematical problems, many were not able to reproduce the same calculations on a written test in a classroom (Barab & Plucker, 2002).

Barab and Roth (2006) delineate three parts of the environment or context: affordances, effectivity sets, and life-worlds. Affordances are the opportunities and possibilities of action that exist in an individual's context across time and space. Effectivity sets are the actual interactions in which the individual engages. Life-worlds are the worlds individuals perceive, "the environment from the perspective of an individual" (p. 3). Life-worlds are not seen as static, but dynamic across time.

The focus of the ecological perspective for the teacher, and for this research, is how an instructor can facilitate "functional transactions through which individuals increase their potential" (Barab & Plucker, 2002, p. 175). In other words, an educator must help learners to participate in and create new affordance networks.

When considering how these networks help and hinder second language learners, it is important to realize that "complex social dynamics and sociocultural and cultural-historical arrangements" cause students to enlist affordances or to reject them (Barab & Roth, 2006, p. 8). Learners need to have the capacity and interest to create new possibilities, even reconstructing social, cultural, and historical relations within their environment, to expand their life-worlds. As Barab and Roth (2006) state, "An important aspect of learning is to support the learner's adoption of a new way of knowing and caring about the world" (p. 6). This is also the "core challenge of education … develop(ing) curricular contexts that extend themselves meaningfully into the personal life-worlds of individuals" (Barab & Roth, 2006, p. 7). Because each individual has her own life-world, her own perception of the world, the teacher needs to address a wide variety of contexts, but ultimately, the learner must desire to engage in situations where the knowledge has value. "Eventually the teacher needs to stop governing the … relations … and the learner needs to create her own contexts-of-use" (Barab & Roth, 2006, p. 11). This is the challenge of this action research—to guide students in the creation of their own contexts where English language will be of use.

Sociocultural theory. Sociocultural theory (SCT), as a praxis-based view, also challenges teachers and learners to create conditions where skills are developed and the environment is mediated by cultural artifacts, activities, and concepts. SCT, as defined

by Lantolf, Thorne, and Poehner (2014), is built on four main concepts: mediation, regulation, internalization, and the zone of proximal development (ZPD). Because SCT is a broad term that is seen in a wide diversity of ways, this definition is the model upon which this action research has been built.

The mind has been socially formed, which means the relationship of mind to world is mediated by artifacts generated by human cultural activity (Edwards, 2001; Lantolf et al., 2014). As Lantolf et al. explain, "Humans do not act directly on the world—rather their cognitive and material activities are mediated by symbolic artifacts as well as by material artifacts and technologies" (p. 221). Symbolic artifacts have consisted of cultural tools such as language, literacy, logic; material artifacts and technologies are tools, both traditional and digital, such as shovels, wheels, time pieces, or even the Internet. Through the use of these tools, higher order thinking is organized and expanded.

Using language as a mediator is a challenge for people who speak a first primary language (L1) and are learning a second language (L2). According to Lantolf et al. (2014), L2 users have a difficult time using the new language to mediate their cognitive activity, even if they are proficient. Learners rely on their first language to mediate learning of the second language. This means teachers and learners need to foster the development of the second language by discussing its features, grammar, and syntax in either the first or second language.

The second concept of SCT has been recognized as regulation or agency. Lantolf et al. (2014) describe three types of regulation: object, in which objects in the environment allow cognition and activity; other, in which mediation is controlled by

others giving feedback; and self, in which the learner mediates the environment for herself. The agency or control has been characterized as being embedded in the "interwoven and shifting contexts" (Edwards, 2001, p. 172). Learner development involves moving from other-control to self-control, or in the words of Lantolf et al. (2014), "development can be described as the process of gaining greater voluntary control over one's capacity to think and act" (p. 209). However, it should be noted the learner moves back and forth across these three regulatory types at different times, and that the transfer of regulation occurs over a period of minutes, hours, months, or years. What is important is that the student sees herself as a learner and as involved in her acting and interacting with the world (Edwards, 2001).

The third important concept of SCT is internalization. Cultural artifacts first appear on the social level, and then they become part of the psychological level—they are internalized after initially being external to individuals. For example, "When we learn to communicate socially, we appropriate the patterns and meanings of this social speech and also utilize it inwardly to mediate our mental activity" (Lantolf et al., 2014, p. 210). Consistent with SCT, this process occurs in the Zone of Proximal Development (ZPD). Tappan (1998) explains the process: "As a result of collaboration that takes place in the ZPD, externally oriented and socially constituted learning processes between persons become internally oriented" (p. 26).

The ZPD has been viewed as the distance or space between what the learner already knows and what the learner is ready to learn, as measured by what she can do with the help of a more knowledgeable other. According to Lantolf et al. (2014), the control of learning and performance first resides in others and then progresses to self-

regulation. Edwards (2001) states that the teacher helps the learner engage with possible methods of learning, as well as the concepts of the subject area. He is concerned with "orchestration of time and space, self and others, learners and knowledge, and affect and cognition" (p. 179). In the ZPD, the mental constructs of both the learner and the more knowledgeable other change through their conversation and dialogue (Tappan, 1998), a fact that validates participation in communities of practice.

The challenge of SCT to this action research is to create zones where students have freedom to inhabit new ways of learning, doing, and relating. Such an approach allows students to capitalize on what Tappan (1998) calls the students' "growing edge," acting to change their future (p. 30). This occurs through interaction with their environment, mediated by social interactions.

Conceptual Framework

The Accountability Group Innovation is designed to organize experiences that allow students to mediate their environment in English through social interactions while allowing them the freedom to create their own contexts that are meaningful to them. A logical type of social interaction that supports the learner's negotiation of the ZPD is communicative partner activities. A review of the literature shows the rationale for using partner work and what types of tasks are most effective for helping students expand their growing edge and create their own opportunities for language learning. Setting goals make students be accountable to themselves for their learning.

Rationale for partner work. Five reasons for using partner work are highlighted in the literature: (1) more negotiation of meaning occurs; (2) students model more

advanced language for each other; (3) scaffolding for better learning takes place; (4) collaborative activities are linked to learning; and (5) student motivation is improved.

Negotiation of meaning. When two people speak to each other, sometimes breakdowns occur in that communication, perhaps because of a lack of shared background knowledge, an unknown or mispronounced word, unclear speech, or any number of other factors. The parties involved in that communication then try to re-establish the message through questioning, replying, restating, modifying, and restructuring. The goal is to achieve comprehensibility; "negotiation can serve as a means of working through perceived or actual gaps in communication" (Pica, 1994, p. 499). In negotiation, the original message is repeated, reduced, or increased in length, by substituting or eliminating words and modifying the message other ways. This act of negotiation assists second language learners with comprehension and with L2 form, which gives it a powerful role in language learning (Pica, 1994).

Pica (1994) studied a corpus of exchanges between native speakers and non-native speakers contained in published studies. She found that negotiation drew attention to comprehension, including clarification of lexical items (vocabulary) and also to structure, such as when a phrase was taken out of context and repeated by itself or when a grammatical form was negotiated. She also reviewed two experimental studies that show that negotiation can have positive impacts on learning.

Despite its benefits, negotiation is rarely valued or taught in the language classroom. Furthermore, "studies by a number of researchers (e.g. Pica and Doughty, 1985; Varonis and Gass, 1985) have shown that compared to teacher-fronted classes or native speaker (NS)- non-native speaker (NNS) pairs, learners in groups of in NNS-NNS

pairs engage in more modified interactions" (Storch, 2001). Pica (1994) posits that this dynamic is related to the power differential between teacher and student and to traditions of how language is taught and how the classroom is structured.

In fact, Klingner and Vaughn (2000) found that groups of fifth grade students who were working in Collaborative Strategic Reading groups used negotiating strategies about half of the time. When confronted with vocabulary and content that they did not understand, they translated for each other, defined words, shared ideas about how to find the meaning through context, elaborated and rephrased ideas, and gave positive or negative feedback. Rather than being controlled by the teacher, each group developed its own style, depending on the personalities and skills of the members.

This type of collaborative group work exemplifies socio-cultural theory, one of the theoretical frameworks for this action research project. In this research, working in accountability groups will give students more opportunities to negotiate meaning with their peers.

Modeling. One way that people learn is by observing others perform some behavior, watch the results and consequences, and then imitate the behavior at a later time. According to Bandura's Social Cognitive Theory, we process the action mentally with four cognitive processes: (1) attention, noticing the behavior; (2) retention, remembering the behavior; (3) reproduction, performing the behavior if we have the ability; and (4) motivation, desiring to perform the behavior (McLeod, 2016).

Pica (1991) found that study participants who watched their classmates complete a task that involved negotiation of instructions were able to comprehend those directions

as well as the classmate who actually completed the negotiating task. Children attend to and imitate people that they perceive are similar to them (McLeod, 2016).

This study proposed that students in accountability groups would model for each other. Because the accountability group consisted of students with many of the same demographic features (age, gender, ethnicity) and of approximately the same ability level, students would attend more closely to the action of their peer than they would of the teacher. Language structures used would more likely be in their Zone of Proximal Development, which means students would remember and be able to imitate the language.

Scaffolding. According to Johnson and Johnson (1999), scaffolding is providing frameworks that allow learners to understand content. Coined by Jerome Bruner, scaffolding moves the student into the Zone of Proximal Development (ZPD) so that learning can take place. Modeling, discussed above, is a form of scaffolding, but there are many others: accessing prior knowledge, talking and writing to process new ideas, pre-learning vocabulary, using visuals, filling out graphic organizers, asking and answering questions, etc.

The Zone of Proximal Development (ZPD) exists between the level at which a student can work independently to the level of language produced when working collaboratively with a peer. Because each student's Zone of Proximal Development is different, it is difficult for a classroom teacher to teach in every student's ZPD. In Ohta's (2001) data analysis, she infers that learners provide and receive developmentally appropriate assistance, or scaffolding, to/from their classmates.

Storch (2001) gives examples of scaffolding exchanges where students extend each other's language and ideas, repeating and confirming what has been said, and maintain their classmates' attention to the task.

When students listened to a mini-lecture, took notes, and then shared their notes, they performed substantially better on a post-listening test than the control group who only studied their own notes. While sharing their notes, students provided scaffolding by peer interaction: questioning of language use, repairing grammatical forms, confirming and code-switching (Garcia & Asención, 2001).

In accountability groups, students will work with partners in their Zones of Proximal Development to optimize language learning.

Collaborative activities. Accountability groups give students the opportunity to engage in collaborative activities that have been linked to learning. As cited by Swain & Lapkin (1998), LaPierre (1994) demonstrated this connection when he had student dyads reconstruct a story in French, their L2, and in so doing, talk about the language they were producing. Specially designed post-tests showed that students remembered what they had collaborated on the previous week. If students had discussed the forms correctly, they got the answers correct on the post-test; if they discussed them incorrectly, they were incorrect on the post-test. Beach (1974) affirms this link between increased student involvement and increased mastery: comparing peer study groups with no tutor to tutor-led groups, the groups without a tutor scored higher on a final achievement test.

Swain and Lapkin (1998) observed two students in a collaborative dialogue in which one student was unsure of the French word for *alarm clock*. After questioning and

receiving responses from his partner several times during the task, he slowly began to move from uncertainty about the word to confidently using the word.

Collaborative pairs also co-construct, extend, and scaffold language, allowing knowledge to be organized in a way that is accessible (De Guerrero & Villamil, 2000; Slavin, 1995; Storch, 2001; Swain & Lapkin, 1998). In addition to increased mastery and consolidation, reorganization, extension, and scaffolding of language, collaborative work creates a classroom that is more conducive to deeper learning (Littlejohn, 1983). Accountability groups will bring increased and deeper learning to students.

Motivation. The best learning occurs when students are motivated. Four dimensions of a classroom that motivate students are proposed by Guilloteaux and Dörnyei (2008), each shown to be met through pair work:

Pleasant and supportive atmosphere with a cohesive learner group. Through both observations and participant's answers to a questionnaire, Littlejohn (1983) found that in student-centered collaborative classrooms, "participants felt free to speak, to make mistakes, and to contribute their own experiences, all of which gave them a feeling of being supported in their difficulties with learning" (p. 598). They no longer felt they were competing against other students, but that they were cooperating.

Use of strategies to increase learners' expectation of success. The classroom is not a place where students are "rehearsing for a performance at some later time and place" (Littlejohn, 1983, p. 597), but a place where real communication is taking place and where students are actively engaged in social and academic discussions with their peers.

Providing learners with successful experiences and helping them to maintain a positive social image. Students are not constrained to a set method to achieve a task, allowing for learning to be personalized. There "may in fact be as many approaches to language learning as there are language learners" (Littlejohn, 1983, p. 597).

Effective and encouraging feedback. In student-directed groups, such as accountability groups, students develop self-efficacy as they decide what to learn, how to learn it, and how to find and correct their mistakes (Littlejohn, 1983).

Summary example. He and Ellis (1999) conducted a mixed-methods study that supports the five rationales for using pair work in the Accountability Group Innovation: students negotiated meaning and modeled language for each other, they scaffolded the task in ways that the teacher did not, the collaborative approach suggested more learning took place, and students were motivated to learn by having a personalized, realistic task and by lowering their affective filters. Three treatment groups were given vocabulary to learn. The first two groups were given oral definitions, but only the second group was allowed to interact with the speaker and ask questions to aid comprehension. The third group was not given definitions, but had to co-construct the meanings with a partner, using scaffolding techniques: interacting, negotiating meaning, modifying their spoken expressions. Afterwards, students were given post-tests covering comprehension, recognition, and production of the vocabulary.

The quantitative data in this study shows the third group, who created their own directions, outperformed the other two groups in comprehension, vocabulary recognition, and vocabulary production. He and Ellis (1999) also included qualitative data in which they analyzed the transactions between the teacher and students in the first two groups

and the peer partners in the third group. The lexical framework for the first two groups, provided by the teacher, used low-frequency words, such as *rock* to describe the action of a rocker, while the student-produced definition used more high-frequency words, such as *move* to describe the action of the rocker. The scaffold of using high-frequency words put the information in the Zone of Proximal Development of the student so that information could be retained. Students were motivated by their peers: they "treated the task as a collaborative problem-solving activity the goal of which was to help each other to arrive at a successful outcome for the task" (He and Ellis, 1999, p. 130). The students were given the opportunity to talk in ways with which they were familiar and comfortable, lowering their affective filters.

This study provides a summary example of the rationale found in the literature for the Accountability Group Innovation. Moreover, the mixed methods approach helps justify the choice of mixed methods for the current study of accountability groups. It provides a model of how utilizing both quantitative and qualitative data strengthen research findings.

Types of tasks. In addition to providing a justification for the Accountability Group Innovation, the literature provides examples of the most effective types of tasks to present to students: those that involve noticing language, those that obligate an exchange of information, and those that elicit language related exchanges.

Noticing. Watanabe and Swain (2007) studied Japanese student dyads who wrote a collaborative essay. Subsequently, a native English speaker rewrote the essays with corrections and revisions to supply the students with a sample of authentic, native English. Student dyads then compared their collaborative essay and the native-English

essay, noticing errors and reformulations. The researchers analyzed recordings of the discussion that took place when the students were writing collaboratively and when they were noticing the differences in the reformulation and found more focus on language-related talk during the noticing activity.

Exchange of information. "Foster found that, overall, dyads with an obligation to exchange information were most likely to talk and to negotiate meaning" (Swain & Lapkin, 1998). Therefore, the tasks that required an exchange of information led to more talking and more negotiation of meaning. Explaining material to someone else has benefits for both the person explaining the information and for the person receiving the information (Slavin, 1995).

Language-related exchanges. Littlejohn (1983) had student dyads record a role-play as they created it. They then listened to the recording and transcribed it, correcting errors as they did so. This activity required students to transfer spoken words to written and incorrect grammar to correct form. To do so, students had to notice the errors and exchange information. This led to many language-related exchanges, which are shown to improve learning.

Summary example. Swain, Brooks, and Tocalli-Beller (2002) reviewed nineteen studies, categorizing those studies by language strand—writing, speaking, reading, and listening—and making suggestions about what types of activities meet the criteria to create the most effective language learning opportunities for student pairs. They suggest collaborative writing and peer revision, in which the peers interact about the written product. In the strand of speaking, they suggest a poster carousel in which students give poster presentations while other students rotate from poster to poster. This allows a

presenter to repeat the presentation several times, improving with each repetition and with questions posed by their audience. Recording role plays is another way to bring students' attention to their language use. In the receptive strands of listening and reading, Swain et. al. (2002) suggest peer to peer dialogue after listening or reading in order to reconstruct the source materials, to answer questions, and to fill in cloze exercises.

Goal setting. The Accountability Group Innovation will harness the power of collaborative pair work, which has been shown to increase language learning through negotiation, modeling, scaffolding, and motivating students. However, accountability groups differ from mere partner work in that students will be consciously setting goals for themselves. This allows them to internalize and take control of their own learning process.

Goals can be classified as two types: learning and performance. Learning goals focus on acquiring skills and knowledge to perform a task while performance goals focus on the completion of the task. Learning goals are better than performance goals (Locke & Latham, 2002, 2006). Setting a learning goal facilitates metacognition—thinking about planning, monitoring, and evaluating progress; in contrast, setting a performance goal tends to lead to the avoidance of tasks in which others may judge the goal-setter unfavorably (Locke & Latham, 2006). In addition, goals may be conscious or subconscious. Over a 25-year study, Locke and Latham (2002) found that a subconscious goal, the intrinsic need for achievement, was not as effective as a conscious goal, the goal of a specific number of promotions. Schippers, Scheepers, and Peterson (2015) found that the academic achievement gap for 700 students was erased in two years by having students write their goals, possibly because goals were made conscious.

Along with setting conscious learning goals, the literature shows other key factors that influence the attainment of goals: plan for implementation, feedback, and sharing of the goal. Creating an if-then plan of where, when, and how the goal setter will take actions to realize the goal has been shown in 94 independent tests to have a positive effect on goal attainment. Thinking through how to implement the goal increased access to opportunities to take action and automated the goal-directed responses (Gollwitzer & Sheeran, 2006).

Receiving feedback on progress towards a goal is necessary to measure movement toward goal completion. Feedback increases effort towards the goal and promotes trying new strategies if the current strategy is not effectively progressing the goal setter towards their desired result (Locke & Latham, 2002).

The final factor that influences goal attainment is especially salient for the Accountability Group Innovation: sharing the goal. In two separate studies, Shteynberg and Galinsky (2011) found that only when both (a) sharing a goal (b) with similar others does the pursuit of the goal intensify. In sharing goals, task relevant information is shared (Locke & Latham, 2006) which makes the goal pursuit more efficient and encourages metacognition around how to progress towards the goal. Locke, in an unpublished study, found that partners who share information perform better than partners that share less information. Knowing that others like you have the same goal leads to a desire to behave like the others, thus increasing goal pursuit behaviors.

Collaborative base groups. Moving to a new country, leaving family and friends at home, and not having a ready support system in the United States can leave students feeling isolated, lonely, and depressed. This is the same condition that Johnson

and Johnson (1999) describe for students in K-12 who have difficulties in school. While feeling isolated, students often see their academic achievements as meaningless because they have no one with whom to celebrate them. Johnson and Johnson (1999) propose a collaborative base group to develop caring and committed relationships that "provide the social support to improve attendance, personalize the school experience, increase achievement, and improve the quality of life within the classroom" (p. 64). Group members provide support, encouragement, and assistance in completing assignments, while holding each other responsible for striving to learn. Students contribute to the well-being of others and the common good, as well as achieve personal milestones. These groups create social norms among peers that promote high achievement, which should lead to increased achievement (Slavin, 1995). In the accountability group innovation, forming this type of collaborative group will fulfill students' desire to continue to contribute to the group well-being, a prominent feature of collectivist cultures.

Implications

Taken together, the theoretical and conceptual frameworks found in the literature suggest three implications for this action research study. Because language is learned in social situations, the first implication is that any intervention should take this into consideration. The teacher needs to help students create opportunities (affordances) and contexts to use English. The related literature shows that partner work is the most effective way to do this as its emphasis on negotiation, modeling, scaffolding, collaboration, and motivation leads to more language learning. The second implication is that any planned tasks should encourage action that encompasses this rationale for pair

work—tasks that will foster noticing and discussing the form of language and exchanging information. The third implication is that consciously setting goals and sharing them with similar others will lead students toward success in their language learning. By forming meaningful relationships with peers, students will overcome feelings of isolation and become engaged in helping others, as well as themselves, to meet these goals.

Chapter 3 - Methods

This chapter provides information about (a) the participants, (b) the innovation or intervention, (c) instruments and data collection, (d) analysis of data, and (e) threats to validity.

Participants

Treatment group. The participants of this research were students in the Basic 2 level of the intensive English program at Arizona State University, previously described in Chapter 1. This is the second level of the intensive English program and is approximately a Common European Framework (CEFR) level A2. Students were placed into this level in one of four ways, which can effect the student's ability level, motivation, and experience with American classroom culture:

1. Promotion from Basic 1 level (6 students in summer treatment group; 2 students in fall treatment group & 1 student in control group). Students who have completed the Basic 1 level with a grade of 70% or higher are promoted into the Basic 2 level. These students have been in the United States for at least two months and have experience with American classroom culture and intensive language learning.
2. Demotion from Intermediate 1 level (no students in treatment or control groups). Any student who has not achieved a grade of 70% or better for two sessions in any of their classes (have failed the level twice) are demoted to Basic 2. These students have been exposed to more advanced language and structures, but have not been successful with language acquisition or with the

expectations of an American academic environment, such as doing homework or attending class.

3. Repeating Basic 2 level (5 students in summer treatment group; no students in fall treatment group & control group). A student who was in Basic 2 during the previous session and did not receive a minimum of 70% in all classes will be repeating the level. These students have been in the United States for at least two months and have experience with American classroom culture and intensive language learning; however, they have not been successful.
4. Placement test (4 students in summer treatment group; 10 students in fall treatment group & 14 students in control group). Most students are in their first session at the intensive English program and are new to the American classroom culture and expectations. After registering with the program, they were given a placement test that indicates their language proficiency as pertains to the intensive English program. However, every learner is unique and has different strengths and weaknesses in the four strands of language. Understanding how students are tested on language proficiency helps clarify the differences among learners in the same class. The test has three sections, with the most weight being given to the writing sample.
    a. CAMLA's English Placement Test, a standardized multiple-choice exam created by the University of Michigan. The test is designed to assess general English proficiency and place students into various levels. It contains 80 test items that cover listening, grammar, vocabulary, and reading comprehension. Twenty-five listening items

are played over the testing room's loudspeakers and students are given a set number of seconds to answer the question from their test booklet on a grid sheet. Two types of listening prompts are used: listening to a short sentence or phrase and choosing the correct response or listening to a short conversation and answering a question about the content of that conversation. Then the remaining parts of the test are self-paced for the duration of the testing period. Twenty grammar questions test grammar in context. Next, twenty cloze-type vocabulary questions are presented with four answer choices given. The last section is fifteen reading comprehension items of two types. The first five questions contain one sentence and a comprehension question. In the second part of the reading comprehension section, the student reads two passages and answers ten questions based on those passages, addressing a range of reading skills. A Basic 2 student would score between 38% and 50% on this test.

b. Speaking test, adapted from a standardized ESL speaking test sometimes known as "The John Test," named after the character in the picture prompts. This one-on-one interview test has three parts. After an initial greeting with the tester, the student is asked to answer a series of 24 questions about eight pictures. The pictures depict the everyday life of either John or Jane, depending on the version of the test. Each question is testing specific grammar, syntax, or usage, and the pictures give a familiar context. Designed to elicit spontaneous

oral production, the second section of the test is an independent speaking prompt that has been created by faculty in the program. The response is rated by the interviewer according to a rubric that weights structure, vocabulary, and delivery equally. Delivery is sub-divided into fluency and pronunciation categories. In the third and final section of the test, the interviewer elicits the construction of questions with specific grammar embedded, such as clauses, verb tenses, and voice.

c. Writing sample. Students are given twenty minutes to respond to one prompt that is designed by faculty to elicit an informational causal or persuasive essay (e.g., What do you think of x? Give reasons.). All prompts are scored with attention to form, vocabulary, organization, grammar, punctuation, and logical development. Two trained raters determine levels by holistic scoring, guided by general parameters. If there is a discrepancy between the two raters, a third rater is used. Inter-rater reliability is 73%.

Students in Basic 2 know some very basic English, but are expanding their vocabularies and becoming more comfortable and confident with their language usage. They are learning to use various tenses—present, past and future—in questions, simple personal narratives, descriptions of people and places, and spatial directions.

Participants in this study were chosen through a convenience sample: they were the students that have been randomly assigned to a particular section of the course. In the treatment group, all students in a particular section of the course participated in all

activities; however, as approved by IRB, only those who signed the consent form had data collected and analyzed.

The Summer 2017 treatment group, Treatment Group A, was composed of 9 males (56%) and 7 females (44%), while the Fall 2017 treatment group, Treatment Group B, was 10 males (71%) and 4 females (29%). In Treatment Group A, there were 8 Arabic speakers and 4 Chinese speakers, with 4 others being native Japanese speakers. In Treatment Group B, 3 Arabic speakers, 5 Chinese speakers, and 4 Japanese speakers predominated, with two speakers of other languages.

Control group. Because the materials and learning outcomes for the course are standardized by the department, the outcomes of students in this innovation could be compared to student outcomes in a control group. The control group was a different section of the same course, taught by a different teacher. They were also students in the basic 2 level, having been placed by one of the four methods as shown above. They were also randomly assigned into the course. The control group had 6 females and 11 males, or 65% male, a slightly lower percentage than the program as a whole (66% male). Also, the control group contained 2 Arabic speakers, 9 Chinese speakers, and 6 native speakers of other languages.

Innovation/Intervention

This innovation was to facilitate accountability groups, in order to see how these groups help students to succeed in terms of curricular objectives, grades, and attendance; to gain confidence with using English; to create their own experiences where English language is used; and to use friends and classmates as support systems in language

learning. The Accountability Group Innovation methodology had the following components: (1) Partners and groups, (2) Student-created goals, and (3) Methodology.

Partners and groups. The researcher matched students with partners, taking into consideration what other classes they took together, their ethnicity, and their gender. These pairs of similar students were then matched with another dyad, forming groups of four. This was designed to allow the project to continue even if someone dropped out or did not fully engage. The partners and groups were assigned to work together on class assignments and homework.

Student-created goals. An essential part of the Accountability Group Innovation was the setting of goals by students in treatment groups. Because the intended outcome of the Accountability Group Innovation was enhancement of language learning that would lead to better grades, attendance, and confidence in using English in a variety of contexts, students set academic or social goals. Academic goals could have included completion of assignments, use of structures in other contexts, vocabulary building with flash cards, or writing in a journal. Social goals could have been ways to adapt to American social norms, especially related to communicating effectively in a variety of situations, such as when conferencing with teachers or when asking for help. Issues of traversing independence and American ways of doing things—doing laundry, complaining to the landlord, taking public transportation, buying a car, etc.—could have been addressed as well. Support from peers can help students to feel less isolated in a new culture.

During the first session, students set an initial weekly goal, considering the types of goals listed above and the lessons of the week. The teacher scaffolded the goal-setting

for students individually with prompts, suggestions, and questions. Because of the low proficiency level of the students and their lack of experience in goal-setting, this part of the innovation had to be very clearly structured. Students were asked to write down where and with whom they would use English and exactly when this would occur. After experience with the first treatment group, Treatment Group B was even more structured. They had to say where, with whom, and when they would use the structure taught in the class. Students shared their goals with their partner, and they decided at least one way that the partner could help them achieve the goals. Then they shared with their group. Writing down their goals and how their partners would help them was a part of the exit slip that they submitted to the teacher (see Appendix C).

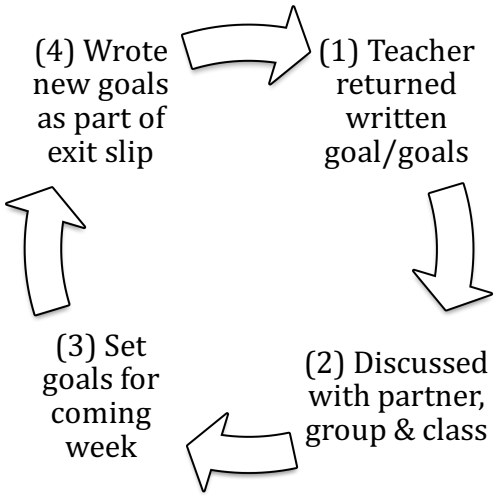


Figure 1. Goal setting cycle used weeks two through eight.

Each week thereafter, the teacher returned the written goal statements, which had no commentary and were ungraded. Students reflected on whether they felt that they had

met their goal or not. With their partner, they shared their successes and challenges, giving real-life examples from the week. Partners were encouraged to share their successes and how their partners helped them with the entire class. The class brainstormed other ways to help each other effectively. Then they set goals for the following week, using the method described above, discussed the goals with their partner, and planned how they could assist each other to meet the goals. This was, once again, written down and given to the teacher on the exit slip, along with individual reflections on how they achieved their goals for the week, how their partner was able to help them, and what they envisioned for the new week.

Methodology. The course, Basic 2 Communications, uses a standardized curriculum based on the textbook, *American English File Book 1, 2nd edition* (Latham-Koenig, Oxenden, & Seligson, 2013). Teachers must teach certain pages during each class period, but have some freedom in the methodology. The control group was taught with methodology chosen by the teacher of their class. The textbook recommends role-play, listening and answering questions, and occasional conversations with a partner. In hopes of increasing student learning, the Accountability Group Innovation methodology consisted of primarily partner work, self-reflection in the form of goal setting and monitoring, and collaborative peer support.

Accountability group methodology. The methodology used in the Accountability Group Innovation centered on partner and group work that enhances learning and on setting and monitoring accomplishment of goals. Partner and group work was designed into activities that aimed to (1) reinforce each accountability group's social bonds, (2) scaffold the learning objectives in a more communicative way, (3) allow members of

accountability groups to be the knowledgeable other within their groups, and (4) model how to transfer this knowledge to the "real" world outside the classroom.

Activities for the students in the Accountability Groups were focused on the task types shown in the literature to be effective for learning. As shown in Figure 2, in-class activities, homework, and outside assignments were designed to give students a chance to work together, utilizing the types of tasks recommended in the literature: tasks that promote noticing language, tasks that involve an exchange of information, and tasks that promote language-related exchanges. Activities for all units in the Accountability Group Innovation are provided in Appendix B.

Unit	In Class Activities (Type of Task)	Homework Activities (Type of Task)
Unit 1: Arriving/Checking in to a hotel	Scaffolded role play; with partner, write a dialogue to share with group (noticing language; language related exchanges)	Create a tip sheet for newcomers to ASU from your country, listing ten pointers about cultural things they need to know. (exchange of information)
Unit 3: In a Clothes Store	Play a game with partner about changes in clothing. (exchange of information)	With your partner, plan what you will wear to class. Email the teacher. (language related exchanges)
Unit 4: Getting Lost	Using a map, explain how to arrive at different locations. (exchange of information)	With partner, write how to go from the classroom to a location on campus. (noticing language)

Figure 2. Sample partner/group communicative activities, enhanced from course standardized curriculum.

Control group methodology. The control group was another section of the same course. They used the methodology that their teacher chose, which was usually the

methods recommended in the departmental course guide and the textbook. These were usually vocabulary review, listening with comprehension and recognition questions, repeating conversations, and role-plays. The students worked in pairs, but switched partners often. The pairwork usually consisted of practicing a structure presented in class in a very controlled fashion, such as reading a conversation or answering set discussion questions. To increase trustworthiness of the study, the researcher met with the teacher of the control group regularly to verify how lessons were taught, as well as reviewing the teacher's lesson plans or course calendar.

Institutional Review Board Approval

This study was conducted with the approval of the Arizona State University Institutional Review Board (IRB). They approved the innovation methods, as well as the data collection and analysis methods. In order to protect participants from any perceived coercion to take part in the study, the consent forms were collected by a trusted outsider who was not a part of the study. This person retained all consent forms until the end of the course, so that the teacher/researcher did not know which students had consented and which had not. This meant that during the course of the study, all students in the treatment group participated in data collection activities, as they were a part of the requirements of the course. However, after the implementation of the innovation, the data collected from those who had not consented to be a part of the study were destroyed. The data collected from the consenting participants were analyzed.

Instruments and Data Collection

This action research utilizes a mixed method approach with a nested concurrent quantitative and qualitative study design. Quantitative data allowed comparison between

the treatment groups participating in the Accountability Group Innovation and the control group that did not participate. Qualitative data are, by nature, constructivist, exploring different perspectives on an issue and reflecting the experience of individuals. They can generate in-depth explanations of what is occurring and why it is happening in a particular environment (Ivankova, 2015). Seven instruments will be used for data collection: three are quantitative and four are qualitative. Figure 3 highlights the instruments, who they will be used on and when, the type of data (quantitative or qualitative), and the timeline. Further explanation of the tools follow, and examples of instruments are in Appendix C.

Instrument	Who	Type	Timeline
Pre and post test (skills)	Control & treatment groups	Quantitative	Pre-intervention 5/16 Post-intervention 6/16
Pre and post survey of confidence level & use of friends as academic support	Control & treatment groups	Quantitative	Pre-intervention 5/16 Post-intervention 6/16
Semi-structured interviews	Treatment group	Qualitative	Throughout May and June 2016
Exit slips	All class members completed; data collected from consenting participants in treatment group	Qualitative	Throughout May and June 2016
Text messages	All class members completed; data collected from consenting participants in treatment group	Qualitative	Throughout May and June 2016
Field notes	Researcher with treatment group	Qualitative	Throughout May and June 2016
Academic records (grades & attendance)	Control & treatment groups	Quantitative	September 2016

Figure 3. Data collection instruments for Accountability Group Innovation (AGI).

Pre-treatment and post-treatment skills tests. Skills tests were created by the researcher. The test had two parts: (1) filling in a common conversation between a server and customer in a restaurant and (2) giving directions based on a map. The conversation

between server and customer had four utterances made by the server and four by the customer, but the customer sections were redacted so that students had to fill them in. The completions were devised so that the student had to create two questions and two declarative sentences. A variety of natural language responses were possible. The second section contained a map of three-block section of a typical town. Students were asked to write directions to a friend explaining how to go from one location to another. They could write as much or as little as needed to complete the task; however, students with greater proficiency should have been able to generate more language. The pre and post tests were identical. Per IRB approval, all students in both the control and treatment groups took the skills tests as part of their course instruction.

Confidence survey. A confidence survey was administered to both the treatment and control groups at the beginning and end of the academic term. Students were given a prompt and had to indicate how often they performed that action or felt that emotion. Questions created to measure confidence included how often students completed certain activities using English (e.g., I use English to order a drink in a coffee shop: more than 6 times a week, 4-5 times a month, 2-3 times a week, or 0-1 time a week) and how confident or anxious they felt when using English in various situations (e.g., I feel very nervous when waiting to speak in English: Strongly agree, Agree, Undecided, Disagree, Strongly Disagree). Questions were also created to measure use of friends and peers as academic support included whether peers helped them with work and how friends supported their academic success (e.g., I practice using new words with my friends and classmates: Always, Often, Sometimes, Not much, Never). Per IRB approval, all students

in both the control and treatment groups took the surveys as part of their course instruction.

Semi-structured interviews. Interviews were conducted with students in the treatment groups. These interviews began with scripted questions, such as "What do you like best about working with your partner?" and "Are you friends outside of class?" The formation of questions was influenced by information supplied by students on exit slips and text messages (explained in the next two paragraphs). According to the answers given by interviewees, the researcher asked follow-up questions to delve more deeply into the interviewee's thoughts and motivations. Because the researcher did not know who had signed the IRB consent form, interviews were conducted and recorded with all members of the treatment groups who agreed to the interview and recording.

Exit slips. Exit slips were completed bi-weekly at the end of class as part of course instruction in the treatment group. All students answered, in writing, questions about what actions they would take, what actions they had already taken, and how interactions with the partner were helping or not helping them acquire English skills. These were collected by the teacher, who read them as a source of information for future instruction.

Text messages. Text messages were sent to students who had supplied their text numbers, separate from the IRB approval process. These messages, sent on Sunday evening after goals set on Friday would have been completed, gathered real-time data on what students were doing with their partners outside of class. They were quick questions asking about how they were helping their partners, what types of interactions they were

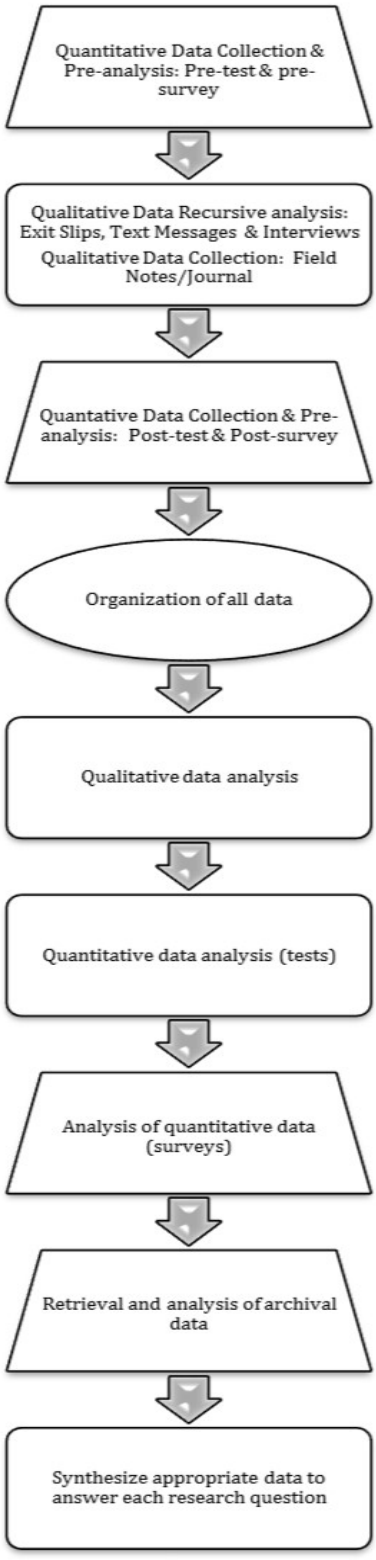
having, and when those interactions occurred. This type of data was used to confirm or disconfirm information given by students on the exit slips.

Field notes. The researcher took field notes throughout the course of the study. She noted how students interacted with each other, whether there were non-verbal cues that demonstrated something about their relationship, how students performed on classwork, and any other relevant information. This data was used to verify students' perceptions of their relationships with their peers.

Student academic records. These records were housed in a secure customer relationship management system. The researcher accessed grades, both midterm and final, and attendance records for consenting participants. These were collected for the period of 2015-2016 to show any changes in achievement or attendance.

Data Analysis

The order in which data is analyzed can affect the outcome of the analysis. Figure 4 shows the order in which data was analyzed during this study.



Pre-tests & pre-surveys were administered to control and treatment groups. Teachers of both groups reviewed pre-tests to inform subsequent instruction. Pre-surveys were secured for later analysis

Throughout the innovation, exit slips, text messages and interviews were completed and reviewed bi-weekly by the teacher/researcher to inform subsequent instruction.

Throughout the innovation, the teacher/researcher kept a journal and field notes. This reflection caused changes to instruction, but no in-depth analysis was done at this time. At the end of the innovation implementation, post-tests and post-surveys were administered to both control and treatment group. Post-tests were reviewed by the teachers; post-surveys were secured for later analysis.

Data was removed for those who did not sign IRB waiver; survey results were input on spreadsheet; interviews were transcribed; tests were labeled by group and pre or post.

Coding of interviews, exit slips, text messages, as well as field notes and researcher's journal occurred.

Pre and post tests were analyzed to make comparisons between pre and post tests, as well as between control and treatment groups.

Survey results were analyzed to make comparisons between pre and post surveys, as well as between control and treatment groups.

Archival data (grades and attendance) were pulled from the departmental database. This data covered any IEP sessions that consenting participants attended in 2016. Data was analyzed to see if there were changes pre- and post-intervention.

Results were organized according to research question. Information was synthesized to formulate answers to questions.

Figure 4. Order of analysis.

The quantitative and qualitative data were analyzed in order to answer four research questions:

1. To what extent do accountability groups help students succeed in an IEP in terms of curricular objectives, grades, and attendance?
2. To what extent do accountability groups help students gain confidence in using English?
3. How do students in an accountability group create experiences where English language is used?
4. How do accountability groups influence the students' perceptions of friends and classmates as support systems in language learning?

The tools and methods used to answer these questions are shown in Table 1.

Table 1

Data Analysis

Tool	Research Question	Analysis
Pre and post test (skills)	RQ 1	Parametric and descriptive statistical procedures
Pre and post survey of confidence level & use of friends as academic support	RQ 2	Parametric and descriptive statistical procedures
Semi-structured interviews	RQ 3 and RQ 4	Process & simultaneous coding
Exit slips	RQ 3 and RQ 4	Summative content analysis
Text Messages	RQ 3 and RQ 4	Summative content analysis
Field notes	RQ 3 and RQ 4	Summative content analysis
Academic records (grades & attendance)	RQ 1	Descriptive statistics of attendance; Parametric statistical procedures of grades

As shown in Table 1, data were analyzed to answer each specific research question:

1. To what extent do accountability groups help students succeed in an IEP in terms of curricular objectives, grades, and attendance? To answer this question, the researcher administered pre-intervention and post-intervention skills tests to both a control group and the treatment group. The results were analyzed according to differences between control and treatment group and according to changes over time (pre-intervention and post-intervention). The researcher also retrieved archival records of

student attendance and grades from the intensive English program. These records showed if there was an improvement in grades, attendance, or progress through the program before and after the treatment or a difference between the control and treatment groups.

2. To what extent do accountability groups help students gain confidence in using English? A pre and post survey was given to students in the control and treatment groups. The survey asked questions about how often students participate in certain activities and how confident they felt about using English in various situations. Descriptive statistics were used to analyze the results and compare pre-treatment to post-treatment, as well as control group responses to treatment group responses.

3. How do students in an accountability group create experiences where English language is used? Analysis uncovered the extent to which participants transferred the use of English to activities outside the classroom and the extent to which they created experiences to further their English language use. Qualitative data were gathered from the confidence surveys (questions asking how often the respondent used English in certain situations, such as English to ask directions), semi-structured interviews, exit slips, text messages, and the researcher's field notes. Qualitative data from interviews were analyzed with process and simultaneous coding. Exit slips, text messages and field notes were analyzed with summative content analysis by using codes created from the analysis of the interviews.

4. How do accountability groups influence the students' perceptions of friends and classmates as support systems in language learning? The pre-treatment and post-treatment survey contained questions about using friends and peers as support for

academic, emotional, and social success. Semi-structured interviews, text messages, and exit slips were explored to discover what partners and small groups were doing to help the participant reach his/her goals. Coding exposed underlying themes and evidence of how students perceived friends and classmates as support systems.

Quantitative and qualitative data were collected in summer and fall 2016 and analyzed in fall 2016 using documented analytical methods. This section describes the methods used to analyze the quantitative and qualitative data.

Quantitative. Three types of quantitative data were collected from control and treatment groups: surveys, tests, and archival data. Each was analyzed using different methods.

Survey. The pre and post survey had four sections for each of four constructs: peers as support, learner autonomy, emotional factors, and frequency of use of English. A reliability analysis was conducted for each construct, using Cronbach's alpha reliability coefficient. The first construct, peers as learning support, was found to have low correlation ($r = 0.70$). Therefore, it was not considered in the quantitative data analysis. The other constructs had strong correlation: learner autonomy, $r = 0.83$; emotional factors, $r = 0.95$; and frequency of use of English, $r = 0.96$.

Because pre and post surveys could be matched by respondent's birthdate and country, those who had not completed both a pre and post survey were identified and the survey results were removed from the data set. If a respondent had not answered all questions in a construct, the data for that section were removed. The average for each construct of each respondent was found for the pre and post survey and the mean of means for each construct and group (control, Treatment Group A and Treatment Group

B) was calculated. A paired samples t-test provided indicators of statistical significance for the change in scores from pre to post survey.

An independent sample t-test was also performed on the post survey results to determine if the groups were different from each other.

Test. The test was constructed so that students would have two tasks: ordering in a restaurant and giving directions from one place on a map to another. These tasks were designed with two different types of tasks: ordering in a restaurant was a gap fill where the student had to fill in the customer's part of a typical conversation; giving directions was a constructive task where the student had to write a paragraph to explain the directions.

The researcher and another educator, who had experience in testing and evaluation and who was a scorer and supervisor for the ETS TOEFL scoring team, created a rubric that scored each task on two criteria (structure /vocabulary and the comprehensibility of their language) on a scale of 1 to 4. The two raters dual rated 20% of all tests and compared scores. Their scores matched on 85% of the tests, with differences in scores being only one point apart. Differences were discussed and resolved, with this further clarification helping standardize later scores.

Four mean scores, each based on an 8-point scale, were calculated from the data for further analysis.

- A gap fill task score, (ordering in a restaurant), based on criteria of structure/vocabulary and comprehensibility,
- A constructive task score (giving directions), based on criteria of structure/vocabulary and comprehensibility,

- A structure and vocabulary score, derived from gap fill and constructive tasks
- A comprehensibility score, derived from gap fill and constructive tasks

Mean pre and post test scores were analyzed using descriptive statistics, based on each task (ordering in a restaurant and giving directions) and on each criterion (structure/vocabulary and comprehensibility).

Archival data. Grades and attendance records for each participant were pulled from the departmental customer relationship management system. Each participant took four classes: reading/writing, listening/speaking, communication, and structure. The grades from all four classes, based on a 100-point scale, were averaged for the session that the innovation was implemented. This allowed the data to show more clearly overall performance and helped validate the findings, since grades were not just from the class where the innovation took place. It also mitigated the variable of having different teachers for the control and treatment groups.

Once the overall grade average was calculated, the three groups were compared with an ANOVA test to see if the difference between the three groups was significantly significant. The ANOVA was negative for the null hypothesis that there was no difference among the three groups. The ANOVA was followed by two t-tests to see where the variance was between the three groups.

Attendance was reported as minutes missed in each class. The minutes missed by a student in all four classes were calculated by adding all minutes missed in each class. A mean number of minutes missed in each of the three groups was calculated and compared.

Qualitative. Qualitative data was collected through interviews and verified by field notes, exit slips, and text messaging. A commercial service was used to transcribe the interviews. The researcher's objectivity was increased because time had passed between the interviews and the data analysis. Due to the low language proficiency of the participants, the interviews did not contain a wide breadth of content. Therefore, deductive analysis was used: five codes were created a priori, based on the research questions, the interview questions, and the expert knowledge of the researcher. As the interviews were read and analyzed, four other codes emerged. The researcher made notes on the transcripts themselves and then created a document with paraphrases and quotes that supported each code. This document was then used to re-read the material and subdivide the material under each code, expanding into sub-codes. This was used to demonstrate the various ways that participants experienced the accountability group innovation.

After analyzing the interviews, the text message exchanges between the researcher and students and the researcher's field notes were used to verify the comments that students had made during the interviews. When real-time data of text messaging was available, it did verify comments in the interviews, but yielded very little further data because of its brevity.

Last, participants' interview remarks about goal-setting were compared to their exit slips to see if any patterns emerged, if any anomalies were noted, or if further insight could be gained.

Threats to Validity

With any type of research, threats to validity exist, and the researcher must work to mitigate them. This mixed methods action research was no exception.

One common threat to validity is the Hawthorne effect. Participants or stakeholders may, consciously or unconsciously, regulate their behavior or edit their remarks as a result of being observed. Students may have wanted to please the teacher and may have answered in ways that she wanted to hear to gain her favor.

Another threat to validity is the experimenter effect. The researcher may have bias that unconsciously influences the participants of the experiment. In educational research, this is called the "Pygmalion effect." The teacher/researcher has a preconceived idea or bias about a student or group of students, and this bias affects the learning of those students.

To mitigate both the Hawthorne and the experimenter effect, a trusted colleague collected the consent forms. This formed a psychological separation for students and teacher between the research and the class. At that time, it was made clear that the teacher did not know who consented and that consent had no effect on the students' grades. After the consent forms had been signed, the research was not discussed with the students. From their perspective, the partner work and goal setting were just the way the class was conducted. The teacher/researcher also stayed mindful of this threat and worked to control her own actions in the classroom.

A third threat to validity is diffusion. As a teacher in the intensive English program, the researcher might have communicated information about the intervention to the teacher of the control group. The control group teacher could have adopted some of

the techniques, which could have affected the performance on post-intervention skills of the control group, not because the intervention was ineffective, but because parts of the intervention were used in the control group. To mitigate this threat, the Accountability Groups Innovation and its methods were not emphasized. The researcher did not talk about the treatment nor its methods until after the data collection had been completed.

When the threats to validity are mitigated, quantitative and qualitative data can be rigorously analyzed, leading to "well-validated conclusions" (Ivankova, 2015, p. 129). In their Legitimation Model, Onwuegbuzie and Leech (2006) talk about weakness minimization. This is the "extent to which the weakness from one approach is compensated by the strengths from the other approach" (Ivankova, 2015, p. 281). It is important for all studies using mixed methods research to ensure that the quantitative and qualitative data collection tools strengthen the overall analysis of the problem and the innovation. In this study, quantitative pre and post intervention surveys collected students' perceptions of change in confidence and use of peers as support over the duration of the study, while qualitative data from interviews, text messages, and exit slips showed what happened in situ to attend to the validity of the quantitative data. The quantitative data from pre and post tests of curricular objectives assessed the extent of change in students' academic skills. The qualitative data described the how and why of the changes, and was helpful in further analyzing the magnitude of change in students' academic progress. Data collected in these two methods were combined, synthesizing the results and the researcher's interpretations of them, so that meta-inferences could be made. Effective triangulation enhances the validity of the inferences made from research

findings to create a larger, consolidated view of accountability groups as a tool for language learning.

Chapter 4 - Results

Based in the ecological perspective and socio-cultural theory, the challenge of the accountability group innovation was to facilitate students' transfer of learning experiences from other-control to self-control, allowing them to create new affordances in which to practice and use the language. This transfer of control began in class when students were assigned partners with whom they worked, both inside and outside of class, for the duration of the 8-week session. Each student created weekly goals to create language learning experiences outside of class, and they shared these goals with their partner. Data were collected to investigate if this innovation could help students succeed in an intensive English program in terms of curricular objectives, grades, and attendance; if their confidence in using English would improve; how they created experiences to use English; and how students perceived their classmates and friends as supports in learning language. This chapter presents quantitative data from pre and post tests and surveys, archival grades and attendance records, as well as qualitative data collected from interviews which were verified through text messaging, exit slips, and field notes. In the qualitative data, the words and language of the participants is, in many instances, retained to honor the experience and expression of the participants. It also gives insight into the language proficiency level of the students.

Working Together

Assignments were created to aid students in working together, and many of these assignments were "two students, one paper" submissions. Students described their various methods of working together on assignments. One student, Summer, said that she and her partner wrote the papers together, taking turns writing and talking about the

work as they wrote. Summer explained how she and her partner would finish each other's sentences: "Sometimes, there's a sentence that I cannot finish, and then she's like, 'What about this?' And I said, 'That's good. That's gonna be a good sentence.' " Some students reported doing the assignment independently and then checking over answers after the work was completed. Others talked of dividing the work for expediency, with one person doing half of the assignment and the other doing half without any real collaboration. Moheen talked about using his partner to understand the assignment and brainstorm before doing the actual work separately. Abdullah and Donna saw their partners as distractions, losing focus on the assignment and speaking their native language. Abdullah explained: "he sometimes talk about with me Arabic, he use the compute in the class when the teacher explain…one time he talked with me about the scholarship (off topic from class)."

Knowledgeable other. After working with the same partner for seven weeks, students viewed their partner as a knowledgeable other. Half of the students, in the process of being interviewed, were descriptive of their partners' characteristics when asked what they liked about their partner. Most recognized their partner as a knowledgeable other by using the adjective "smart;" Mahwah, Risa, and Taka said their partner was "nice" or "kind;" and Yazan and Mahwah described their partner as easily approachable. For example, Yazan had this exchange with the interviewer:

> Yazan: He (my partner) is smiling and help.
> Interviewer: That makes you feel good about English because he is smiling.
> Yazan: Yes, smiling and help.
> Interviewer: And helping, yes.
> Yazan: Sometime when you ask someone and help you, he help you but he don't want to help you, like inside.
> Interviewer: And you can tell, because he's not smiling.

Only Yazan described his partner as "just partner…like every partner."

When students talked about how they worked with their partner, most talked of helping each other. Mayumi even described the help as teaching: "I like writing English, but she likes speaking…So she teach me about speak English. I taught her writing English…She speak, I write…I can teach her." Moeen, from Saudi Arabia, who had a partner of lower level proficiency, felt an obligation to his partner and did not have a sense of reciprocity: "I am with Abdul because I want to help Abdul." Elvin expressed doubt about his Japanese partner's ability, but turned to another classmate, Lynn, as a knowledgeable other: "Sometimes the word I don't know, but he can tell me the word. Some were, I not sure this right, so I need to ask to (Lynn), this word it really is this means? She tell me, 'Yes, this is.'" However, the Japanese student's vocabulary and proficiency level was higher than Elvin's level.

Five students talked of learning vocabulary with their partner, giving specific examples of words that they discussed with their partner in class, such as "flip-flops and high heels" or the always-confusing directional words "right and left." A few moved beyond simple vocabulary to sentence construction, like Summer's earlier description of working with her partner, who helped finish her sentence when she didn't know how to complete it.

Attendance. Working with an accountability group created engagement and a sense of responsibility to attend class. Attendance for each participant is stored in the departmental Salesforce CRM in the form of "minutes missed" in each class. Those minutes were added to give a total minutes-missed for the entire session in all classes, not

just the class using the accountability group innovation. The mean number of minutes missed for each group is shown in Table 2. Students in both treatment groups had better overall attendance during the treatment session than those in the control group. While no one in the treatment groups failed for missed class time in any of their four classes on their schedule, five students in the control group (n = 15) missed more than the allotted time in at least one of their classes and consequently failed the session.

Table 2

Mean Missed Class Time, in Minutes, Number of Students with Perfect Attendance and with Attendance Failure, by Group

Group	Mean missed class time, in minutes	Number who missed no class	Number who failed for missed class time
Treatment Group A, n = 16	190.7	7	0
Treatment Group B, n = 12	87.7	5	0
Control, n = 15	1152.4	7	5

Upon reviewing the nationalities of the students in the control group, it was found that the five students who failed were Chinese. In this intensive English program, archival records show that 38% of Chinese students fail a class for which they are registered due to non-attendance. For all other students, the failure rate due to absences is 5.2%. In fact, 82% of all absence failures are Chinese students.

Elvin, a Chinese student in Treatment Group A, explained a relationship he had with his English teacher back in China. The teacher had given him a grammar book before he came to America. The teacher made regular assignments in the book, which

the student did, taking pictures of the completed work. He then sent the picture to his teacher, who corrected it and sent it back. The student felt that he didn't need to focus on his grammar class in the U.S. university because he had this support from back home.

As evidenced by their attendance, Chinese students continued study throughout the session and were more engaged with the accountability group innovation. However, the Chinese students who finished the course in the control group performed better than those who retained in the treatment groups on curricular objectives as shown through pre and post tests, although the *n* of 3 for the control group is not powerful and some of the standard deviations are large. In the treatment group, scores of Chinese students on the restaurant task rose 0.56 points (*SD* = 1.42). In the control group, scores of Chinese students improved 1.33 points (*SD* = 4.04). On the directions task, Chinese students in the treatment group improved an average of 0.44 points (*SD* = 1.51) while those in the control group improved an average of 1.33 points (*SD* = 1.53). On the scores for structure/vocabulary and comprehensibility, the treatment group improved 0.33 points (*SD* = 1.12) and 0.67 points (*SD* = 1.32) respectively, while the control group students improved 1.67 points (*SD* = 1.53) and 1.00 point (*SD* = 3.61).

Table 3

Comparison of Pre, Post Scores on Each Task and Criterion on 8-Point Scale, for Chinese Speakers in the Control and Treatment Groups

	Ordering in Restaurant						Giving Directions					
	Control			Treatment			Control			Treatment		
	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change
Mean	3.67	5.00	1.33	4.22	4.78	0.56	2.67	4.00	1.33	3.78	4.22	0.44
SD	2.08	2.00	4.04	1.92	1.30	1.42	1.15	1.00	1.53	1.39	1.09	1.51
n	3	3		9	9		3	3		9	9	
	Structure and Vocabulary						Comprehensibility					
Mean	3.00	4.67	1.67	4.22	4.56	0.33	3.33	4.33	1.00	3.78	4.44	0.67
SD	1.00	0.58	1.53	1.30	1.01	1.12	1.15	2.52	3.61	1.30	1.33	1.32
n	3	3		9	9		3	3		9	9	

Different languages. As students described their experience with their partner, six students talked about the partner's native language—whether it was the same or different than their own. Five of the six expressed a desire to pair with someone who spoke a language different than their own because this would help them improve their English skills. According to Lynn,

> different country partner is best. They (same language partners) promise 'I will speak English, I will speak English.' So, we just use our Chinese. No one want to check it…no one will check on the dictionary…They just let it go. They will never learn the word, I think. If your partner is foreigner, and you don't know the word, 'Oh my god. What should I do? I don't know translate (translation).' You know all these word. You can know the word and use word and understand. This is the way to know some new word.

This sentiment was echoed by others: they liked having a partner from another country because it expanded their affordance network and promoted actual interactions in which they would engage (effectivity sets); thus, they viewed it as the best way to practice and learn English.

Meanwhile, Moeen had a partner who shared the native language of Arabic. When asked if he helped his partner with English, Moeen replied that he was not helping very much because he was just translating from English to Arabic for his partner.

Lynn also explained another aspect that could affect the acquisition of language in the intensive, academic program: teaching styles.

> If you are same country, the education similar. You will make some mistake. If you are same country, you can't found the mistake. That means, let it go, let it go. In different country, because education is not the same, so maybe this sentence, they know correct answer. You don't know and he can tell you. When you talk with him or she, if you have some mistake and she will tell, 'Oh, this is mistake.' In same country, they just, 'Okay, you know.' That's all.

Overall, the students had a sense that they had to use English to improve, and that the best way to do that was to be partnered with someone from a different country. However, the quantitative data collected from the pre and post tests of curricular objectives was inconsistent and did not strongly support this supposition, except in regards to vocabulary and structure where the data were convergent.

As shown in Table 4, the scores of students who had a partner with a different native language and the scores of those who had a partner with the same native language were compared. Scores of students with partners of a different language improved more on structure and vocabulary, with an increase of 0.31 points (SD = 1.08), as compared to those whose partner had the same language, with scores that remained constant (SD = 0.89).

Table 4

Comparison of Pre, Post Scores on Structure and Vocabulary Criterion on 8-Point Scale, by Partner Language

		Structure and Vocabulary		
		Pre	Post	Change
Same Language Partner	*M*	4.17	4.17	0.00
	SD	0.75	1.17	0.89
	n	6	6	
Different Language Partner	*M*	4.38	4.69	0.31
	SD	1.02	0.95	1.08
	n	16	16	

Gender. Overall, students saw their accountability partner as a friend. When asked if they were friends with their partner, eight of twelve said yes. Partners were generally of the same gender, with the exception of one pair. Interestingly, the students who considered their partner a friend were divided by gender—all female interviewees said their partner was a friend, while only 3 of 7 males did. Hiro said that "sometimes" they were friends, and other males said that they were not friends with their partners. Mayumi described her friendship with her partner "She interested in Japan…I ask about her country, so we are good friend." Risa reported, "We cooked together and watched a movie…we hang out." Summer highlighted the communication between she and her partner: "We can communicate very well. She's also very smart. That's good. That's nice."

Although females reported stronger social bonds, the quantitative data gathered by pre and post tests of curricular objectives was divergent, showing that males made more improvement. On the gap-fill restaurant task, males and females made approximately the same improvement: males improved 0.47 points (*SD* = 1.66) and

females improved 0.50 points (*SD* = 1.69). Males improved more than females on the constructive giving directions task: males improved 1.06 points (*SD* = 1.68) and females improved 0.38 points (*SD* = 1.60). However, females' scores on the giving directions task were higher on the pre-test (4.25, *SD* = 1.28) than males' scores on the post-test (4.18, *SD* = 1.33). Overall, females did better on the task, with a mean post-test score of 4.63 (*SD* = 1.69), as shown in Table 5.

Table 5

Comparison of Pre, Post Scores on Each Task on 8-Point Scale, by Gender

		Ordering in Restaurant Control			Giving Directions Control		
		Pre	Post	Change	Pre	Post	Change
Male	Mean	5.00	4.33	-0.67	3.67	4.33	0.67
	SD	2.65	1.53	1.53	2.08	1.53	2.52
	n	3	3		3	3	
Female	Mean	4.50	5.00	0.50	4.00	4.00	0.00
	SD	1.76	1.79	2.81	2.10	1.26	1.79
	n	6	6		6	6	

Looking at the different criteria, not much difference was noted on the structure and vocabulary, but scores of males increased more than females in terms of comprehensibility. As seen in Table 6, the males gained 1.24 points (*SD* = 1.35), while the females gained 0.63 (*SD* = 2.00).

Table 6

Comparison of Pre, Post Scores on Each Criterion on 8-Point Scale, by Gender

		Structure and Vocabulary Control			Comprehensibility Control		
		Pre	Post	Change	Pre	Post	Change
Male	Mean	4.33	4.67	0.33	4.33	4.00	-0.33
	SD	2.08	0.58	1.53	2.52	1.73	1.53
	n	3	3		3	3	
Female	Mean	4.33	4.50	0.17	4.17	4.50	0.33
	SD	1.86	1.05	1.94	1.60	2.17	2.42
	n	6	6		6	6	

Role of goal-setting in experience creation

One important part of the accountability group innovation was setting goals each week to use English in a context beyond the classroom. Because of students' low level of proficiency and the newness of goal-setting, the process of setting goals was difficult for them to understand. When interviewed about their goals at the end of the program, it was difficult for them to understand what was being talked about. The first few interviewees did not understand completely what the interviewer was referencing, even though goal setting was an assignment completed in class each week. When the interviewer revised the interview protocol and showed later students a copy of their paper, they could better articulate their thoughts about the goal-setting process. The students' responses were sorted into three groups: those that felt goal setting was beneficial, those that felt it was not beneficial, and those who only identified large, long-term goals and did not talk about the weekly goal-setting process.

Five of the eleven students who were interviewed felt that the weekly process was "good." Two of those five felt that they would not have completed the goal if they had not written it down. Hiro said, "I like this. I don't forget if I have this paper." He also said that he plans to continue to write his goals throughout his studies. Risa said that it was good to write the goals, but only "so-so" in helping her fulfill the goal. She thought it was a good way to think about English and how it is used. Three of the five students told what their specific goals generally were (e.g., who they were going to talk with) and how they fulfilled the goals.

Three Chinese students, Lynn, Donna, and Valerie, said that they didn't like writing the goals. Valerie reported that it was repetitive to write a goal each week, as she and her friends generally repeated the same goal week after week. Donna said that defining these goals did not push she and her partner because they chose to write goals of activities and conversations that they would have done every day regardless of the goal-setting exercise. For example, Donna listed her goal as "I will go to the supermarket to talk with the cashier about buying some food," and Lynn created the goal of talking to her roommate when she walked the dogs.

Summer, Abdul, and Moeen could only articulate a large, long-term goal: to learn English well. When asked how to achieve that goal, or the smaller goals that were written in class, Summer talked about her study methods rather than goals written in class.

Experience with others (non-classmates). Many students set goals to talk to people who were not their classmates. Almost half of the students mentioned that they

use English with friends. Moeen and Mayumi highlighted friends from many different countries. Mayumi specifically noted that she goes to parties to learn English:

> Mayumi: Party ... There are many language because I'm Japanese, but the party has another country friend. We have American, Arabic, Japanese, Chinese, Korean so it's a national party. International party. My another country friend teach me about English, and I was teaching Japanese.
> Interviewer: So they taught you English and you taught them Japanese?
> Mayumi: Yeah…So I think party is so good, learning English.
> Interviewer: Yes. Why?
> Mayumi: I can made many friend, another country friend.
> Interviewer: Yes.
> Mayumi: I learn a lot of speak English if I am learning, talking just Japanese…But the party is no. Party, if I go to party, I can speak English right.

Another student, Moeen, also puts himself in situations where he can improve his English. He met an American and a South African at the mosque with whom he speaks English. He also goes to departmental activities that are planned for students to make social connections in an English-speaking environment: "Almost I am staying in the (program) groups and conversation club, and some students from Japan and South Korea, I talk with them in English."

As students talked about their interactions outside of class, another theme was revealed: family members, specifically cousins, effect their language learning outside of class. Yusef, an Arab student, said that he had lived in a homestay where he spoke English, but was spending Ramadan with his cousins, with whom he never spoke English. Yazan said that he asked his cousin to do his homework, but his cousin wouldn't do it. Summer, from Djibouti, was living with her aunt and cousins. Although she spoke Somali with her aunt, the cousins "don't speak our first language. They speak English. They understand the first language, but actually they cannot…I speak English with them for sure because they don't understand what I'm saying if I start talking."

Host families and roommates were also commonly mentioned. Taka said he wanted to practice English in the house, and always talked to the host son. Abdullah said he talks to his host family for about an hour each day. Mayumi said that her host mother, a lawyer, was very busy and didn't usually have much time to talk. Moeen, the only student who mentioned his roommates, said that he never spoke English with them.

Perception of autonomy

The goal of the accountability group innovation was to increase students' autonomy as they wrote goals and identified situations where they could practice the English language they were learning in class. A survey with seven Likert-scale items was presented to gauge students' autonomy in the use of English. Presented in Appendix C, the items included statements such as "I start conversations in English," "I try to find many ways to use English," and "I ask for help from English speakers." Participants had to choose from five adverbs, and the adverbs were subsequently coded with a numerical value: always (5), often (4), sometimes (3), not much (2), never (1). Each participant's scores were averaged, so a score closer to 5 is a student whose self-perception is more autonomous and a score closer to 1 is a student who does not perceive themselves as autonomous. As shown in Table 7, students in the treatment groups had a larger change in their self-perceptions of autonomy during the study period. Treatment Group A showed the most progression, with an average increase of 0.33 ($t = 1.49$, $p < 0.17$), followed by Treatment Group B with an average increase of 0.14 ($t = 1.40$, $p < 0.19$). Although these differences did not show evidence of significance at the 90% confidence level, they showed more improvement than increases experienced by the control group (mean difference of 0.11, $t = 0.57$, $p = 0.58$).

Table 7

Treatment and Control Groups: Autonomy Differences, Pre and Post, on 5-Point Scale

Group	N		Pre Survey	Post Survey	*M2 - M1*	*p*	*t*
Treatment A	10	*M*	3.51	3.84	0.33	0.17	1.49
		SD	0.47	0.65			
Treatment B	12	*M*	3.94	4.08	0.14	0.19	1.40
		SD	0.31	0.42			
Control	10	*M*	3.67	3.79	0.11	0.58	0.57
		SD	0.52	0.61			

*mean difference is significant at $p \leq 0.05$
**mean difference is significant at $p \leq 0.10$

In addition to the survey responses, students told of their increasing autonomy with language during interviews with the researcher. Abdullah, in his fourth month in the U.S., said:

> Maybe in the first day in United States I didn't have any word or I don't talk with people…in the first, eh, month, eh, I, eh, I understand some of words and the second month I uh talk with people in the basic…Right now, in the third month I understand with people but I don't have any answer about question.

Elvin said that he never went to the coffee shop and ordered coffee because he didn't know how. After learning about ordering coffee in class, he now goes to coffee shop and enjoys coffee often. Mahwah told how she always went shopping with her brother and never asked questions, always relying on him to help her and speak for her. But, now she goes to the market and talks with the cashier, asking any questions that she needs. Students recognized their growing autonomy.

Experience with classmates. The students' sense of autonomy can be seen in their interactions with classmates outside of class. When asked directly, each of the

students reported getting together with their accountability partner outside of class. However, upon further discussion, the nature of this interaction seemed to fall in to three categories: those who socialized with their partner in English around joint interests, those who socialized in their first language, and those who reported meeting outside of class but lacked details or had accounts that conflicted with their partners.

First, some students talked about joint interests they had with their partner or social activities they had done together. Elvin and Taka independently reported getting together around campus to talk about a joint interest—anime:

> Taka: We talking always…anime?
> Interviewer: Anime? You both like anime?
> Taka: Umm-hmm (affirmative) We like both.
> Interviewer: Do you ever read anime together? Or is the language different?
> Taka: Yes, a few time.

Mayumi and Summer liked to compare daily life and customs in their countries (Japan and Djibouti); however, Summer lived far from campus which impeded the feasibility of their socializing away from class. Summer did express a desire to socialize with her partner though. Mayumi actually got together with the other members of her accountability group. They would get together at the apartment of Mawah, an Arab student, and cook or watch movies, or as Risa explained, they liked to "hang out." Mawah showed the interviewer photos and a video of the group cooking and talking about the ingredients in English. Actually, they were planning to get together the afternoon of the interviews and cook clams. Mawah showed me a text exchange where Risa asked if she (Mawah) could eat clams, but Mawah didn't know what clams were. In response, Risa sent a picture from Google of a clam. Mawah responded, "I do not know,

but we can try." All of these groups did not have their native language in common, so they used English.

Second, two pairs reported meeting outside of class, but because they could speak their native language together, they did: Moeen saw his partner at the mosque at prayer time, but spoke in Arabic; Donna and Valerie worked together in every class and cooked food together, but spoke only in Chinese.

> Interviewer: Do you speak English while you are cooking food or Chinese?
> Valerie: Chinese.
> Donna: Yeah. Chinese.
> Interviewer: Chinese? Why?
> Valerie: Because we are Chinese.
> Donna: Yeah.
> Interviewer: Do you ever think, "Oh, we should speak English."
> Valerie: Maybe we talk about another things. We can't say English but, it's [inaudible] when we can speak English quicker. Many words we don't know yet.
> Interviewer: Could you look them up?
> Valerie: What?
> Interviewer: Look. Translate. "Words, I don't know. What is this in English? Let me look."
> Valerie: Yeah.
> Interviewer: Do you do that, or no? Too hard?
> Donna: It's too hard.
> Valerie: Yeah.

The third category were students who reported getting together outside of class, but their account differed greatly from their partner or lacked any details of what they did. Two students reported that they met their partner outside of class to talk or socialize with friends, but the partner, interviewed independently, said they did not get together. One pair had identical responses—that they met outside of class, but neither could say where, when, or what they did.

Frequency of language usage. The survey about student confidence and autonomy also included four items concerning how often students use English in certain

situations that were simulated and taught in the course curriculum: ordering a drink in a coffee shop, shopping for clothes, ordering food in a restaurant, and asking or giving spatial directions. Respondents could choose one of four frequencies: more than 6 times a month, 4-5 times a month, 2-3 times a month or 0-1 times a month. For ordering a drink in a coffee shop, the choices were based on a week time span rather than a month. These responses were then assigned a numerical value from 4 for the most frequent to 1 for the least frequent. Each participant's responses were averaged to a number between 1 and 4. Therefore, a higher score is more frequent use of English and a lower score is less frequent.

Table 8 illustrates that Treatment Group A used English in these situations slightly more at the end of the course than they did at the beginning (0.15, $p = 0.48$); Treatment Group B used English less (-0.23, $p = 0.14$). However, the control group increased their reported use of English (0.25) at a 90% confidence interval.

The treatment groups reported higher frequency of using English in the pre-survey with average scores of 2.73 ($SD = 0.89$) and 3.00 ($SD = 0.63$), while the control group showed less frequency of using English with average scores of 2.45 ($SD = 0.60$). At the end of the innovation, both treatment groups showed higher frequency with average scores of 2.88 ($SD = 0.69$) and 2.77 ($SD = 0.72$), and the control group showed lower frequency with average scores of 2.70 ($SD = 0.50$). Although the control group improved their frequency score more than the treatment groups, the treatment groups were using English more frequently than the control group at the end of the study.

Table 8

Treatment and Control Groups: Frequency Differences, on 4-Point Scale

Group	N		Pre Survey	Post Survey	*M2 - M1*	*p*	*t*
Treatment A	10	*M*	2.73	2.88	0.15	0.48	0.73
		SD	0.89	0.69			
Treatment B	13	*M*	3.00	2.77	-0.23	0.14	-1.59
		SD	0.63	0.72			
Control	10	*M*	2.45	2.70	0.25	0.06**	2.12
		SD	0.60	0.50			

*mean difference is significant at $p \leq 0.05$
**mean difference is significant at $p \leq 0.10$

Affective Factors

The survey also contained seven Likert-scale items related to participants' emotions related to using English. Items included "I enjoy speaking English," "I feel very nervous when waiting to speak in English," and "I have no fear of speaking English." Likert-scale choices were later assigned a numerical value according to whether the item described a negative attribute or a positive attribute: Strongly agree (5 or 1), Agree (4 or 2), Undecided (3), Disagree (2 or 4), Strongly disagree (1 or 5). An average of each participant's responses was calculated, with a score closer to 5 being more positive emotions towards using English and a score closer to 1 being more negative emotions towards using English. While all groups reported more positive emotions towards using English during the study (Table 9), the control group had the largest increase (0.37, $p = 0.20$). However, the slightly lower positive shift in the treatment groups was statistically significant at the 90% confidence interval (Treatment A 0.27 improvement, $p = 0.07$; Treatment B 0.31 improvement, $p = 0.09$).

Table 9

Treatment and Control Groups: Affective Differences, on 5-Point Scale

Group	N		Pre Survey	Post Survey	*M2 - M1*	*p*	*t*
Treatment A	11	*M*	3.21	3.48	0.27	0.07**	2.07
		SD	0.60	0.49			
Treatment B	11	*M*	3.32	3.64	0.31	0.09**	1.87
		SD	0.48	0.56			
Control	10	*M*	3.23	3.60	0.37	0.20	1.37
		SD	0.79	0.63			

*mean difference is significant at $p \leq 0.05$
**mean difference is significant at $p \leq 0.10$

Traditional measures of success

Taken together, the quantitative and qualitative data, although somewhat inconsistent, show that changes were occurring in how the students in the treatment group experienced the use and learning of the English language. Students talked about how they worked together, viewing each other as knowledgeable others, as well as seeking out speakers of other languages with whom to practice. They began to set goals and used those goals to increase their perceived autonomy and to increase the frequency that they used English. Researchers often attempt to measure these changes with traditional measures. Here overall grade averages and pre and post tests of curricular objectives will be presented.

Grades. Class grades are an accepted method of measuring change. Archival grades for each consenting student in the treatment and control groups were pulled from the departmental SalesForce data management system. The data were organized into

three groups: control, Treatment A (Summer 2016) and Treatment B (Fall 2016). Each student had four grades for each session, corresponding to each course taken. Each grade was based on a 100-point scale. The four grades were averaged together to compute an average grade for the session. This average grade explains the overall performance of the student, not just in the course with accountability groups, but in all courses.

A one-way analysis of variance was conducted to evaluate the relationship between the mean overall grade average and the student's group. The independent variable was the group (treatment A, B, or control), and the dependent variable was the mean overall grade average. The ANOVA was significant, $F(2, 40 = 7.10, p = 0.00)$. The obtained F value of 7.10 is higher than the critical value of F which is 3.23. Therefore, the obtained F value is likely to occur by chance with a $p \leq 0.01$. The difference is significant, so follow-up tests were conducted to evaluate pairwise the differences among the means.

Because the variances among the three groups ranged from 2 to 40, it was not assumed that the variances were homogeneous and post hoc comparisons were conducted with the use of two-sample *t*-tests. There was a significant difference in the means between and Treatment Group A and the control group, as well as between Treatment Group B and the control group, but no significant difference between Treatment A and Treatment B groups. The treatment groups showed higher mean overall grade averages in comparison to the control group. The 95% confidence intervals for the pairwise differences, as well as the means and standard deviations for the three groups are reported in Table 10.

Table 10

95% Confidence Intervals of Pairwise Difference in Mean Changes in Overall Grade Averages

Group	*M*	*SD*	Treatment A	Treatment B	Control
Treatment A	82.2	10.52		t (18) = 2.10 p = 0.02	t (16) = 1.70 p = 0.01
Treatment B	89.2	3.12	t (18) = 2.10 p = 0.02		t (14) = 1.80
Control	55.8	40.60	t (16)= 1.70 p = 0.01	t (14) – 1.80 p = 0.00	

While the mean overall grade point average shows that there was significant difference between the treatment groups and the control group, it is also enlightening to note the range of individual grade averages within each group. In Treatment Group A (n = 16), only three students had an overall grade average less than 80 points, and only one of those was below 60 points. In Treatment Group B (n = 12), no students had an overall grade average less that 80 points. However, in the control group (n = 15), eight students (53.3%) had overall grade averages less than 80 points, and six of those eight were below 60 points. The overall grade averages, calculated from all four courses taken during the session, trended higher in the treatment groups than in the control group.

Table 11

Range of Mean Overall Grade Averages, Based on 100-Point Scale, by Group

	Overall Grade Average, of 100 Points				
Group	0 – 20 points	21 – 40 points	41 – 60 points	61 – 80 points	81 – 100 points
Treatment A	0.0%	0.0%	6.3% (1 student)	12.5% (2 students)	81.3% (13 students)
Treatment B	0.0%	0.0%	0.0%	0.0%	100.0% (12 students)
Control	26.7% (4 students)	6.7% (1 student)	6.7% (1 student)	13.3% (2 students)	46.7% (7 students)

Curricular objectives. As explained earlier in the chapter, a pre and post test was given to participants in the control group and treatment groups, covering the curricular objectives of the class in which the accountability group innovation occurred.

Response length on constructive task. In the lower proficiency levels, the amount of language a student can produce is often linked to a gain in proficiency. As students' vocabulary expands, they are able to convey fuller and richer meanings, and are more adept at producing language. One task on the test, giving directions to a location on a map, was an open-ended constructive task that allowed students to write freely. It was not timed. By comparing the average word count on the pre and post task for the treatment and control groups, it can be seen that the treatment groups became more adept at writing longer responses than the control group. As shown in Table 12, students in Treatment Group A increased the average word count of their responses, with an average increase of 9.00 words (*SD* = 13.28). In the Treatment Group A post-test, only three

students wrote two to four words less than the pre-test, while the other nine students wrote from two to 35 words more.

Treatment Group B responses increased an average of 12.35 words (*SD* = 13.20). In the Treatment Group B post-test, two students wrote from 7 to 17 words less than they had on the pre-test, while the remainder wrote from 3 to 33 more words.

Conversely, the control group responses actually became shorter, with an average decrease of 1.10 words (*SD* = 14.58). In the control group post-test, six of 10 students wrote shorter answers on the constructive task than they had on the pre-test, ranging from one word less to 19 words less.

Table 12

Average Word Count for Pre, Post Giving Directions Task by Group

Group	Word Count Pre-Test	Word Count Post-Test	Change in Word Count
Treatment Group A (*n* = 11)	35.73	44.73	9.00
	SD 13.14	*SD* 14.66	*SD* 13.28
Treatment Group B (*n* = 17)	41.12	53.47	12.35
	SD 18.14	*SD* 14.48	*SD* 13.20
Control Group (*n* = 10)	51.80	50.70	-1.10
	SD 10.86	*SD* 19.25	*SD* 14.58

Further analysis of the pre and post tests included the steps outlined in the procedure for scoring and compiling the scores of the various parts of the pre and post test that was explained in Chapter 3. There were two tasks—giving directions and ordering in a restaurant—and two criteria, structure / vocabulary and comprehensibility.

Table 13 shows the mean scores for each treatment and control group for each task and criteria, and the difference between pre and posttest means was calculated for each score.

Table 13

Mean Scores on Each Task and Criteria, Based on 4-Point Scale, by Group

		Ordering in Restaurant Structure/ Vocabulary			Ordering in Restaurant Comprehensibility			Giving Directions Structure/ Vocabulary			Giving Directions Comprehensibility		
		Pre	Post	Change	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change
Treatment A $n = 11$	Mean	2.55	2.18	-0.36	2.00	2.00	0.09	1.91	2.18	0.27	1.82	1.91	0.09
	SD	0.69	0.75	0.81	0.89	0.89	1.04	0.70	0.75	0.79	0.75	0.83	0.70
Treatment B $n = 17$	*Mean*	2.29	2.35	0.06	1.71	2.59	0.82	1.76	2.29	0.53	1.53	2.35	0.82
	SD	0.69	0.49	0.75	0.85	0.94	0.88	0.56	0.59	0.80	0.72	0.86	1.07
Control $n = 10$	*Mean*	2.20	2.50	0.30	2.50	2.30	0.30	2.00	2.10	0.10	1.80	2.00	0.20
	SD	0.92	0.53	1.16	0.97	1.16	1.40	1.05	0.57	0.99	0.92	0.82	1.03

A one-way analysis of variance was conducted to evaluate the relationship between the mean difference of the overall score and the group. The independent variable was the group: Treatment Group A, Treatment Group B, and the control group. The dependent variable was the mean difference of the overall score from pre to post test. The ANOVA was significant, F (2, 35) = 3.87, p = 0.03. However, a Tukey post hoc showed significance between Treatment Group A and Treatment Group B (2.29, SD = 0.91, p=0.04), but did not show evidence of a statistically significant difference between Treatment Group A and the control group (1.89, SD = 0.93, p = 0.12) nor Treatment Group B and the control group (-0.40, SD = 1.20, p = 0.92).

Next, the mean post test scores were analyzed using descriptive statistics, based on each task (ordering in a restaurant and giving directions) and on each criterion

(structure/vocabulary and comprehensibility). Table 14 shows the mean group posttest scores for each task and criterion, based on an 8-point scale. Treatment Group B, which occurred in the fall session of 2016, had the highest post-test scores across both tasks, ordering in a restaurant and giving directions, and both criteria, structure/vocabulary and comprehensibility. However, Treatment Group A, which occurred in the summer session of 2016, had lower scores than the control group in all tasks and criteria.

Table 14

Mean Post Test Scores, by Task and by Criterion, on 8-Point Scale

		Post Test Scores, based on 8-point scale			
		Ordering in Restaurant	Giving Directions	Structure / Vocabulary	Comprehensibility
Treatment A	Mean	4.18	4.09	4.36	3.91
$n = 11$	*SD*	1.40	1.51	1.12	1.22
Treatment B	Mean	4.94	4.65	4.65	4.94
$n = 17$	*SD*	1.30	1.32	0.79	1.48
Control	Mean	4.80	4.10	4.60	4.30
$n = 10$	*SD*	1.55	1.20	0.84	1.83

No patterns concerning the effect of the treatment emerged from the overall data by group. However, two interesting patterns not discussed previously were noted in the analysis of various sub-groups.

Japanese speakers. Overall, Japanese speakers had the highest pre and post scores, in both treatment and control groups, in the restaurant and directions tasks, as well

as the structure/vocabulary and comprehensibility criteria. However, the change in scores was pronounced between the treatment and control groups. Japanese in the treatment group improved on all tasks and criteria; however, Japanese students in the control groups had lower scores on the post test, as shown in Table 15.

Table 15

Comparison of Pre, Post Scores on Each Task and Criterion on 8-Point Scale, for Japanese Speakers in the Control and Treatment Groups

	Ordering in Restaurant						Giving Directions					
	Control			Treatment			Control			Treatment		
	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change
Mean	6.33	6.00	-0.33	4.86	5.29	0.43	6.33	4.67	-1.67	4.00	5.43	1.43
SD	0.58	1.00	1.15	1.21	1.60	2.23	0.58	0.58	0.58	1.15	0.98	1.72
n	3	3		7	7		3	3		7	7	
	Structure and Vocabulary						Comprehensibility					
Mean	6.33	5.00	-1.33	4.57	5.14	0.57	6.33	5.67	-0.67	4.29	5.57	1.29
SD	0.58	1.00	0.58	0.98	0.69	1.27	0.58	0.58	1.15	1.11	1.51	2.36
n	3	3		7	7		3	3		7	7	

Length of time in program. The second pattern involved the length of time participants had been in the intensive English program prior to the session that was studied. Participants in the study were in the Basic 2 proficiency level, or second of six levels, of the English program. Many students in this level have recently arrived from their home countries and have not studied in an intensive program nor in the United States. Some participants had studied previously in the program, either completing the first level of instruction or unsuccessfully attempting the second level. A few had completed both the first and second levels previously. The mean scores of students in the

treatment groups with previous instruction in the program and the mean scores of students who were in their first session were compared.

Predictably, the students who had no previous time in the program improved their scores the most. This often happens because of the initial learning spurt when constantly exposed to a new language. Scores on the restaurant task rose 1.07 points (*SD* = 1.44) of 8.00 points for those with no previous time in the intensive English program, while declining 0.27 points (*SD* = 1.62) for those who had attended classes previously. On the giving directions task, those who had no previous intensive English instruction gained 1.00 point (*SD* = 1.62), while those who had previous time in the program increased 0.64 points (*SD* = 1.75). On the structure/vocabulary criterion, those with no previous time in the program improved 0.50 points (*SD* = 1.09) and those who had previous time in the program remained the same (*SD* = 1.00). Last, on the comprehensibility criterion, those with no previous time in the program improved their scores by 1.57 points (*SD* = 1.60), while those with previous time in the program gained only 0.36 points (*SD* = 1.29).

In addition to analyzing those in the treatment groups who had previous instruction in the intensive English program to those who had not, it is useful to compare their scores to the control group. Because there was only one student in the control group who had previously attended classes in the program, a comparison of this group was not possible. However, looking at students who had not previously studied in an intensive English program, those in the treatment group outperformed those in the control group in all areas. On the restaurant task, the treatment group (n = 14) improved scores by 1.07 points (*SD* = 1.44) while the control group (n = 8) remained constant (*SD* = 2.56). On the giving directions task, the treatment group scores rose by 1.00 point (*SD* = 1.62), and the

control group scores rose by 0.13 points (*SD* = 2.03). Structure and vocabulary criterion scores rose 0.50 points (*SD* = 1.09) for the treatment group, while remaining constant for the control group (*SD* = 1.69). Comprehensibility criterion scores increased 1.57 points (*SD* = 1.60) for the treatment group, and 0.13 points (*SD* = 2.23) for the control group.

Table 16

Comparison of Pre, Post Scores on Each Task on 8-Point Scale, by Previous Time in Program

		Ordering in Restaurant						Giving Directions					
		Control			Treatment			Control			Treatment		
		Pre	Post	Change	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change
Previous time in program	Mean	2.00	3.00	1.00	4.45	4.18	-0.27	2.00	3.00	1.00	3.64	4.27	0.64
	SD				1.44	1.54	1.62				1.43	1.62	1.75
	n	1	1		11	11		1	1		11	11	
No previous time in program	*M*	5.00	5.00	0.00	4.00	5.07	1.07	4.13	4.25	0.13	3.36	4.36	1.00
	SD	1.77	1.60	2.56	1.47	1.21	1.44	1.96	1.28	2.03	1.08	1.34	1.62
	n	8	8		14	14		8	8		14	14	

Table 17

Comparison of Pre, Post Scores on Each Criterion on 8-Point Scale, by Previous Time in Program

		Structure and Vocabulary						Comprehensibility					
		Control			Treatment			Control			Treatment		
		Pre	Post	Change	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change
Previous time in program	Mean	2.00	3.00	1.00	4.45	4.45	0.00	2.00	2.00	0.00	3.64	4.00	0.36
	SD				0.93	1.13	1.00				1.63	1.18	1.29
	n	1	1		11	11		1	1		11	11	
No previous time in program	*M*	5.00	5.00	0.00	4.07	4.57	0.50	4.50	4.63	0.13	3.29	4.86	1.57
	SD	1.77	1.60	2.56	1.00	0.85	1.09	1.69	1.85	2.23	0.99	1.56	1.60
	n	8	8		14	14		8	8		14	14	

Summary

The quantitative data showed that accountability groups helped students retain in their course of study and improved attendance, but had little or no effect on their overall academic success in terms of grades or curricular objectives. Qualitative data showed that it was a new concept to think of a classmate as a learning support, but students have an emergent sense of their autonomous learning and of the value of their own knowledge and that of their partners. They are learning their own and each other's strengths. Themes that emerged from the qualitative data show some inconsistencies between class members' experiences. While some sought out opportunities to connect with others outside of class with whom they could speak English, others stayed close to a social group of compatriots and spoke their native language. Goal setting was helpful to those who understood the concept and participated; others did not fully embrace the methodology as it was a concept that had not been previously explored. Overall, students saw their partners as knowledgeable others and as helpful supports. There was also an expressed preference for partners with a different native language, as they were viewed as a better support system for language learning. However, tests of curricular objectives did not show a quantifiable difference between those who had partners who spoke the same language and those who did not.

Chapter 5 - Discussion of Results

This chapter presents the discussion of findings, limitations of the study, implications for practice, recommendations for future direction, and conclusion.

In the accountability group innovation, students worked with partners in the classroom on tasks that promoted noticing language features, sharing information, and explaining language constructs and vocabulary. Partners also lent each other support, help, encouragement, and assistance needed to progress academically. Students set individual goals, which helped partners to hold each other accountable in order to create new affordances in their own environment to make English their own.

Findings

The effect that accountability groups could have on student success in language acquisition and on creation of their own opportunities to use language was explored through mixed methods research focusing on four research questions; each question will be responded to in turn.

1. To what extent do accountability groups help students succeed in a university intensive English program in terms of curricular objectives, grades, and attendance?

Pre- and post-tests of skills taught in the course, as well as archival records of grades and participant attendance help respond to this question.

Data varied from the pre and post tests, reflecting the fact that each student is on his/her own learning journey. This indicates that students are working in their own Zone of Proximal Development, with appropriate scaffolding and support. Students need encouragement and aid in taking autonomy for their own learning, such as the goal-writing that students did in the accountability group. As will be discussed later in this

chapter, written tests have the inherent limitation that they are not focused on the context where language is used. Comparisons of overall grade averages and attendance of the control group to the treatment groups showed a marked difference, with the treatment group having better grade averages and better attendance. Because students in the accountability groups were more engaged in their learning, they had better attendance and thus, better grade averages, indicating progress in their language acquisition, even though tests of curricular objectives were inconsistent.

The most striking results of the accountability group innovation were its contribution to the overall attendance and overall grade average of the students. The original design of the innovation was that each student would have a partner and then partners would be paired with another partner-group. This design was utilized in case a student withdrew or became disengaged with the class, the partner could continue on with his/her group. However, it quickly became apparent that this was an unnecessary feature. In both treatment groups, no students failed any of their courses for missed class time. All classes total about 132 hours, or 7920 minutes, of class time. Treatment Group A students missed a mean of 190.7 minutes (2.4%) across all their classes and Treatment Group B missed a mean of 87.7 minutes (1.1%) across all classes. Conversely, the control group students missed a mean of 1152.4 minutes (14.6%) across the same time span. This gap demonstrates the engagement and social support that Johnson and Johnson (1999) and Slavin (1995) postulate create high academic achievement, improved attendance, and personalization of the learning experience. Accountability group members felt a sense of responsibility to their partner to participate and be present. Because students in the treatment groups were present in classes and were more engaged

in their own learning, students in the accountability groups had higher grades, as measured by the mean grade average of all courses, not just the course with the innovation.

Chinese students greatly benefited from the accountability group innovation in terms of attendance, a measure of engagement, and thus, because of the policy mandating a failure for excessive minutes missed, in overall grade average. In the treatment groups, which had similar proportions of Chinese students to the control group, no students had averages below 40% and no students failed because of excessive absences. In the control group, five students had overall grade averages below 40%. All of these students were Chinese. Four Chinese students failed because of excessive absences. However, there was an inconsistent relationship between the grade averages and attendance and the tests of curricular objectives: Chinese students in the treatment group showed less improvement from the pre to post test of curricular objectives than Chinese in the control group. This comparison shows that the tests of curricular objectives do not show the entire situation, mostly because the Chinese students who were failing due to missed class time probably did not take the posttest.

Japanese students showed a marked difference between the treatment groups and the control group on the pre and post tests. Japanese in the treatment groups improved in all areas—gap fill and construct tasks, structure, vocabulary and comprehensibility; Japanese students in the control group actually decreased in all areas, as shown in Table 15. This indicates that Japanese students need the social support to transition to new ways of learning in an immersive environment of a new culture. Interventions such as

the accountability group, which provide scaffolding in the form of partner work, support networks, and goal setting, help them to maintain and improve their language skills.

Similarly, students with no previous time in the intensive English program need the support offered by accountability groups. The students who were in their first session in the program improved more from pre-test to posttest than those who had studied previously in the program. As Charles Duhigg (2012) explains, people change what they do and how they do it when they have a change of environment. All the old cues and triggers of previous behaviors are changed; therefore, this is the best time to change habits and ways of doing things. Barab and Roth (2006) noted that social dynamics and a person's culture and history cause them to accept or reject affordances. The accountability group supports students in changing the way they interact with language and in learning a new way of knowing about language.

2. To what extent do accountability groups help students gain confidence in using English? Data collected and analyzed to measure students' confidence in using English were drawn from a survey given to study participants before the study and at the end of the study. Survey items were arranged to elicit information about three constructs and were found to be highly reliable: students' perceptions of autonomy (7 items, α = 0.82), affective factors (7 items, α = 0.95), and frequency of language use (4 items, α = 0.95). Qualitative data were also gathered from interviews with participants at the end of the study.

Students in the accountability groups reported using English more frequently than those in the control group. Although one could hypothesize that using English more frequently would help lower affective factors, such as nervousness when speaking

English, the surveys did not support this view. Students in both the control and treatment groups were more comfortable, but those in the control group reported increasing their comfort with English use more than those in the treatment groups. Perhaps as students used English more and in a wider variety of situations, they noticed their nervousness more as they took risks and overcame cultural barriers to communication.

Students in the treatment groups did express more confidence with their language usage, giving details of being able to ask questions and converse in common situations, such as the grocery store. Some of them clearly delineated their journey of language acquisition and expressed a new confidence in using language as well as controlling their own learning experience. Because the quantitative data from the survey were divergent from the qualitative data from the interviews, understanding the students' confidence levels requires further study through qualitative methods, such as interviewing the participants to explore why their survey responses were chosen.

3. How do students in an accountability group create experiences where English language is used? To explore this research question, qualitative data were collected through interviews and verified through text messages and exit slips.

The qualitative data showed that students in an accountability group had various ways of creating their own experiences to use English. To move toward self-control of learning, students wrote goals each week about where, when, and with whom they would use structures learned in class. About half of the students said that they completed the goal because they had written it down, while others said it helped them think about using English. In fact, the student attitudinal survey showed that students in the treatment groups had more gains in their perceived learning autonomy than the control group.

This was further demonstrated by the students who reported creating affordance networks outside the classroom to practice English. Some created experiences with their classmates, getting together to talk about joint interests, to cook, to watch movies, and to hang out. Others sought out international groups and people who did not speak their native language to expand their effectivity sets, the actual interactions where they used English (Barab & Roth, 2006). Overall, students found that putting themselves in proximity of those who did not speak their native language increased their chances to use English outside of class. The post surveys showed that the students in the accountability groups, on average, used English more frequently than those who were in the control group.

Each student had a unique story to tell of the affordances that he/she created, reflecting the diversity of learners within the same proficiency level, the same English language program, and the same native culture. Each one also had an emergent sense of autonomy, reflecting his or her own life experiences up to this point. These varied accounts bolster the sense that accountability groups are a way to help students establish autonomy in their language learning and to work in their own and others' Zone of Proximal Development. They acquire English at a pace that is right for them, in contexts that are natural to them. Even though tests of curricular objectives did not capture this type of improvement, it is essential to language acquisition.

4. How do accountability groups influence the students' perceptions of friends and classmates as support systems in language learning?

After working with the same partner for seven weeks, students viewed their partner as a helpful resource. Partners took on the role of knowledgeable other, teaching

each other at a level that they understood. They talked of learning vocabulary as they worked through assignments together. They saw that they were valuable to their partners also. Through this conversation and interchange of information, both learner and knowledgeable other change (Tappan, 1998). When teachers step out of the way and let students work through tasks on their own, they work in their Zones of Proximal Development and begin to take the control of their own learning process and assist that of their partner. This transfer of regulation from teacher to other to self is an on-going process and is essential to learner development (Edwards, 2001; Lantolf et. al., 2014).

Overall, students expressed a desire to work with partners whose native language was not the same as their own. They explained that they did not learn as much English when their partner shared their language, as they were drawn to speak their first language which was considered much faster and expedient for communication. Some students sought out places and situations where they would be in the proximity of speakers of other languages, which helped take away the tendency to speak in their own language. The qualitative data was divergent from the quantitative. Within the quantitative data, inconsistencies were noted: the fill-in-the-gap task and comprehensibility criterion improved more for those with partners of the same language, but the constructive task and vocabulary/structure criterion improved more for those with partners with a different first language. Since the quantitative testing in this research has certain limitations, discussed below, it can be assumed that the students benefited from partners with a different language as they created experiences and support for language learning.

Limitations of the Study

As with every study, this research had limitations that should be noted. Analyzing the quantitative data in any educational or social science research is problematic. First, the population studied was quite small: each group contained approximately 15 students. When these sections were separated into sub-groups for analysis, some sub-groups, such as Arab students or students with previous time in the program, became very small, sometimes a little as one. Trends cannot be discovered, although sometimes patterns can be noted.

The quantitative data collection instruments are also limiting. Surveys, such as the one used in this research, can be highly reliable, but only measure participants' perceptions on a given day. Because the survey was given as a pre and post survey, participants' perceptions may change, not because of a change in the quality being measured, but because they have a fuller picture of what the behavior can entail or because they have different expectations of themselves. Likewise, a test of skills learned does not capture any change in the ability of that student to apply that skill in the "real world." In this research, the measurement tool was text-based, and students had to write sentences of what they would say in specific situations. This is a different task than speaking to another person in an actual situation.

The last limitation to be mentioned here is that there were no interviews with the control group. The students in the accountability group were able to communicate what happened in their lives in regards to language acquisition during the study with their simple yet revealing language. However, it is unknown if students in the control group would have made similar observations and statements.

Implications for Practice

While results were somewhat mixed, reflecting learners' emergent sense of autonomy, the accountability group innovation does show that practitioners need to be cognizant of the need to facilitate interactions between students that build relationships. It has been a generally accepted practice in this intensive English program to reassign partners each class period, so that students can work with a variety of others. However, the accountability groups showed that students may improve language acquisition with one assigned partner, as this helps them to form a relationship, to learn to work together in a style that fits each of them, and to recognize the partner as a knowledgeable other who can help them in their Zones of Proximal Development.

Many teachers know that working with partners and small groups are effective methods in the language-learning classroom because they provide practice using targeted language structures. However, the literature shows that specific types of partner activities are more effective than others. The accountability groups used activities that promote language acquisition—collaborative activities that promote noticing of language (Pica, 1994; Watanabe & Swain, 2007), exchange of information (Foster, 1993), and language-related exchanges (Littlejohn, 1983). Practitioners should carefully plan partner activities to allow students to negotiate meaning, model and scaffold for each other, and allow pairs to organize knowledge in a way that is accessible to them. Assignments such as the "two people, one paper" tasks encourage students to work together in a method that fits their regulatory style and to discuss the vocabulary and language structures that they are using.

Last, practitioners should be facilitating ways to connect learning in the classroom to learning outside the classroom. Writing and sharing goals helps students do this in an

autonomous way, allowing them to think about their own context and how it can best be shaped to further their own language learning. Students in the accountability groups identified a specific time, place, and person with whom they could use English: at parties, with host families, with people in the community. Sharing these small goals gave some of the students the motivation to practice English in their own context. In addition, each student brings a socio-cultural history that effects his/her perception of the world that causes affordances to use language to be accepted or rejected (Barab & Roth, 2006). Changing these perceptions is sometimes key to language learning, and accountability groups provide the support needed. Mahwah, an Arab lady, was used to letting her male family member talk for her at the grocery store. Through the accountability group innovation, she gained the support and confidence to ask questions and be more autonomous in her interactions in the grocery store. This shows that the assignments that practitioners make in class may not be where the students are in their language learning journey; what students are doing and what teachers expect they are doing are actually two different things. If practitioners truly believe in a Zone of Proximal Development, they must give students the autonomy to create their own affordance networks that allow them to move forward in their language learning in the way that most suits their own personal story.

Recommendations for Future Research

This study encompassed two classes using the accountability group methods and one class in a control group. All participants were lower proficiency level, and many had difficulty understanding the goal-setting concept due to language comprehension ability. Future research could consider if the results found here are more valid at a larger scale

and at higher proficiency levels. Because of the marked improvement in attendance, it would be beneficial to explore partnerships that last longer than one 8-week session to see if these results could be sustained over a longer period of time.

It would also be of interest to see what correlation students make between their grades in their classes and their perceptions of their language learning. Since educational institutions regularly use class grades as a measure of proficiency and achievement, it makes the researcher wonder if those grades are viewed in the same light by the students. Related to this is whether students are motivated or de-motivated in their language acquisition by the grades that they receive—does making a better grade inspire them to use English in their daily lives?

Conclusion

The Accountability Group Innovation aimed to apply the ecological perspective to a language classroom, allowing students to take charge of their own learning and create affordances to expand their interactions with the language. The activities were supported by their peers and learning took place between the partners. This made their learning much more engaging than a worksheet to which everyone shared the answers on a WhatsApp group.

When this research began, I was worried about the accountability that students traditionally have to the teacher: Would students actually complete the assignments that they made for themselves through their goals, or would they just say that they had done something? Could peers actually teach other, or did they need me, the teacher, to oversee and correct all their writing and utterances?

One day in about the middle of the term, I was beginning the part of class when the students would write their goals for the week. To prompt them about applying the structures and skills learned in class, I asked, "What have you learned in this class?" I expected answers such as introducing myself, describing clothing, or ordering coffee in a coffee shop. Moeen looked up with a light in his eyes, the kind that warms any teacher's heart—the one that says "I've got this!" I called on him, happy that someone had paid attention in class. His reply was "We've learned how to connect with other people in English."

References

Alon, I., & McIntyre, J. R. (Eds.). (2005). *Business and management education in China: Transition, pedagogy and training*. Singapore: World Scientific.

Barab, S. A., & Plucker, J. A. (2002). Smart people or smart contexts? Cognition, ability, and talent development in an age of situated approaches to knowing and learning. *Educational Psychologist*, *37*(3), 165-182.

Barab, S. A., & Roth, W. M. (2006). Curriculum-based ecosystems: Supporting knowing from an ecological perspective. *Educational Researcher*, *35*(5), 3-13. http://dx.doi.org/ 10.3102/0013189X035005003

Barab, S. A., Squire, K., & Barnett, M. (1999, May). *From teachers' fixed curricular objectives toward students' emergent practices.* Paper presented at the Annual Meeting of the American Educational Research Association. Montreal, Canada.

Barakat, H. (1993). *The Arab world: Society, culture, and state*. Berkley, CA: University of California Press.

Beach, L. R. (1974). Self-directed student groups and college learning. *Higher Education*, *3*(2), 187-200. http://dx.doi.org/10.1007/BF00143791

De Guerrero, M., & Villamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *The Modern Language Journal*, *84*(1), 51-68. http://dx.doi.org/10.1111/0026-7902.00052

Duhigg, C. (2012). *The power of habit: Why we do what we do in life and business.* New York: Random House.

Edwards, A. (2001). Researching pedagogy: A sociocultural agenda. *Pedagogy, Culture and Society*, *9*(2), 161-186. http://dx.doi.org/10.1080/14681360100200111

Fan, Y. (2000). A classification of Chinese culture. *Cross Cultural Management: An International Journal*, *7*(2), 3-10. http://dx.doi.org/10.1108/13527600010797057

Flaitz, J., & Eckstein, L. K. (2003). *Understanding your international students: An educational, cultural, and linguistic guide*. Ann Arbor, MI: University of Michigan Press.

Garcia, P., & Asención, Y. (2001). Interlanguage development of Spanish learners: Comprehension, production, and interaction. *Canadian Modern Language Review*, *57*(3), 377-401. http://dx.doi.org/10.3138/cmlr.57.3.377

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta- analysis of effects and processes. *Advances in*

Experimental Social Psychology, 38, 69-119. http://dx.doi.org/10.1016/S0065-2601(06)38002-1

Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly*, 55-77. http://dx.doi.org/10.1002/j.1545-7249.2008.tb00207.x

Hammad, A., Kysia, R., Rabah, R., Hassoun, R., & Connelly, M. (1999). *Guide to Arab culture: Health care delivery to the Arab American community*. Dearborn, MI: Arab Community Center for Economic and Social Services (ACCESS) Community Health and Research Center.

He, X., & Ellis, R. (1999). Modified output and the acquisition of word meanings. In *Learning a second language through interaction* (pp. 115-132). Philadelphia, PA: John Benjamins Publishing.

Institute for International Education. (2016). Open doors data. Retrieved December 2016, from http://www.iie.org/Research-and-Publications/Open-Doors/Data

Ivankova, N. V. (2015). *Mixed methods applications in action research*. Thousand Oaks, CA: Sage.

Johnson D., & Johnson, R. (1999). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Needham Heights, MA: Allyn and Bacon.

Klingner, J. K., & Vaughn, S. (2000). The helping behaviors of fifth graders while using collaborative strategic reading during ESL content classes. *TESOL Quarterly*, *34*(1), 69-98. http://dx.doi.org/10.2307/3588097

Lantolf, J., Thorne, S., & Poehner, M. (2014). Sociocultural theory and second language development. In Van Patten, B. & Williams, J. (Eds.), *Theories in second language acquisition: An introduction* (pages 207-226). New York: Routledge.

Latham-Koenig, C., Oxenden, C., & Seligson, P. (2013). *American English file book 1* (2nd ed.). Oxford, U.K.: Oxford University Press.

Littlejohn, A. P. (1983). Increasing learner involvement in course management. *TESOL Quarterly*, *17*(4), 595-608. http://dx.doi.org/10.2307/3586615

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705-717.

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*, *15*(5), 265-268. http://dx.doi.org/10.1111/j.1467-8721.2006.00449.x

McLeod, S. (2016). Bandura - Social learning theory. Retrieved March 03, 2016, from http://www.simplypsychology.org/bandura.html

Ohta, A. S. (2001). *Second language acquisition processes in the classroom: Learning Japanese*. New York: Routledge.

Onwuegbuzie, A. J., & Leech, N. L. (2006). Linking research questions to mixed methods data analysis procedures. *The Qualitative Report*, *11*(3), 474-498. Retrieved from http://nsuworks.nova.edu/tqr/vol11/iss3/3/

Pica, T. (1991). Classroom interaction, participants, & comprehension: Redefining relationships. *System, 19*(4)*,* 437-452. http://dx.doi.org/10.1016/0346-251X(91)90024-J

Pica, T. (1994). Research on negotiation: What does it reveal about second- language learning conditions, processes, and outcomes? *Language Learning, 44*(3), 493-527. http://dx.doi.org/10.1111/j.1467-1770.1994.tb01115.x

Schippers, M. C., Scheepers, A. W., & Peterson, J. B. (2015). A scalable goal-setting intervention closes both the gender and ethnic minority achievement gap. *Palgrave Communications*. Retrieved from http://www.palgrave-journals.com/articles/palcomms201514

Shteynberg, G., & Galinsky, A. D. (2011). Implicit coordination: Sharing goals with similar others intensifies goal pursuit. *Journal of Experimental Social Psychology*, *47*(6), 1291-1294. http://dx.doi.org/10.1177/1745691615589104

Slavin, R. (1995). *Cooperative learning* (2nd ed.) Needham Heights, MA: Allyn and Bacon.

Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research*, *5*(1), 29-53.

Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). Peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics*, *22*, 171-185.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, *82*(3), 320-337. http://dx.doi.org/10.1111/j.1540-4781.1998.tb01209.x

Tappan, M. B. (1998). Sociocultural psychology and caring pedagogy: Exploring Vygotsky's "Hidden Curriculum." *Educational Psychologist*, *33*(1), 23-33. http://dx.doi.org/10.1207/s15326985ep3301_2

Watanabe, Y. & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult

ESL learners. *Language Teaching Research*, *11*(2), 121-142. http://dx.doi.org/10.1177/136216880607074599

Wunderle, W. D. (2006). *Through the lens of cultural awareness: A primer for United States Armed Forces deploying in Arab and Middle Eastern countries*. Retrieved from: http://usacac.army.mil/cac2/cgsc/carl/download/csipubs/wunderle.pdf

Zaharna, R. S. (1995). Understanding cultural preferences of Arab communication patterns. *Public Relations Review*, *21*(3), 241-255. https://doi.org/10.1016/0363-8111(95)90024-1

APPENDIX A

VIGNETTES

Standard Methodology. Most units are centered around everyday life situations where a student would need to use English, such as the Practical English unit "Getting Lost."

In this particular unit, the methodology suggested by the text, and used by many teachers, is to watch three short videos that show a character in this exact situation, modeling the language for the students. The first video, which is a continuation of a storyline that runs throughout the course, sets the scene of Jenny arranging to meet her friend Rob at the Tate Modern Museum in London. He shows concern that she might not know the way, but she responds confidently, "I have a map." Then students learn some limited vocabulary needed for understanding directions (e.g., the corner, turn left, go past a certain building). Students then watch the second video in which Jenny, predictably, gets lost and asks for directions. Students then watch the video again to fill in the blanks on a transcript of the conversation. After practicing saying the conversation individually, trying to mimic the pronunciation and rhythm, students practice the dialogue with a partner. This scaffolds the next task, which is to role-play the dialogue with their partner, asking for directions to a building indicated on a map shown in the textbook. The final video shows Jenny and Rob meeting and deciding what activity they would like to do for the afternoon. This video is designed to recycle language that has been used in previous lessons and is followed by eight true/false comprehension questions. Students are then asked to watch the video again and explain why the false sentences are false. Attention is called to typical social English phrases shown in the three videos. Small clips of just these expressions are condensed into one short video so that students can hear them

again. The lesson concludes with a three-part checklist for students: "Can you…ask for and understand directions? give simple directions? ask someone to do something in a polite way?" (Latham-Koenig, Oxenden, Seligson, 2013, p. 59)

Accountability Group Methodology. Taking the same example from the unit "Getting Lost" in the Standard Methodology, the Accountability Group Methodology will introduce the following interactions. After being reminded of language for making suggestions, which was taught in a previous lesson, students will role-play with their group a conversation between Jenny and Rob, discussing what they could do with a free morning in London. The class will then watch the video clip from the text to see if their conversation matches what happens in the video; the twist is that the conversation between Jenny and Rob is interrupted when Rob's boss calls him in to work. Then students will go over the vocabulary for giving directions in the textbook with their partners, and complete a jigsaw cloze activity designed by the teacher. In this activity, directions are given to a place on a map. One student has all the directions written out, but the other has missing words. The first student reads the directions to the second student who fills in the blanks, calling attention to the target vocabulary and how it is used in context. Next, they watch the second video clip in which Jenny gets lost and asks for directions. During the second viewing, they fill in the cloze activity in the textbook. Next, they read over the conversation with their partner. The next activity will begin in class, but be finished together as homework. Partners will choose a location on campus and write instructions how to get to the location, without naming the exact location. As explained in the literature review, working together to write the instructions will allow partners to help fill gaps and reinforce each other's language abilities. During the final

five minutes of class, the third video clip will be shown and any directions that are unfinished will be completed for homework. The following day, each partner group will give the written instructions to the other pair in their accountability group, who will leave the classroom and follow the instructions. They then report the ending location of the directions. An additional homework assignment is to go with their partner to a location like the mall, a busy shopping street, or a public location where people can be found. They are to approach an American and politely ask directions to a location nearby. Doing this activity together will possibly give them confidence to approach a stranger, and they will also reinforce their English learning as they follow the directions and review what the speaker has said. After completing this lesson, students will be able to self-evaluate the outcome questions about asking for and understanding directions, giving simple directions, and asking someone to do something in a polite way.

APPENDIX B

INTERVENTION ACTIVITIES

Figure 1. Partner/group communicative activities, based on course standardized curriculum.

Unit	In Class Activities	Homework Activities
Getting Acquainted	Interview your partner. Introduce your partner to your group.	Connect with your partner outside of class—by phone, text, or face-to-face. Find out some study tips that they use.
Unit 1: Arriving/Checking in to a hotel	Scaffolded role play; with partner, write a dialogue to share with group	Create a tip sheet for newcomers to ASU from your country, listing ten pointers about cultural things they need to know.
Unit 2: In a Coffee Shop	Scaffolded role play	In a coffee shop, order something.
Midterm Presentation	Make partner presentation about interviews in class	Interview Americans.
Unit 3: In a Clothes Store	Play a game with partner about changes in clothing.	With your partner, plan what you will wear to class. Email the teacher.
Unit 4: Getting Lost	Using a map, explain how to arrive at different locations.	With your partner, write instructions how to go from the classroom to a location on campus. With your partner, ask Americans for directions.
Unit 5: In a restaurant	Practice telling your partner good news & responding appropriately. Practice turning down an invitation. Class role play: eating in a restaurant with your group.	Go to a restaurant with your partner and order food. Create a video with restaurant etiquette tips.
Final Presentation	Present video in class.	Make a video tour of your house or apartment.

APPENDIX C

DATA COLLECTION TOOLS

Student Success Survey
Pre and Post Instrument

1. Where are you from (city & country)?

2. What is your birthday?

Part 1

1. I speak English with classmates outside of class.
   Always Often Sometimes Not much Never

2. I study English with friends and classmates.
   Always Often Sometimes Not much Never

3. I do my homework by myself.
   Always Often Sometimes Not much Never

4. I practice using new words with my friends and classmates.
   Always Often Sometimes Not much Never

5. I review the class lessons with a classmate.
   Always Often Sometimes Not much Never

6. I ask a friend or classmate for help with American life (washing clothes, shopping, cooking, registration, etc.)
   Always Often Sometimes Not much Never

Part 2

7. I start conversations in English.
   Always Often Sometimes Not much Never

8. When I can't think of a word in English, I use gestures or a different word.
   Always Often Sometimes Not much Never

9. I try to find many ways to use English.
   Always Often Sometimes Not much Never

10. I look for people I can talk to in English.
    Always Often Sometimes Not much Never

11. I am afraid of making a mistake when I speak English.
    Always Often Sometimes Not much Never

12. I ask for help from English speakers.
    Always Often Sometimes Not much Never

13. I ask questions in English.
    Always Often Sometimes Not much Never

Part 3

14. I find it hard to look at the other person when I am speaking English.
    Strongly agree Agree Undecided Disagree Strongly Disagree

15. I feel very nervous when waiting to speak in English.
    Strongly agree Agree Undecided Disagree Strongly Disagree

16. I enjoy speaking English.
    Strongly agree Agree Undecided Disagree Strongly Disagree

17. I get nervous and confused when speaking English.
    Strongly agree Agree Undecided Disagree Strongly Disagree

18. I get nervous when the teacher asks me to speak in English when I have prepared before.
    Strongly agree Agree Undecided Disagree Strongly Disagree

19. I want to speak less because I feel shy while speaking English.
    Strongly agree Agree Undecided Disagree Strongly Disagree

20. I have no fear of speaking English.
    Strongly agree Agree Undecided Disagree Strongly Disagree

Part 4

21. I use English to order a drink in a coffee shop.
    more than 6 times a week 4-5 times a week
    2-3 times a week 0-1 times a week

22. I use English when shopping for clothes.
    more than 6 times a month 4-5 times a month
    2-3 times a month 0-1 times a month

23. I use English to order food in a restaurant
    more than 6 times a month 4-5 times a month
    2-3 times a month 0-1 times a month

24. I use English to ask or give directions to places (stores, bathroom, buildings on campus, etc.)

more than 6 times a month 4-5 times a month
2-3 times a month 0-1 times a month

Exit Slip

Your birthday _____

Answer these questions and give to the teacher before you leave class today.

1. What did you learn about today? Is it useful for you? Why or why not?

2. What is your goal for this week?

3. How will your partner or group help reach the goal?

4. What did your partner do to help you this past week? What did you do to help your partner?

Interview Questions

What do you like best about working with your partner?

What do you like least about working with your partner?

Give one example of how your partner has helped you.

Give one example of how you have helped your partner.

What is one change you and your partner can make that would improve your language learning?

In my group, I am good at….

I feel my group is ….

What English communication goal did you set? Did you achieve the goal? Who helped you achieve the goal?

Text Messaging

Below are sample questions that will be sent to each consenting participant once or twice a week. They will be sent one question, which should prompt a conversation of 2-4 exchanges.

Sample exchange:

Researcher: Have you used English today outside of class?
Student: Not yet.
Researcher: What do you think you will try to do to speak English?
Student: Maybe asking for help at a store.
Researcher: Check in with (your partner) and see what he is doing. Ask him to give you ideas on how to use English.
Student: Okay

Have you used English today?

When was the last time you spoke to _____ (partner's name)?

Did you talk to your partner about school work?

How have you helped _____ this week?

What can you do to help _____?

How is _____ doing today? Have you talked to him/her?

APPENDIX D

INSTITUTIONAL REVIEW BOARD APPROVAL

ASU Knowledge Enterprise Development

EXEMPTION GRANTED

Kathleen Puckett
Division of Teacher Preparation - Polytechnic
480/727-5206
Kathleen.Puckett@asu.edu

Dear Kathleen Puckett:

On 3/29/2016 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Accountability Groups to Improve Student Success in an Intensive English Program
Investigator:	Kathleen Puckett
IRB ID:	STUDY00003776
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	• AG_Lippincott_ConsentFormsAB_v5a.pdf, Category: Consent Form; • Student Recruitment Letter, Category: Recruitment Materials; • AG_HRP-503a_Mar2016_v8a.docx, Category: IRB Protocol; • DataCollectionTools, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions);

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (1) Educational settings, (2) Tests, surveys, interviews, or observation on 3/29/2016.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,

IRB Administrator

cc: Dianna Lippincott
Dianna Lippincott