8-9-2008

Tutoring instrument flight: patterns of instructor and student communication

Adnan Okdeh

TUTORING INSTRUMENT FLIGHT: PATTERNS OF INSTRUCTOR AND STUDENT COMMUNICATION

By

Adnan Okdeh

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Experimental Psychology
in the Department of Psychology

Mississippi State, Mississippi

August 2008

TUTORING INSTRUMENT FLIGHT: PATTERNS OF INSTRUCTOR AND STUDENT COMMUNICATION

By

Adnan Okdeh

Approved:

Gary L. Bradshaw
Professor of Psychology
(Director of Thesis)

Deborah K. Eakin
Assistant Professor of Psychology
(Committee Member)

Carrick C. Williams
Assistant Professor of Psychology
(Committee Member)

Kevin J. Armstrong
Associate Professor of Psychology
Graduate Coordinator for the
Department of Psychology

Gary L. Myers
Dean of the college of
Arts and Sciences

Name: Adnan Okdeh

Date of Degree: August 9, 2008

Institution: Mississippi State University

Major Field: Psychology (Experimental)

Major Professor: Dr. Gary L. Bradshaw

Title of Study: TUTORING INSTRUMENT FLIGHT: PATTERNS OF INSTRUCTOR AND STUDENT COMMUNICATION

Pages in Study: 109

Candidate for Degree of Master of Science

Individual tutoring has been successful in facilitating learning in domains such as LISP, physics, and algebra. These tasks are static in that problems do not change while the student is trying to solve them. Dynamic tasks such as flying, where the problem changes spontaneously over time, represent different challenges for tutors.

To understand tutoring in dynamic tasks, we conducted a field observation of students being given messages by a flight instructor. Five low flight time student pilots were asked to perform nine instrument flight tasks while being tutored by an instructor pilot in both a virtual simulator flight and in a real airplane flight. The data from our study were compared to two prominent models of one-on-one tutoring. Only a small portion of the utterances made by the tutor or by the student matched previous accounts, suggesting that a new approach is needed to address tutoring during instrument flight instruction.

DEDICATION

I would like to dedicate this research to my parents Subhi and Layla, my sister Rana, my brother-in-law Issam, and my niece “little” Layla.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all the people who assisted me in this project. First and foremost, a sincere thanks to Dr. Gary L. Bradshaw, my thesis director and a committee member, who kindly provided me with continuous feedback and direction during the entire process. I would also like to extend my sincere thanks to Dr. Stephanie Doane who helped me get work pertaining to my thesis research. Expressed appreciation is also due to Mr. Thomas E. Hannigan III for graciously giving his time to being the instructor pilot in this study. I would like to acknowledge the Office of Naval Research for sponsoring this research (Award #: N000140210152). Finally, I would like to thank Dr. Deborah K. Eakin and Dr. Carrick C. Williams for serving on my committee and sharing their feedback with me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

One-on-One Tutoring

The ability to learn quickly and effectively is an important part of every person's success and growth. Teaching plays a major role in learning and our 'teachers' include not only professionals (e.g., school teachers, college faculty, etc.) but also paraprofessionals (e.g., student teachers) and peers as well. Thus, it is no surprise that pedagogy (i.e., how to teach) is an important topic of research in psychology and education.

Researchers in these areas have investigated whether there are certain methods of teaching that are more effective than others, and if so, what makes them so effective. Studies that compared one-on-one tutoring with classroom teaching showed that the average student who learned through one-on-one tutoring had a performance gain of 0.4 to 2.3 standard deviations greater than the average student who learned through a classroom setting (Bloom, 1984; Cohen, Kulik, & Kulik, 1982). Furthermore, Slavin (1987) found evidence that students understood concepts better, learned faster, and were more motivated to learn during one-on-one tutoring than during traditional classroom teaching.

Other evidence suggests that tutors themselves improved their understanding of the material they were assigned to teach. A study by Juel (1991) looked at college athletes with a grade nine reading level who tutored first, second, and special education students with reading difficulties for two semesters. Before the college athletes began tutoring the students, they took an evening class to give them background about literacy acquisition and practical ideas for encouraging reading and writing. Each tutor worked with one child for 45 minutes twice a week. The results showed that after two semesters the mean grade equivalent of the college athletes soared up to a grade level of 13.1 in reading comprehension and vocabulary, significantly surpassing that of the control group. Also, the tutors began enjoying for the first time the experience of teaching reading and writing and some even decided to become teachers.

Anatomy of One-on-One Tutoring

To understand the one-on-one tutoring process, a few early studies examined the dialog between an instructor and a student (Fox, 1991, 1993; McArthur, Stasz, & Zmuidzinas, 1990; Schoenfeld et al.; 1992). In these studies, the tutors were professionals who were highly skilled in their field. One representative study done by McArthur et al. (1990) looked at how three high school teachers tutored three nine and 10^{th} grade high school students in individual tutoring sessions. Each tutor had at least five years of experience teaching high school algebra and had won awards for teaching excellence. They found that the tutors gave the students specific and lengthy instructions on what to do when the student did not know the answer. For example, in one of the dialogs a student is asked how to solve for x in the equation x/3 = b/2 + c/6. When the student did not know where to start, the instructor immediately recommended that the student get rid of the fractions by multiplying every term by the lowest common multiple. The pattern of immediately giving instructions to the student continued throughout the tutoring process. Thus, the expert tutor walked the student through the problem.

To capture the behavior observed in these sessions, McArthur developed a model of the tutoring process based on the idea that tutors employ *scripts* to direct the student's behavior when they make an error. They also examined behaviors other than errors as triggers for a script. Their scripts were referred to as tutorial *microplans* which can be triggered based on six main conditions. Each condition and its description are shown in Table 1.

McArthur et al. (1990) identified the frequency of occurrences for each script in three one and a half hour tutoring sessions that included 54 different problems. The most frequently triggered microplan was the feedback microplan (n = 135). The decision microplan was the second most frequent script given by the instructors to the students (n = 121). The third most triggered script was the new-problem microplan (n = 54). The instructors gave the students a slightly lower frequency of remediation microplans during the tutoring sessions (n = 40). Motivational and redirection microplans were triggered the least during the tutoring sessions (n = 4).

Examining the frequencies of the various microplans reveals several patterns about the tutoring sessions in McArthur's study. The instructor employed the feedback microplan 135 times during the tutoring session but only employed the new-problem microplan 54 times. The high proportion of feedback microplans to new-problem

Table 1

McArthur's Tutoring Model

Type of Microplans		Microplan Descriptions
New-problem microplan	Beginning a new problem	Tutor tells the student what problem will be covered based on (a) the curriculum or (b) an assessment of the student's overall performance on the previous problem.
Feedback microplan	Student generates a response whether correct, incomplete, or incorrect	Tutor provides quick immediate feedback to the student on the status of their answer (e.g., "Yes", "Okay", or "No").
Decision microplan	As students generate responses	Tutor monitors the student's performance and decides what to say next.
Remediation microplan	Student gives an incomplete or incorrect answer during problem solving	Tutor provides specific and lengthy instructions (remediation).
Motivation microplan	Student gives several incomplete or incorrect answer in a row during problem solving	Tutor provides motivation.
Redirection microplan	Student asks a question or makes a comment that is irrelevant to the problem	Tutor directs the student back to solving the problem at hand.

microplans suggests that students typically received feedback on more than one occasion as they worked on each problem. This disproportionate ratio also suggests that throughout the problems, the instructor is either encouraging the student to think along the same line of reasoning or to encourage the student to change their line of reasoning.

Decision microplans were triggered 121 times suggesting that the instructors are closely monitoring the student’s intermediate steps as they worked towards a solution to the problems. The process of the instructors closely monitoring the student’s steps allows for quick feedback and possible remediation to be given to the students. As we will see in later sections of this chapter, quick feedback and remediation are two important aspects of one-on-one tutoring that makes it an effective approach to learning.

During the tutoring sessions the instructors also employed the remediation microplan 40 times compared to 54 new-problem microplans. Thus, the instructors employed an average of less than one remediation for every problem that was solved, suggesting that the students performed fairly well on the task. Motivation microplans and redirection microplans were employed four times each by the instructors to the students which further suggests that the students performed fairly well on the problems given during the tutoring sessions.

Note the style of tutoring did not involve asking probing questions that would encourage students to think about how to answer the questions. Instead the instructor provided explicit instructions about what step needed to be performed next: The style of tutoring was explicit and specific.

A different line of investigation about one-on-one tutoring was initiated by Arthur Graesser and several of his colleagues (Graesser, 1992, 1993a, 1993b; Graesser, & Person, 1994; Person, Graesser, Magliano, & Kreuz, 1994). Their research focused on how tutoring works when the instructor is a paraprofessional or a peer rather than a professional. Graesser and his colleagues were motivated by the Fitz-Gibbon (1977) study that found most tutoring was done by paraprofessionals and peers rather than professionals.

Graesser and his colleagues examined pedagogical techniques, remediation, feedback, and task assessment in tutoring sessions with the goal of identifying the various techniques the instructors used to tutor students. There are two main settings for these studies: The first setting involved three graduate student instructors who had previously received A's in a graduate-level psychology research methods course, but had never tutored in research methods. The second setting involved 10 high school students who had an average nine hours of tutoring experience prior to the study. The goal of the instructors in the first setting was to tutor 27 undergraduate students on six troublesome topics in an undergraduate-level psychology research methods course. The goal of the instructors in the second setting was to provide assistance to 13 students in the seventh grade who where having difficulties in algebra. All tutoring sessions were videotaped for later analyses.

Graesser, Person, and Magliano (1995) sought to identify the most salient techniques (e.g., remediation, feedback, assessment, etc.) used by non-professional instructors to tutor students. They examined the results from their previous studies and devised a model of one-on-one tutoring that leads to deep learning. The model could be used to tutor students effectively by instructors of any skill level (professional, paraprofessional, or peer). They referred to their model as the “tutoring frame” which consists of five sequential steps (Table 2).

Table 2

Graesser’s Tutoring Model

Step	Action
1	Tutor asks an initiating question
2	Student provides an answer
3	Tutor gives the student immediate feedback on the status of their answer (e.g., “Yes”, “Okay”, or “No”) followed by a hint in the form of a question to prompt the student to rethink their answer
4	Tutor improves quality of answer through a collaborative conversation usually taking 5-10 steps (involves giving slightly more specific feedback in the form of questions)
5	Tutor assesses student's understanding of answer and decides what question to ask next

In a typical tutoring session the instructor begins by asking an initiating question which serves to narrow the focus of the material. Chi (1996) notes: “The questions tutors ask to initiate a dialog tend to be consistent with their curriculum, that is, a set of subtopics and example problems that are consistent with the standard materials that have to be covered" (p. 2). Narrowing the focus of the material allows the tutor to cover key concepts in the curriculum so that the tutor can diagnose the student's knowledge in order to provide them with the information they need to understand a particular idea

After receiving an answer from the student the tutor is able to give them feedback. In the feedback process the tutor assesses the student's answer and provides positive feedback for a correct answer, neutral feedback for an incomplete answer, and negative feedback for an incorrect answer. If the student gets the answer correct in the second step the tutor will give a short affirmatory response (e.g., "Yes, that's correct") in the third step. The tutor will then skip the fourth step and ask a question to gauge the student's understanding (step 5).

If the student gives an incomplete answer the tutor will give the student neutral feedback followed by a question to prompt the student to provide more information (e.g., "Okay, is there something else you think is missing here?"). If the student gives an incorrect answer the tutor gives the student negative feedback followed by a question in the form of a hint to prompt the student to rethink/rework their answer. So, for example if the tutor asks the student, "What is numerical value of (23+2) x 2 – 1?", and the student incorrectly responds by saying "25" the tutor might say, "No, 25 is not the correct answer, are you sure you are following the right steps?"

The process of giving the student hints continues in step 4 until the student gets the correct answer. Graesser et al. (1995) reported that step 4 usually takes 5-10 turns for a student to get the correct answer. The instructor eventually gives more specific hints if the student's answer is incorrect or incomplete. Therefore, if the student answers the question incorrectly a second time the tutor might say, "No, that is not correct, what do you have to solve for first?" If the student gets the answer wrong a third time the tutor may then give a more specific hint by saying, "No, think about order of operation, you have to solve the expression in parenthesis first, then what?" Graesser et al. do not provide a concrete rule on how specific the tutor's hints will be if the student's answer is incomplete or incorrect, but a review of the examples provided suggests that specific hints are often provided after the third incorrect attempt.

In step 5, once the student reaches the correct answer the tutor will ask a question to gauge the students understanding of the material (e.g "Do you understand all that?"). The purpose of the question is to provide extra information, if needed, on a certain concept covered by the initial question. Also in this step, the tutor assesses the student's performance to decide if the student needs more practice in the same line of questioning.

Comparing and Contrasting One-on-One Tutoring Models

McArthur et al. (1990) and Graesser et al. (1995) proposed models of tutoring that share several features. The tutors in both models provide students with quick immediate feedback on the status of their answer (e.g., "Yes", "Okay", "No"). Merrill, Reiser, and Landes (1992) also found that tutors were very fast at providing feedback when a student made a mistake. This feedback decreases the likelihood of the student searching down the wrong path of a search space (Anderson, Boyle, & Reiser, 1985). In their 2001 article, Chi, Siler, Jeong, Yamauchi, & Hausmann state:

> Feedback alone can guide students, in the sense of encouraging students to stay on the same track of reasoning or problem solving, in the context of confirmatory feedback, or to change direction or goal, in the context of negative feedback (p. 473).

Thus, the immediacy of the feedback given by the instructor is a key element that makes one-on-one tutoring an effective approach to learning.

Another element that both these representative models of one-on-one tutoring share is they both follow a curriculum. Before the tutoring session begins the instructor lays out a series of questions or topics that they will cover with the student. A third element these models share is an overall assessment of the student's performance at the end of a problem by the instructor in order to decide what the next problem will be. Thus, both approaches have found that tutors work to select problems of an appropriate level of difficulty reminiscent of Vygotsky's (1978) "zone of proximal development." Such problems can be solved with some help or scaffolding by the tutor, but are not so easy that the student would learn nothing from their solution nor so difficult that they could not effectively be solved.

Although the models have several similarities, there are some important differences between the tutoring methods as well. As mentioned previously, both models rely on feedback from the tutor. However, the type of utterance that follows the feedback differs considerably between the two models. In the model proposed by McArthur et al. (1990) the tutor provides specific and lengthy instructions after giving initial feedback to the student. The tutor essentially walks the student through the solution when they get stuck, in an attempt to fill in any gaps in knowledge the student may have. However, in the model proposed by Graesser et al. (1995), the instructor asks the student probing questions after giving them feedback. The act of asking the student a question rather than giving instructions forces the student to work out the answer on their own. The tutor then continues giving hints in the form of questions (that get more specific) until the student gets the correct answer. Thus, Graesser's model encourages deep learning that occurs when the student realizes they have made a mistake and need to correct it on their own (Chi, 1996).

Another important difference between the two models is that in McArthur's model the tutor motivates the student when they give several incomplete or incorrect answers in a row. The tutor using Graesser's model of tutoring does not provide any explicit motivation to the student when the student makes mistakes.

A third difference arises during the assessment phase after a problem has been solved. In McArthur's model, the tutor does not ask the student if they need additional tutoring on the problem or how it was solved. Instead, the tutor asks the student if they need more practice on the same kind of problem. Thus, the students can control to some extent the course of tutoring. In Graesser's model the tutor asks a question to gauge the student's understanding at the end of a problem (e.g., "Did you understand all that?") but does not explicitly ask the student afterwards if they need more practice on the same kind of problem. In this case, the tutor is in control of the path through the curriculum. Thus, in Graesser's model the tutor attempts to specifically diagnose any gaps in knowledge the student might have in order to provide them with the information they need to understand a particular concept. In McArthur's model the tutors do not address the student's specific difficulties, instead the tutors makes an inference as to what those difficulties are and then use their understanding to select the next problem.

Thus, we see important similarities and differences in both models. Similarities include the instructors giving the same type of initial feedback when responding to the student ("Yes", "Okay", "No"). Additionally, the instructors in both models follow a curriculum when tutoring the student and assess the student's performance at the end of a problem. There are, however, important differences between the two models. In McArthur's model after the instructors give initial feedback, they provide the student with specific instructions on how to solve the problem. In contrast, the tutors in Graesser's model give hints in the form of questions to prompt the student to think through the problem on their own. Motivation is given to the students by the instructors in McArthur's model after the student makes several mistakes in a row, however, no motivation is given by the instructors in Graesser's model. Finally, in assessing the student's performance at the end of a problem, the instructors in McArthur's model give the student a choice of what the next problem might be. In Graesser's model the tutors do not give such a choice to the student.

Domains Studied by the Tutoring Models

Both tutoring models have spawned further research in various domains. McArthur’s approach was studied primarily in the domain of algebra (Robyn, Stasz, Ormseth, Lewis, & McArthur, 1991; Schoenfeld et al., 1992). Graesser’s model was studied in domains such as LISP (Merrill, Reiser, Merrill, & Landes, 1995), physics (Gertner, & VanLehn, 2000; VanLehn, 1996), and algebra (Koedinger, Anderson, Hadley, & Mark, 1997). Unfortunately, little work has been done to compare these two models to determine their generalizability and completeness.

One important factor that may limit the generality of both models of tutoring arises because of the similarity of the tasks investigated by McArthur and Graesser. These tasks share three distinct characteristics. First, they involve problem solving of one sort or another. For example, the student might be asked to solve an arithmetic expression such as (23+2) x 2 – 1. Another characteristic these domains have in common is that they are usually governed by an ordered set of principles. So, to solve (23+2) x 2 – 1, the student must follow the precedence rules of operation by first calculating the expression in the parentheses, then multiplying, and finally subtracting. A third characteristic these domains share is that the tasks are performed in a relatively stable environment where the problem does not change over time.

Tasks Performed in Dynamic Domains

Many tasks that are learned through one-on-one tutoring do not necessarily require this kind of problem solving. Learning how to drive, fly, or perform surgery requires *manual control* rather than problem solving to perform. Manual control is, according to Wickens and Hollands (2000, p. 386), "a dynamic systems approach examining human abilities in controlling or tracking dynamic systems to make them conform with certain time-space trajectories in the face of environmental uncertainty." For example, driving a car requires an individual to notice other cars while manipulating the controls and monitoring the car's instruments in order to maneuver the vehicle safely to a desired destination. In particular, noticing the car in front of you while driving and making adjustments to the speed that conform to certain time-space trajectories is important in order to avoid a collision. Driving requires the individual to track/manipulate more than one element at a time (e.g., looking at the speedometer, looking at other cars, pushing on the pedals, etc.). Tasks requiring manual control raise the question of how well McArthur and Graesser's models of tutoring (which mostly deal with one problem at a time in a stable domain) generalize to tutoring students who are learning dynamic tasks.

Also, driving, flying, and surgery are not completely governed by an ordered set of principles due to the unpredictability of the domains. Therefore, a certain sequence of actions that may apply in one context to achieve a specific outcome would not apply in another context to achieve that the same outcome. For example, in an airplane, climbing to an altitude of 4000 feet from an altitude of 3000 feet in windy weather is based more on skill rather than a fixed set of ordered principles/rules (which may apply when the weather is clear). The reason is because in windy weather there are external factors that could complicate the task. Lastly, driving, flying, and surgery are performed in volatile environments where the task changes spontaneously over time (e.g., wind gusts pushing the plane off the desired heading forcing the student to correct the plane's heading while achieving the task objective). In contrast, tasks performed in domains such as LISP, physics, and algebra remain stable despite any manipulation on behalf of the student.

Unfortunately, there is little or no past research on tutoring in aviation and pilot training that we have been able to identify. A literature search of the *PsychInfo* database for keywords including ("tutoring" or "instruction") and ("aviation" or "pilots" or "flight") did not yield a single empirical study of human tutors helping pilots learn to fly. A second search through the *International Journal of Aviation Psychology* also failed to produce any matching articles. A few articles were identified that described automated tutors for different aspects of piloting (e.g., Chappell, Crowther, Mitchell, & Govindaraj, 1997; Remolina, Ramachandran, Fu, Stottler, & Howse, 2004) but these systems were not based on research with human tutors.

Importance of Research

A review of the literature on McArthur and Graesser's models of one-on-one tutoring reveals nothing about how these models apply to tutoring student dynamic tasks such as driving, flying, or performing surgery. How well do the models generalize to tutoring students these dynamic tasks? Which one fits the data better, and why? Are there certain elements in each model that apply to tutoring dynamic tasks? If so, what are they? Also, do the models apply differently to professionals versus non-professionals? What is the influence of the task characteristics on how tutoring occurs? To investigate these questions we looked at how a certified flight instructor tutors 5 student pilots through a sequence of nine instrument flight maneuvers as described in Chapter II.

CHAPTER II

METHOD

Participants

Five local student pilots and one certified flight instructor participated in the study. Student pilots had between 28-130 hours of experience flying and ranged in age from 18 to 40. Each student pilot received 50 dollars for their participation in the simulator session and 50 dollars for their participation in the airplane session. Student pilots were also able to log one hour of flight experience for their time in the simulator and one hour for their time in the airplane.

Design

The study is observational and consists of one subject variable and two research variables. The subject variable is the student pilot's level of skill, which is measured by the number of hours the student pilot has logged. The first research variable is the flight setting: Students performed the same tasks first in a computer simulator and then in a real airplane. Given the safety concerns of the IRB, the order of the flight settings (first the simulator then the airplane) was held constant across all participants. The second research variable is the type and difficulty of the tasks given to the students during the flight sessions. Nine tasks were given to each student pilot to perform, beginning with the easiest task and ending with the hardest in both sessions (simulator then airplane).

Because the two research variables were invariant across participants, we cannot separate treatment effects from history or practice effects, and therefore we cannot isolate treatment effects from these confounding factors. Thus, the study is not a true experiment; rather it is an observation of how these variables affect tutoring.

Materials

Flyers were posted at a local flight school to recruit student pilots for the study. The flyer explained the basic flight requirements and the incentive for participating.

A consent form was given to the student pilot to fill out if they chose to volunteer for the study. The consent form outlined in more detail the study, the requirements, the incentives and the student's right to withdraw at anytime.

A questionnaire was administered to the pilots before participating in both the simulator flight and the airplane flights. The questionnaire was designed to gather information about where the pilot learned to fly, how long the pilot had been flying, and what kind of planes they have flown (see Appendix A).

The flight simulator is a Jeppesen FS-200 PCATD (Personal Computer-Based Aviation Training Device), which is a hardware and software unit that is certified by the Federal Aviation Administration (FAA) for flight instruction. The PCATD consists of a control box with a yoke and a throttle that is connected to a computer screen (Figure 1). The student pilot is presented with a screen showing the main instruments necessary for flight (Figure 2). A video camera was placed behind the instructor and student in both the simulator and the airplane to capture audio and video data for later analysis.

Figure 1.   Student pilot performing a task on the PCATD.

Figure 2.   PCATD instrument panel screenshot.

Procedure

Five student pilots were asked to perform nine tasks, first in a simulator and then in an airplane, that pose different levels of difficulty. A two-hour time commitment and a three-hour commitment were required for the flight simulator and airplane, respectively. A consent form (see Appendix B) was given to each student to read and sign explaining the student's right to withdraw and discontinue participation at any time during the study.

In each case, the instructor pilot was present to give the student the task instructions and to tutor them during the flight sessions. Each task was performed using seven instruments essential for instrument flight (Tachometer, Airspeed Indicator, Attitude Indicator, Altimeter, Turn Coordinator, Directional Gyro, and Vertical Speed Indicator). During airplane flights the student pilots wore "foggles" or goggles that blocked any view outside the cockpit allowing only a view of the instruments and flight controls. Task 1 to 4 involved a single axis maneuver, Task 5 to 7 involved a double axis maneuver, Task 8 involved a triple axis maneuver, and Task 9 involved an instrument failure (Table 3). The progressive difficulty of the tasks was designed to examine how communication might be affected under more complex situations. Between each task the instructor readjusted or reset the plane to prepare for the next task, allowing a short break for the participants.

The audio track of the video recording was transcribed for coding. Each utterance was broken into constituent clauses, for example the instructor might say, "Watch your heading, you need to pull up." This utterance would be broken into "watch your heading"

Table 3

Task Instructions

Task	Instruction
1	Decrease airspeed to 80 knots while maintaining a heading of 360 degrees and an altitude of 3000 feet.
2	Turn left (standard rate) to a heading of 180 degrees while maintaining an airspeed of 90 knots and an altitude of 3000 feet.
3	Turn right (standard rate) to a heading of 180 degrees while maintaining an airspeed of 90 knots and altitude of 3000 feet.
4	Climb to 4000 feet at 500 feet per minute while maintaining an airspeed of 90 knots and a heading of 360 degrees.
5	Increase airspeed to 100 knots while turning right (standard rate) to a heading of 180 degrees and maintaining an altitude of 3000 feet.
6	Decrease airspeed to 80 knots while descending to 2500 feet at 500 feet per minute and maintaining a heading of 360 degrees.
7	Turn left to a heading of 90 degrees and climb to 4000 feet at 500 feet per minute while maintaining an airspeed of 90 knots.
8	Increase airspeed to 100 knots while turning right a heading of 1-8-0 and descending to 2000 feet at 500 feet per minute
9	Turn left (standard rate) to a heading of 180 degrees while maintaining an airspeed of 90 knots and an altitude of 3000 feet.

which would be coded as a warning and “you need to pull up” would be coded separately as a control command. The coding system was developed based on the broad types of messages given by the instructor to the student. Four main categories were initially defined: commands, warnings, comments, and explanations. The coding system was further divided into 15 subcategories. For example there are control commands (e.g., “Pull back on the yoke”) and instrument commands, such as “Look at the attitude indicator.” The full coding system is shown in Appendix C.

The goal of the coding system is to understand the purpose behind each utterance: Is the instructor trying to warn the student, tell the student what to do, or to teach the student about flight dynamics? The coding system was designed to answer such questions.

The simulator recorded and stored the airplane’s altitude, heading, and airspeed every 10 milliseconds during each simulator flight. Since each flight was also recorded using a video camera, the time from the simulator data and the time from the videotape were synchronized. Synchronizing the simulator flight data and the video tape recording is important in relating the state of the plane (in terms of altitude, heading and airspeed) to the instruction given by the tutor.

Objectives

Five objectives were proposed to examine the nature of individual tutoring during instrument flight. The first objective focuses on whether the tutoring models proposed by McArthur et al. (1990) and Graesser et al. (1995) generalize to the domain of instrument flight. Objectives two, three, and four focus on the frequency (and proportion) of overall utterances and the type of utterances generated by the instructor. Objective five investigates whether the student’s skill level improves across tasks. Each objective contains a short description of how the data were used to address the objective. Some objectives also have a brief description of what we expect to see once the data are analyzed.

Objective 1

The first question we wanted to investigate was: To what extent do the models proposed by McArthur and Graesser account for the behavior of our tutor (the flight instructor) and the student pilots. To investigate this issue, we examined the nature of the dialog between the instructor and student pilots during instrument flight to find out what aspects of McArthur and Graesser's tutoring models applied. The first task was to develop a coding system that classifies each utterance made by the tutor and the student during the tutoring process (Appendix C). An utterance is a statement made by the instructor or student followed by a brief pause. After the coding system was developed, all utterances generated by the tutor and the student were coded by one investigator. Coding allows formal comparisons to be made between our data and each of the two models. Specifically we can measure the percentage of the instructor/student-pilot communication that is captured by each model. We propose that if a high percentage of our utterances match to one or the other model (for example, 80% - 90%), that would provide strong support for the generality of that model to our different domain. We also propose that a low percentage of utterance captured (perhaps 10%-20%) suggests the model does not generalize well to our domain or our tasks. Finally, we propose that an intermediate level of capture (perhaps 40% - 60%) suggests a good fit to a model that might be expanded to encompass the remaining unexplained utterances we observed.

Objective 2

How sensitive is the instructor to student pilots of different skill levels? The level of sensitivity the instructor has to different pilots can be determined by examining the kind of utterances employed during tutoring. Because each utterance is coded, the frequency of different codes can be calculated for each student pilot. The frequency of different codes reflects the level of sensitivity the instructor pilot has towards different student pilots. Given the apparent absence of prior research on this topic, we speculate on several trends in the data. First we speculate (based on the coding system) that a higher proportion of commands and warnings will be given to lesser-skilled pilots because the instructor will be occupied with keeping the airplane under control. With more skilled pilots, who will likely deviate less from the ideal flight state, the instructor will perhaps spend more time commenting on the status of the plane and explaining the dynamics of flight (giving more instructional utterances).

Objective 3

What is the difference in how the instructor tutors students in the airplane versus in the simulator? This difference can be examined by comparing the frequency and type of utterances in each of the two situations, similar to the analysis described in objective 2. We speculate that the instructor will produce more utterances in the simulator than in the airplane for two reasons: 1) the simulator does not have a high level of background noise from the engine and propeller; and 2) the airplane uses an awkward voice-activated intercom that increases the difficulty of communication. Furthermore, we may see a higher proportion of instructional utterances employed by the instructor in the simulator because the instructor is not as concerned with maintaining controlled flight in a simulator compared to a real airplane. A chi-squared test will be performed to see if there are any significant differences between the two situations.

Objective 4

How does the complexity of the task affect tutoring? All flight sessions in both the airplane and simulator require the student pilot to perform nine tasks while being tutored by an instructor pilot. Task 1 through 4 involves a single axis maneuver, Task 5, 6, and 7 involves a double axis maneuver, Task 8 involves a triple axis maneuver, and Task 9 involves a single axis maneuver with an instrument failure (see Table 3). Again, the frequency of the various coded utterances can be calculated for the different tasks for each flight session, similar to objective 2 and 3 as an index of the relationship between complexity and tutoring. A chi-squared test will be performed to evaluate any statistical differences between the different task complexities.

Objective 5

Do the student's flight skills improve during the tutoring session? One way to determine whether the student's flight skills improve is to look at the simulator flight data that is automatically recorded every 10 milliseconds. This data includes a measure of the simulator's airspeed, heading, and pitch (the airplane does not have a similar flight data recorder). An anchor point can be set at the beginning of each task for the appropriate axis or axes to measure how much deviation occurs from the ideal flight state. The ideal flight state is based on a linear extrapolation of the pilot's performance on a given axis. For example if the student pilot needs to increase their altitude from 3000 to 4000 feet, a standard rate of climb of 500 ft/min indicates the climb should take 2 minutes. Any deviation by the student pilot (in feet) can then be calculated based on their performance in that task. If the task requires an axis to remain constant (i.e., maintaining a heading of 360 degrees), all deviations from that steady value are computed as well. Finally, if a task requires more than one axis to perform, each ideal state will be extrapolated from the anchor point until the maneuver is supposed to be completed, and root-mean squared deviations can be computed. Deviations for each axis on each of the nine tasks will be computed for each of the five pilots in the simulator.

If the instructor's tutoring is effective, we would expect to see progressively smaller deviations from the ideal flight state for each task. Unfortunately, the increasing difficulty of flight maneuvers followed in our experimental protocol makes it difficult to interpret the root mean squared deviation values we compute. As students learn, their deviation values should decrease. But as they perform ever more complex maneuvers, their deviation score is likely to increase. Fortunately, there are two maneuvers in our protocol that are of approximately equal difficulty. Task 2 and 3 require the student to perform the same maneuver but in the opposite directions as shown in Table 4.

If the instructor is doing a good job of tutoring we expect to see less deviation from the ideal flight state across pilots and flight settings in Task 3. We expect to see the same amount or more deviation from the ideal flight state across pilots if the tutoring is ineffective.

Table 4

Task Comparison

Task 2	Task 3
A standard left turn to 180 degrees while maintaining 90 knots of airspeed and 3000 feet.	A standard right turn to 180 degrees while maintaining 90 knots of airspeed and 3000 feet.

The next chapter describes the results that answer our five objectives. The aim of each objective will be described followed by an analysis of the data. A brief explanation will be provided as to why certain trends in instrument flight tutoring might emerge.

CHAPTER III

RESULTS

Chapter three presents results of our observational study on tutoring during instrument flight instruction. Each of our five objectives will be considered in turn with two goals in mind. First, we consider how tutoring in our situation compares and constrasts with the previous accounts of tutoring offered by McArthur et al. (1990) and Graesser et al. (1995). Then, we consider more specific questions about details of the tutoring process that have not been carefully reported elsewhere.

Objective 1

How well do the models proposed by McArthur et al. (1990) and Graesser et al. (1995) conform to our results? To address this question we examine two aspects of the dialog between the instructor and students during instrument flight. First, we compare the frequency of utterances produced by the instructor with the frequency of utterances produced by student pilots during all the flight sessions. Thus, the ratio of instructor-to-student utterances can then be computed to gain insight into how tutoring works during instrument flight. Second, we examine the proportion of different utterance types employed by the instructor and the students during instrument flight. We can then determine how these proportions of utterances generated by our tutor and students relate to the different kinds of utterances produced by the tutors and students in McArthur and Graesser’s models.

To provide some context for our results, recall that in McArthur’s model the instructor begins the tutoring session by employing a new-problem microplan. The student then provides the instructor with a brief response. If the student provides an incorrect or incomplete response the instructor employs the remediation microplan which provides a detailed description about how a student can accomplish the next step. If the student provides several incorrect or incomplete answers in a row the instructor will employ the motivational microplan. If the student asks a question or makes a comment that does not pertain to the problem the instructor employs a redirection microplan. Thus, based on McArthur’s model we see the instructors dominating the tutoring session, but they rely on students to tell them what they do or do not know.

In contrast, the tutors in Graesser’s model begin by asking the student an initiating question about a certain topic. Again, the student provides the instructor with a brief response. If the response is incorrect or incomplete the instructor gives the student a hint in the form of a question. The process continues with the instructor giving more specific hints in the form of questions until the student provides the correct answer. Thus, in Graesser’s model we see a collaborative conversation between the instructor and the student.

Instructor Utterances

To begin exploring how both models generalize to tutoring student pilots we first examine the ratio of total utterances made by the instructor to those made by the student. The total number of utterances across all pilots and across both flight sessions (simulator and airplane) is 3961. The instructor makes 3789 utterances (96%), in contrast, the students only make 172 utterances (4%). The data reveals that the instructor does nearly all the talking while the students do very little.

The overall pattern of the instructor dominating the tutoring session and the student talking very little is similar to the instructor-student interaction in McArthur's model. However, in Graesser's model the tutoring dialog is a collaborative conversation between the instructor and the student. Thus, the overall pattern of tutoring used in Graesser's model does not mirror the pattern we see in our study.

Next, we examine the proportion of different utterance types employed by the instructor and students during instrument flight. Again, we can compare how much our tutoring dialog is reflected in McArthur and Graesser's models.

Recall that the tutoring sessions between the instructor pilot and the students were recorded. The audio portion of the recording was transcribed and a coding system was developed to define the various utterance types employed during the tutoring sessions. The coding system includes four main categories and 15 subcategories of utterance types employed by the instructor. In contrast, student utterances were divided into five main categories with no subcategories. We examine the type of utterances produced by the instructor and students in our study and how they relate to the types of utterances produced by the instructors and students in McArthur and Graesser's models.

If all or some of McArthur's model generalizes to instrument flight we ought to see that the instructor produced instructional utterances to guide the student through the task. We should also see evidence that the instructor employed motivational utterances when the student made several mistakes in a row. Additionally, we should see evidence that the instructor produced utterances that redirected the student back to the problem at hand when they asked an irrelevant question or made an irrelevant comment.

Graesser's model relies solely on questions from the instructor, therefore in order for the model (or some parts of the model) to generalize to instrument flight we ought to see several kinds of questions produced by the instructor. First, we should see evidence that our instructor asked a question to introduce a new problem. Also, we should see evidence that the instructor produced probing questions to guide the student through the problem. Finally, we should see evidence that the instructor asked a question at the end of the problem to gauge the student's understanding of the solution.

Table 5 shows that more than half the utterances employed by the instructor were commands (55%). Commands are specific and short directive statements generated by the instructor that require immediate action on behalf of the student. Command utterances play no obvious role in McArthur's model, but are richly represented in the flight transcripts.

Commands bear a surface similarity to the instructions given by McArthur's tutors: Both tell the student what to do. However, commands in our transcripts are specific and need to be executed immediately. For example, a command might be: "Pull

Table 5

Overall Frequency (and Proportion) of the Four Main Categories of Utterance Types Employed by the Instructor Across Flight Settings and Students

Codes	Overall Frequency (and proportions) of utterances for all flight sessions
Commands	1964 (55%)
Warnings	550 (15%)
Comment on Status	414 (12%)
Instructions	625 (18%)
Total	3553

up on the nose." McArthur's tutors gave higher-level instructions: "Multiply both sides of the equation by the least common multiple." Note that the least common multiple is an abstract quantity that must be determined before the instruction can be performed, so the student cannot immediately respond. Also, the student may not know how to determine the least common multiple and may request additional help before beginning to work on the problem. McArthur's instructions lack the immediacy we see in our commands. Of course, McArthur's tutors could have given a specific command: multiply both sides by six. That would have the immediacy and specificity we find in our commands. But students would probably learn the wrong thing with that sort of input, as the student would not know where the number 6 came from.

We also found instructions in our model that are defined as utterances that explain a certain rule about flight. For example, the instructor might say, “Now that you got the new airspeed you readjust your power to maintain airspeed, heading and altitude.” Note that our instructions have the same abstract character of those provided by McArthur’s tutors and lack the requirement for immediate execution. For this reason, we believe that our instructions best match those observed by McArthur and his colleagues, while our commands represent a different kind of utterance altogether.

The tutors in Graesser’s model employ questions throughout the tutoring sessions to guide the students through the problem. Nowhere in Graesser’s model do we see any type of utterances produced by the instructors that resemble commands in our flight transcripts. Commands are counter to the philosophy of deep learning where students discover the answers for themselves that is so much a part of Graesser’s approach.

Why are commands so dominant in our tutorial sessions whereas they have apparently not been observed in tutoring previously? We speculate that the volatility of the task plays an important role. The instructor not only has to tutor the student, but more importantly, he must keep the state of the plane under control. As we have noted, commands are usually immediate and specific action statements telling the student what to do. For example the instructor might say, “Make a left turn” or “Add more airspeed.” Thus, commands act as a way to guide the student in a manner that conveys a sense of urgency. The instructors in McArthur and Graesser's models tutor students in task domains that are stable (e.g., LISP, algebra, physics, etc.). Therefore, their utterances have less of a sense of urgency and more of an element of guiding the student through the problem in a manner that focuses on teaching.

A more detailed comparison can be made by investigating the sub-categories of utterances. Table 6 shows the breakdown of the utterances based on the frequency and proportions employed by the tutor. Instructions, which overall are the second largest proportion of utterances employed by the tutor, resemble the instructions that are the most common type of utterance employed by tutors in McArthur's model. However, several of the subcategories of instructional utterances generated by the tutor in our study have no obvious parallel to the utterances generated by the tutors in McArthur's model.

McArthur's model is based on the tutors employing instructions to guide their students through the problems during the tutoring sessions. We find that 4 of the subcategories of instructions produced by the tutor are similar to the utterances produced by the tutors employing McArthur's model.

Instructions on flying, which make up 8% of the total utterances produced by our tutor, are similar to the instructional utterances employed by the tutors in McArthur's model. For example, the instructor in our study might say, "You could use your pitch on the airspeed and your power to adjust your vertical velocity." In McArthur et al.'s 1990 study there is a table of tutoring techniques used by the instructors. One example in the table is of an instructor tutoring a student on an algebra problem. In the example the instructor says, “I think you probably want to put all *xs* on one side and all *bs* on the other” (p. 239). In both cases the instructor is providing the student with a suggestion on how to proceed with a step in the task/problem.

Table 6

Overall Frequency (and Proportion) of the Fifteen Subcategories of Utterance Types Employed by the Instructor Across Flight Settings and Students

Codes	Specific Codes	Overall frequency (and proportions) of utterances for all tutoring sessions
Commands	Control Command	1621 (46%)
	Instrument Command	142 (4%)
	Task Reminder	54 (1%)
	Instrument Target	25 (1%)
	Target Value	122 (3%)
Warnings	Direct Warning	522 (15%)
	Indirect Warning	3 (0%)
	Forewarning	25 (0%)
Comments	Comment on Status	414 (12%)
Instructions	Instruction on Flying	276 (8%)
	Action Rule	49 (1%)
	Action Outcome	43 (1%)
	Plan	78 (2%)
	Explanation	19 (1%)
	Pilot Critique	160 (5%)
	Total	3553

Action outcomes represent another subcategory of utterance produced by our instructor that is similar to utterances produced by the instructors in McArthur’s model.

An action outcome is a statement explaining the outcome of a particular course of action. For example, the instructor might say, "At a standard rate turn it takes one minute to change course by 180 degrees." Action outcomes make up 1% of the instructions produced by our tutor. In the table of tutoring techniques in McArthur et al.'s 1990 study (p. 242) an instructor tutoring a student in algebra says, "If you put *b/x* it would turn into a 1." Thus, we see the tutors in McArthur's model employing the same type of statements as action outcomes in our coding system.

Explanations are a third subcategory of instructions our tutor produces that are similar to the utterances employed by the tutors in McArthur's study. Explanations make up 1% of the instructions produced by our tutor and convey a cause-and-effect relationship. One example from our flight transcripts is when the tutor says, "If you roll too quickly it stops the turn all at once." We see the same type of cause-and-effect statements made by the instructors in McArthur's et al.'s 1990 study. One example is when an instructor tutoring a student says, "So once you eliminate fractions in any problem, you see how smoothly things begin to go after that." (p. 241).

The final subcategory of instructions generated by our flight instructor that mirrors the type of utterances employed by instructors in McArthur's model is critiques. Critiques, which make up 5% of the utterances employed by the instructor, are instructional messages telling the student what they needed to do on a previous step to achieve the correct result. For example the instructor in our study might say to the student after the student has performed a certain maneuver, "So during the climb there when you meant to make a little bit of a power reduction, you actually reduced the power too much.” The tutors in McArthur et al.’s 1990 study (p.240) employ the same kind of utterances (e.g., “What you could have done is to multiply both sides by negative 1. It doesn't have to be written down but (- l)(-x) = (- 1)(21) that's what you could have done.”)

Two of the six subcategories of instructions produced by our tutor do not match the instructions employed by the tutors in McArthur’s model. Action rules, which make up 1% of the instructions produced by the tutor, are context specific utterances dealing with a possible future event. The tutors in McArthur’s model do not employ instructions that explain how to handle a future event if it arises. Instead, the tutors employ instructions that deal with the step that the student is currently having difficulty with. Another subcategory of instructions produced by our tutor that are not employed by the tutors in McArthur’s model is plans (2%). Plans are utterances produced by our tutor at the beginning of a new task that outlines the steps the student needs to take in order to complete the task. McArthur does not report any instances where the tutor summarizes the basic steps the student needs to take to solve a particular problem.

The proportion of instructional utterances employed by the tutors in McArthur’s model and the tutors in our model are the same (18%). The majority of utterances employed by the tutors using McArthur’s model are performance feedback utterances (42%). The length of the instructions employed by the tutors in McArthur’s model is longer than the instructions produced by the tutor in our study. Each instruction employed by the tutors in McArthur's model is specific and provides the student with considerable detail on how a step is done and, often times, why it is done. Furthermore, most of the instructions produced by the tutors in McArthur's model deal with the step the students are currently having difficulty with. However, the tutor in our study produces instructional utterances pertaining to a problem the students might have had several steps after the mistake was made. For example, after the student climbs to a certain altitude the tutor might say, "When you were climbing you added a little bit too much power, you want to always maintain a rate of climb of 500 feet per minute" Thus, there are similarities in the nature of the instructional utterances produced, however, the way the instructions are employed differ.

Turning to a comparison between our data and Graesser's tutoring frame, we discover that there are no utterances employed by the tutors in Graesser's model that resemble instructions. The tutors rely on three kinds of questions: Initiating questions, probing questions (which consists of hints that get more specific), and knowledge gauging questions at the end of a problem. Graesser's model is designed to prompt students to discover the solutions to problems on their own as a way to promote deep understanding of the material. Providing explicit information to help students undermines this process. The goal of the tutor is to simply guide the student through this process by providing them cues as to whether they are on the correct path of reasoning or not.

Warnings in our study make up 15% of the utterances employed by our instructor during the tutoring sessions. There are three subcategories of warnings: Direct warnings, indirect warnings, and forewarnings. However, based on the data in Table 6, the proportions of indirect warnings and forewarnings were less than 1%. The proportion of direct warnings employed by our tutor was 15%. The tutors in McArthur's model do not employ any utterances resembling warnings.

In Graesser's model hint/prompts employed by the tutors could be thought of as warnings because the instructor is not explicitly telling the student how to correct a problem. A warning simply brings the student's attention to a problem (e.g., "Your airspeed is low") without telling the student how to correct the problem. However warnings in our coding system are usually short statements identifying what the problem is compared to hints in the form of questions (as is the case in Graesser's model).

Comments on status employed by our tutor make up 12% of the utterances during the tutoring sessions. The tutors in McArthur and Graesser's models provide the student with feedback on the status of their performance (e.g., "Yes", "Okay", "No"). On the surface, feedback on the status of the student's performance in both models may be similar to comments on status in our coding system. However, in both models performance feedback is given frequently throughout the tutoring sessions and is used as a cue to encourage or discourage a student from following a certain line of reasoning. A comment on status as defined by our coding system is given infrequently by the instructor and typically occurs at the end of the task. The instructor produces comments on status as a way to tell the student that they have achieved the task objective. Thus, if the task objective is to climb to 4000 feet while maintaining 90 knots and 180 degrees, the instructor would make a comment at the end of the task by saying, “We are at 4000 feet, 90 knots, and 180 degrees, good.”

Student Utterances

The next issue that we want to examine is the nature of the utterances produced by the students. Although the instructor pilot in our study makes fifteen different kinds of specific utterances, the student pilots only make five. The breakdown of the student utterances (as a percentage of the total) across both flights is shown in Figure 3.

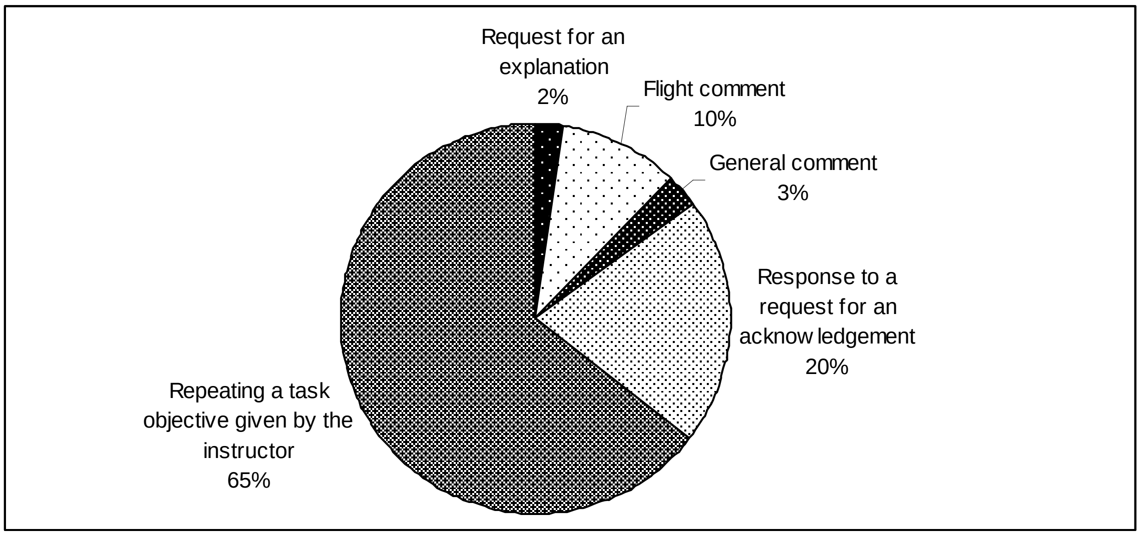

Figure 3. Percentage breakdown of the type of utterances made by students across both flight settings during instrument flight.

In McArthur’s model the student is given a new problem and is asked to provide an answer. Typically the student cannot immediately answer the problem but just provides a brief response. In Graesser’s model the students provide a series of brief answers to probing questions posed by the instructor during the problem. Thus, in both models the students reveal some level of their knowledge (more so in Graesser's model because it relies on a series of probing questions). So according to both models, we should see students attempting to provide the answer to the question given to them or to express their lack of knowledge.

The analysis of the utterances produced by the students reveal that, of the 172 utterances made by students, the majority (65%, n = 112) were made when students repeated back the task objective given to them by the instructor. For example the instructor might say, "CESSNA 3-4 FOXTROT reduce your airspeed to eight zero knots." The student would repeat back the exact utterance produced by the instructor because it is demanded by protocol, so students were simply speaking as instructed. Repeating back a task objective does not give the instructor any information about the student's knowledge of flight. The data suggests that the type of student utterances produced in McArthur and Graesser's models do not generalize to the type of student utterances produced during instrument flight. More than half of the types of utterances produced by the students are repeating back a task objective which reveals no information about their knowledge of flight. However, all the utterances produced by the student's in McArthur and Graesser's models reveal some information about their knowledge of the subject at hand.

The next-most-frequent student utterances are responses to a request for an acknowledgement (e.g., Instructor: "Did you hear me?" Student: "Yes Sir."). Note that 20% (n = 34) of the utterances made by the students are acknowledgements.

The third most common class of student utterance is the flight comment: a statement made by the student pertaining to the flight task. For example the student might say, “I think I added too much power.” Flight comments comprise 10% or 17 of the utterances made by the students. A flight comment generally reveals that a student has made a mistake or doesn’t understand why the plane is in the state that it is in. This seems consistent with the utterances provided by the students in both McArthur and Graesser’s models. But the very low frequency (17 utterances across 18 tasks from 5 different students) implies that this is not a common occurrence.

A general comment is a statement that does not pertain to the flight task. For example the student might say, “It is windy today.” General comments comprise 3% or 5 of the utterances made by the students. Clearly, a general comment made by the student does not offer any insight into their knowledge of flight. General comments are similar to irrelevant comments and/or questions made by the students in McArthur’s model that trigger a redirection microplan by the instructor.

Lastly, a request for an explanation is a question by the students about the state of the plane. For example the student might ask “Why am I losing altitude?” A request for an explanation comprises 2% or 3 of the utterances made by the students. A request for an explanation shows the student has a missing piece of knowledge as to why the plane is in the state it is in or how to get the plane to a desired state.

Therefore, only two out of the five types of utterances made by the students reveal their knowledge about flight. Both the flight comment and the request for an explanation make up 12% or 21 of the utterances made by the students that reveal anything about the student's knowledge of flight. Given that five students performed 9 tasks in two different environments, a total of 90 student tasks were performed. In the majority of these tasks, students said nothing that would help the instructor track their level of understanding.

The first analysis in this section revealed that the five student pilots in this study only made 4% of all the utterances across both flight settings. The finding is inconsistent with the proportion of utterances students made in McArthur and Graesser's models. Furthermore, of the 5 type of utterances produced by the students during instrument flight only two types of utterances reveal anything about their level of knowledge. A total of 21 utterances made across all pilots reveal their understanding of instrument flight instruction. Again, only a very low proportion of utterances made by the students during instrument flight mirrors the utterances made by the students in McArthur and Graesser's model.

The students are revealing very little information about their knowledge of instrument flight, which raises the question, how does the instructor figure out what a student knows and does not know? One plausible explanation is that the instructor is watching the instruments and the control movements in order to figure out what the student knows or does not know about instrument flight (this possibility will be considered further in the discussion chapter).

What we observe in this objective is very little of what the instructors and the students are saying in either model generalizes to instrument flight. Only 16% of the instructional utterances produced by our tutor correspond to the utterances generated by the tutors in McArthur's model. Warning utterances, which make up 15% of the utterances employed by our tutor, generalize to hints/prompts produced by the tutors in Graesser's model. Additionally, only a small fraction of utterances produced by the students during instrument flight generalize to the utterances employed by the students in either McArthur or Graesser's models.

The remainder of our objectives explores more subtle aspects of the tutoring process. Here the existing literature provides little theory or empirical data to build upon. Thus, we provide a largely descriptive account. We begin with an analysis of how the tutor alters his instruction when he works with student pilots of varying levels of skill.

Objective 2

How sensitive is the instructor to student pilots of different skill levels? Does the instructor tutor all pilots the same? Or does the instructor tutor pilots of varying skill levels differently? First we look at the overall totals of utterances by the instructor, then we look in greater detail at the specific types of utterances and how they vary.

We speculate that the instructor will produce a higher frequency of utterances in tutoring sessions involving lesser-skilled pilots. Generating a higher frequency of utterances is one of the ways the instructor can exercise more control over the state of the plane while tutoring lesser-skilled pilots.

It is important to note that instrument flight instruction requires hands-on experience to be learned. Thus, the number of hours flown by a pilot is a good index of their level of skill, and is used by the FAA in judging the qualifications of pilots. Table 7 below shows the number of hours flown by each pilot including the two hours logged after flying in the simulator and airplane.

Table 7

Total Number of Hours Flown Prior to the Study Including the One Hour Simulator Flight and the One Hour Airplane Flight for the Study

Pilot	Total number of hours flown
1	56
2	130
3	84
4	28
5	82

If our prediction is true we ought to see an inverse relationship between flight hours and total utterances (since pilots with low flight hours need more help). Figure 4 graphs the relationship between the number of hours flown and the total number of utterances given by the instructor.

The data show a strong linear relationship between the number of hours flown by the student pilot and the total frequency of utterances given by the instructor ($r^2 = 0.8168$, $p < .005$). Thus, the instructor is sensitive to the skill level of the student pilots and progressively speaks more to pilots with lower flight hours.

Next we examine the proportion of different utterance types given by the instructor to the student pilots. The instructor may give more utterances requiring immediate action (commands and warnings) to student pilots with low flight hours. We speculate that pilots with low flight hours have more gaps in knowledge pertaining to the immediate action steps necessary to complete a task and therefore require more commands and warnings to keep on track.

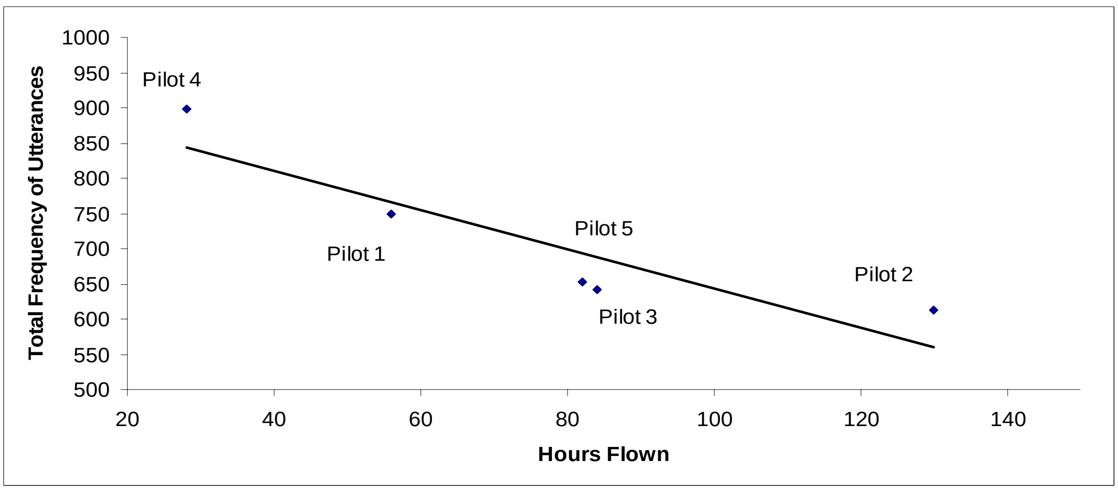

Figure 4. Scatter plot of the number of hours flown by each pilot and the total number of utterances given to each pilot.

We also speculate that the instructor may give more commands and warnings to pilots with low flight hours to avoid the plane reaching an unsafe flight configuration. We further speculate that the instructor might give more instructions and comments (that do not require immediate action) to pilots with high flight hours because they have more knowledge of the action steps needed to complete the task. The pattern would be reflective of the instructor's sensitivity to the student's zone of proximal development:

more competent student pilots receive higher-level instructions that allow them to refine their basic skills that the more novice pilots lack.

Table 8 shows a summary of the proportion of the four main utterance types. The instructor employs approximately the same proportion of commands, warnings, comments, and instructions to pilots 2, 3, and 5. Pilot 2 had the greatest amount of prior flight experience, while pilot 3 and pilot 5 both flew an intermediate number of hours. The instructor employs the highest proportion of commands and comments on status to pilot 4 (who flew the least number of hours). Additionally, the instructor employs the lowest proportion of instructions to pilot 4. Conversely, the instructor employs the lowest proportion of commands and the highest proportion of instructions to pilot 1 (who flew the second least number of hours). Thus, we find pilot 4 and pilot 1 are outliers based on the proportion of utterances the instructor employs.

Table 8

Frequency (and proportion) of the Four Main Types of Utterances Employed by the Instructor Across Flight Settings to Different Pilots

Codes	Pilot 1	Pilot 2	Pilot 3	Pilot 4	Pilot 5
Commands	357 (48%)	340 (55%)	364 (57%)	550 (61%)	353 (54%)
Warnings	123 (16%)	101 (16%)	82 (13%)	137 (15%)	107 (16%)
Comment on Status	92 (12%)	65 (11%)	85 (13%)	82 (9%)	90 (14%)
Instructions	177 (24%)	107 (17%)	110 (17%)	129 (14%)	102 (16%)
Total	749	613	641	898	652

Performing a chi-squared test using the five student pilots and the four major coding categories reveals a significant difference for the frequencies the instructor employs to the various pilots χ^2 (12, N= 5) = 50.3, p < .005.

Next we examine the proportion of the 15 subcategory of utterances the instructor employs to the 5 student pilots (Table 9). Again, the proportion of utterances that pilots 2, 3, and 5 received is similar across all the subcategories. The only exception is pilot 3 receives a slightly lower proportion of direct warnings from the instructor. Pilot 4 receives the highest proportion of control commands while pilot 1 receives the lowest proportion of control commands. Conversely, pilot 1 receives the highest proportion of pilot critiques while pilot 4 receives the lowest proportion of pilot critiques. Another notable pattern is pilot 4 receiving a lower proportion of comments on status compared to the other pilots.

Performing a chi-squared test using the five student pilots and the 15 specific subcategories reveals a significant difference for the frequencies the instructor employs to the various pilots χ^2 (56, N= 5) = 105.8, p < .005.

It is difficult to see where the major differences in the proportions lie for the sub-categories of utterances. To find which sub-categories of utterances yield the greatest amount of variation, we compute a difference score between the pilot with the highest proportion of each utterance sub-category and the pilot with the lowest proportion of the same sub-category. The difference score is equivalent to the range of scores for each sub-category. Difference scores for each utterance type should be between the pilot who flew

Table 9

Frequency (and Proportion) of the Fifteen Specific Types of Utterances Employed by the Instructor for Each Pilot

Main Codes	Specific Codes	Pilot 1	Pilot 2	Pilot 3	Pilot 4	Pilot 5	Difference score
Commands	Control Command	288 (38%)	297 (48%)	293 (46%)	451 (50%)	292 (45%)	12%
	Instrument Command	31 (4%)	12 (2%)	22 (3%)	46 (5%)	31 (5%)	3%
	Task Reminder	13 (2%)	6 (1%)	14 (2%)	8 (1%)	13 (2%)	1%
	Instrument Target	8 (1%)	2 (0%)	4 (1%)	6 (1%)	5 (1%)	1%
	Target Value	17 (2%)	23 (4%)	31 (5%)	39 (4%)	12 (2%)	3%
Warnings	Direct Warning	116 (15%)	99 (16%)	79 (12%)	130 (14%)	98 (15%)	4%
	Indirect Warning	0 (0%)	0 (0%)	0 (0%)	1 (0%)	2 (0%)	0%
	Forewarning	7 (1%)	2 (0%)	3 (0%)	6 (1%)	7 (1%)	1%
Comments	Comment on Status	92 (12%)	65 (11%)	85 (13%)	82 (9%)	90 (14%)	5%
Instructions	Instruction on Flying	67 (9%)	47 (8%)	46 (7%)	71 (8%)	45 (7%)	2%
	Action Rule	16 (2%)	8 (1%)	10 (2%)	7 (1%)	8 (1%)	1%
	Action Outcome	14 (2%)	8 (1%)	7 (1%)	7 (1%)	7 (1%)	1%
	Plan	21 (3%)	14 (2%)	16 (2%)	14 (2%)	13 (2%)	1%
	Explanation	8 (1%)	4 (1%)	3 (0%)	1 (0%)	3 (0%)	1%
	Pilot Critique	51 (7%)	26 (4%)	28 (4%)	29 (3%)	26 (4%)	4%
	Total	749	613	641	898	652	

the greatest number of hours and the pilot who flew the least number of hours in order to suggest a systematic difference in tutoring across pilots.

Based on our analysis we find three ranges of difference score values: 12%, 3-4%, and 1-2%. The utterances with a difference score between 1-2% shows very little variability and will therefore not be discussed further.

Table 9 shows that control commands have the largest difference score across pilots (12%). Pilot 4 (who flew the least number of hours) receives the highest proportion of control commands while pilot 1 (who flew the second least number of hours) receives the lowest. Because the difference score is between the two pilots who flew the least number of hours it suggests that the differences we observed are either unsystematic or reflect differences in skill that are not captured by prior flight-time experience.

There is a 3% difference score for instrument commands that the instructor employs. The instructor employs the highest proportion of instrument commands (5%) to pilot 4, who flew the least number of hours, and pilot 5, who flew an average number of hours. On the other hand, the instructor employs the lowest proportion of instrument commands (2%) to pilot 2, who flew the greatest number of hours.

The instructor employs target values with a difference score of 3%. The instructor employs the highest proportion of target values to pilot 3 (5%) who flew an average number of hours. Conversely, the instructor employs the lowest proportion of target values (2%) to pilot 1, who flew the second least number of hours, and pilot 5 who flew an average number of hours.

We also see in Table 9 that the instructor employs direct warnings with a difference score of 4%. Pilot 2, who flew the greatest number of hours, receives the highest proportion of direct warnings (16%). Pilot 3, who flew an average number of hours, receives the lowest proportion of direct warnings (12%).

Comments on status have a difference score of 5%. Pilot 5, who flew an average number of flight hours, receives the highest proportion of comments on status (14%). Pilot 4, who flew the least number of hours, receives the lowest proportion of comments on status (9%). Thus, the difference score lies between two pilots with skill levels that are not pronounced.

A difference score of 4% is present for pilot critiques employed by the instructor. Pilot 1, who flew the second least number of hours, receives the highest proportion of pilot critique utterances (7%). On the other hand, pilot 4, who flew the least number of hours, receives the lowest proportion of pilot critiques (3%).

In summary, for objective 2 we saw a significant correlation between the skill level of the pilot (measured by past flight experience) and the number of utterances produced by the instructor. The instructor gave a higher number of utterances to pilots with low flight hours compared with pilots with high flight hours. Also, there was a significant difference in the proportion of utterances types employed by the instructor across pilots. Significant score differences were found for control commands, instrument commands, target values, direct warnings, comments on status, and pilot critiques. We predicted that score differences should lie between the most skilled and the least skilled pilot. However, we find that most differences did not lie between the most skilled and the least skilled pilot. We do not have an adequate explanation of why these differences were observed.

Objective three explores the pattern of utterances employed by the instructor in the simulator compared to the airplane. Does the instructor employ a higher frequency of utterances in the simulator or in the airplane? Also, does the instructor employ a higher frequency of certain utterances in the simulator compared to the airplane and vice versa?

Objective 3

Our third objective was to determine whether the instructor varies his technique from the simulator to the airplane. To examine this issue we first compared the total frequency of utterances and then compared the proportion of utterance types (4 broad categories and the 15 subcategories) given to the students in the simulator versus the airplane.

The total frequency of utterances helps us see the degree of control the instructor exerts in each setting. One might perhaps expect to see the instructor giving more total utterances (exerting more control) in the airplane due to the negative consequences of allowing the plane to reach an unsafe flight situation. Alternatively, ease of communication in the simulator might promote more discussion.

The data show that the instructor made a total of 2068 utterances in the simulator and 1485 in the airplane, talking nearly 42% more in the simulator than in the airplane. There are four possible reasons why the instructor produced 42% more utterances in the simulator compared to the airplane. First, the lack of background noise in the simulator may facilitate communication. Next, the airplane uses an awkward voice-activated intercom that likely increases the difficulty of communication. A third reason might be that the instructor does not have to worry about the real life consequence of the student crashing the simulator, which allows the instructor more freedom to talk. A final reason could be due to practice effects. All the students flew in the simulator before flying in the airplane. It could be that, due to practice, the students need less tutoring in the airplane compared to the simulator. We do not really know which reason or combinations of reasons are responsible for the difference in the frequency of utterances employed in the simulator and the airplane.

Next we examined the proportion of different kinds of utterances given by the instructor in the simulator and in the airplane. First we looked at the four main categories of utterances to get a picture of how the instructor tutors student pilots in both contexts (Table 10). We speculated that the instructor would be interested in exerting a greater degree of control in the airplane compared to the simulator due to the dangers inherent in allowing an actual airplane to reach an unsafe flight situation. Thus, we might expect to see the instructor giving a higher proportion of commands and warnings in the airplane. We also might expect to see the instructor giving a higher proportion of instructions and comments in the simulator due to the less volatile nature of the setting.

Table 10

Frequency (and Proportions) of the Four Main Types of Utterances Made by the Instructor Across All Pilots in the Simulator and the Airplane

Codes	Simulator	Airplane
Commands	1158 (56%)	806 (54%)
Warnings	316 (15%)	234 (16%)
Comments on Status	222 (11%)	192 (13%)
Instructions	372 (18%)	253 (17%)
Total	2068	1485

A chi-squared test was performed to determine whether a significant difference arose in the four main utterance types made in the simulator compared to the airplane. Contrary to our expectations, no significant differences were found χ^2 (3, N= 5) = 4.6, $p > .2$.

To further examine how the instructor tutors students in the simulator compared to the airplane we looked at the frequency (and proportion) of utterances given by the instructor in the 15 sub-categories of utterance types. The aim is to see whether there is a significant difference, and if one is found, to illuminate which sub-categories contribute the most to that difference (see Table 11).

Performing a chi-squared test based on the 15 subcategories of utterances to see if there was a significant difference in how the instructor tutors students in both settings shows a significant difference χ^2 (14, N= 5) = 36, $p < .005$. Although we did not find a significant difference for the four main categories (Table 10), this analysis did reveal significant differences within the subcategory breakdown.

Table 11

Frequency (and Proportions) of the Fifteen Types of Utterances Made by the Instructor Across All Pilots in the Simulator and the Airplane

Main Codes	Specific Codes	Simulator	Airplane
Commands	Control Command	968 (47%)	653 (44%)
	Instrument Command	98 (5%)	44 (3%)
	Task Reminder	23 (1%)	31 (2%)
	Instrument Target	11 (1%)	14 (1%)
	Target Value	58 (3%)	64 (4%)
Warnings	Direct Warning	298 (14%)	224 (15%)
	Indirect Warning	3 (0%)	0 (0%)
	Forewarning	15 (1%)	10 (1%)
Comments	Comment on Status	222 (11%)	192 (13%)
Instructions	Instruction on Flying	164 (8%)	112 (7%)
	Action Rule	35 (2%)	14 (1%)
	Action Outcome	30 (1%)	13 (1%)
	Plan	39 (2%)	39 (3%)
	Explanation	12 (1%)	7 (0%)
	Pilot Critique	92 (4%)	68 (5%)
	Total	2068	1485

It is not apparent from the chi squared analysis which utterance types contribute to the significant difference obtained. To locate the source of the significant effect, we computed the difference in proportions for all 15 utterance types between the simulator and airplane. Many of the differences were within ±1%; those were split away from the remainder. Table 12 shows the resulting breakdown. Each sub-table was subject to a second chi-squared test. The table with 5 differences larger than ±1% again led to a significant result χ^2 (4, N= 5) = 23, p < .001. However, the table that had differences with less than ±1% produced an insignificant result χ^2 (9, N= 5) = 13, p > .15.

Table 12

List of Large (>1%) and Small (< 1%) Differences in the Utterances Types

Large difference in the utterance types	Small differences in the utterance types
Higher Proportion of Utterances in the Simulator	Instrument Target
Control Command	Direct Warning
Instrument Command	Indirect Warning
	Forewarning
Higher Proportion of Utterances in the Airplane	Instruction on Flying
Task Reminder	Action Rule
Target Value	Action Outcome
Comment on Status	Plan
	Explanation
	Pilot Critique

Control commands were the highest proportion of utterances given to the students in both the simulator and airplane. A higher proportion of control commands was produced in the simulator (47%) than in the airplane (44%). Control commands can be abstract “make a left turn” or specific “turn left at standard rate (3 degrees per second) to a heading of 90 degrees.” One assumption why control commands were the highest proportion of utterances given is because flying (whether in a simulator or in a real airplane) is volatile and needs to be closely controlled by the instructor.

A higher proportion of instrument commands were given in the simulator (5%) compared to the airplane (3%). Instrument commands direct a student's attention towards a particular instrument that is being neglected. Presumably, by calling attention to a particular instrument, the instructor is hoping the student will recognize a developing problem and address it.

The instructor gave a lower proportion of task reminders in the simulator (1%) compared to the airplane (2%). Task reminders are utterances that remind the student of the overall objective of the task at hand. One plausible explanation for why the instructor gave a lower proportion of task reminders in the simulator is because there is no real life-threatening consequence of crashing the simulator flight compared to the airplane flight.

The instructor gave a lower proportion of target values in the simulator (3%) compared to the airplane (4%). A target value is a specific utterance given to the student without any context telling them what the ideal state of the plane should be. For example, the instructor might say, “80 knots and 360 degrees”.

The instructor gave a lower proportion of comments on status in the simulator (11%) compared to the airplane (13%). Comments are affirmative utterances by the instructor that the student has the plane in the correct state.

There are two main patterns we noticed about the difference in utterances. First, the difference in the proportion of utterances between the simulator and the airplane range between one percent and three percent. The finding suggests that the difference in tutoring in the simulator compared to the airplane is relatively small. Secondly, with the exception of comments on status, the rest of the significant utterances fall under the main category of commands. Thus, it is practical to assume that the instructor's main priority during tutoring (both in the simulator and airplane) is to make sure the plane stays on course. In the airplane this can be reasonably attributed to the consequences of allowing the airplane to reach an unsafe flight situation. In the simulator, we can reasonably assume (based on the small difference in proportions) that the instructor is consistent in his style of tutoring by maintaining his main priority of keeping the plane on course. We do not really have a good explanation of the differences we observed, especially given the lack of counterbalancing in our experimental design.

The next objective deals with how the difficulty of the task objective affects how the instructor tutors students across flight settings.

Objective 4

How does the complexity of the task affect tutoring? To address this question we first have to look at the total frequency of utterances given during each task across pilots and across flight settings. We then want to look at the proportion of different utterance types employed by the instructor based on the complexity of the task.

It is important to point out that the difficulty of the task is based primarily on the number of axes the student must change in order to accomplish the task objective. Task 1 through 4 requires the student to change one axis, Task 5 through 7 requires the student to change two axes, Task 8 requires the student to change three axes, and finally Task 9 requires the student to change one axis with an instrument failure. It is reasonable to assume that the instructor will give a higher frequency of utterances in tasks that require the student to change more axes to accomplish.

To test this assumption we took the average frequency of utterances for each task with one, two, and three axes maneuvers. The goal is to see if there is a linear trend between the number of axes maneuvers and the total frequency of utterances given by the instructor (data shown in Figure 5).

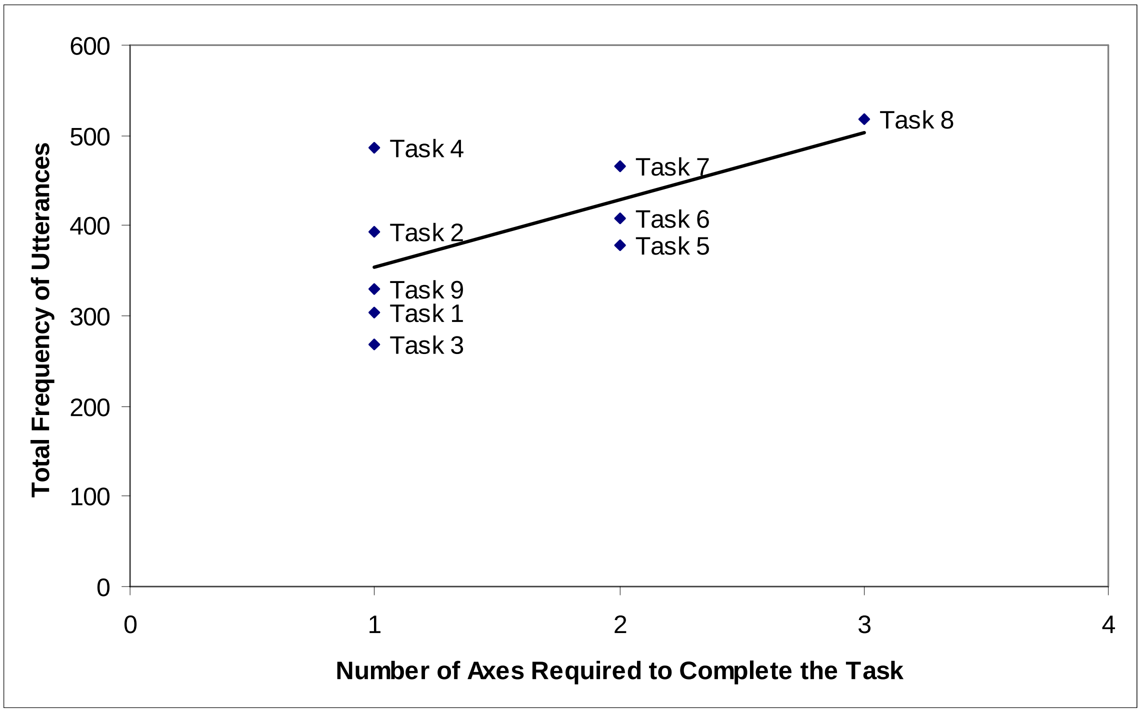

Figure 5. Scatter plot of the average total frequency of utterances given for tasks requiring 1, 2, and 3 axial maneuvers to complete.

We find that the average frequency for tasks requiring one maneuver is 357, the average frequency for tasks requiring two maneuvers is 417, and the average frequency for the task requiring three maneuvers is 518. Thus, we find a significant positive relationship ($R^2 = 0.41$, $p < .05$) between the number of axis maneuvers required to complete the task and the total frequency of utterances given by the instructor to the student. The finding suggests that the more complicated the task gets the more utterances the instructor produces to keep the plane under control.

Students heard an unusually high frequency of utterances in Task 4 even though it required a change on only one axis. Task 4 requires the student to climb to a certain altitude while maintaining all the other axes. Ascending requires the student to not only pull back on the yoke to pitch the airplane upwards, but to also add the appropriate amount of power because of the loss of airspeed as the airplane climbs. Furthermore, leveling off the plane when the altitude is reached again requires manipulation of both the yoke and the throttle (power). The complexity of this maneuver may help to explain the relatively large number of utterances.

The next issue we want to focus on is the proportion of different utterances the instructor gives during each task across students and flight settings. One might expect to see a higher proportion of commands and warnings (utterances requiring immediate action) in tasks requiring more maneuvers. As the complexity of the task increases, so do the opportunities for things to go wrong. Thus, to keep the plane 'on track,' we anticipate seeing a larger number of commands and warnings for multi-axis tasks. One might also expect a higher proportion of comments and instructions (utterances that are informative rather than action-oriented) in tasks requiring fewer maneuvers to accomplish. On simpler tasks, students are more likely to be able to keep the airplane close to the projected track freeing the instructor to operate at a higher level of instruction. To test our predictions, we examined the average proportion of utterances given in tasks requiring 1, 2, and 3 maneuvers. Table 13 shows the average frequency (and proportions) of utterance types given by the instructor to the student in tasks requiring 1, 2, and 3 axis maneuvers to complete.

Table 13

The Average Proportion of Utterances for the Four Main Codes Made for Tasks With 1, 2, and 3 Maneuvers

Codes	1 Axis maneuver	2 Axes maneuver	3 Axes maneuver
Commands	192 (54%)	231 (55%)	312 (60%)
Warning	60 (17%)	61 (15%)	68 (13%)
Comment on Status	38 (11%)	51 (12%)	73 (14%)
Instructions	67 (19%)	75 (18%)	65 (13%)
Total	357	417	518

A chi-squared test was performed and again, contrary to our expectations, no significant difference was found in how the instructor tutors students in tasks with varying difficulty χ^2 (6, $N = 5$) = 12.4, $p > .05$.

Despite the chi-squared value of 12.4 being close to the critical value of 12.592 we do not have definitive evidence that the proportions in Table 9 differ. Even if the values from the analysis in the table did significantly differ, the variations observed in the proportions are small. We do see a suggestion of the trends we projected: commands are most frequent in our 3-axes task and instructions were most common in our 1-axis maneuvers. But the absence of a significant chi-squared effect and the modest differences in frequency do not support a strong difference between our conditions. Thus, it can be said that although the instructor gives a higher total number of utterances in more complex tasks, there is little variation in the kinds of messages the instructor uses to tutor students across tasks of varying complexity.

Objective 5 examines whether the students skill level improves during the tutoring sessions in both the simulator and the airplane.

Objective 5

Do the flight skills of the student pilots improve during the tutoring session? The first challenge here is to measure the skill of each pilot in performing a maneuver. To examine this issue we set an anchor point at the start of each task for the appropriate axis or axes to measure the amount of deviation that occurs from the ideal flight state. The anchor point is based on the first significant control movement. The ideal flight state is based on a linear extrapolation of the pilot’s performance on a given axis (regardless of whether the task requires the student to change the state of the plane on an axis or maintain it). For example, if the student pilot needs to turn right from a heading of 180 degrees to a heading of 360 degrees, the standard rate of turn of 3 degrees per second means it should take the student pilot 60 seconds to achieve the desired heading. Accordingly, 10 milliseconds after the turn had begun, the student should have turned 30 degrees and reached a heading of 210 degrees. Any deviation from the ideal flight path by the student pilot (in this case in degrees) can then be calculated based on their performance in that task. If a task requires more than one axis maneuver to perform, the ideal state of those two or three axes will be assessed using anchor points and any deviation from those axes can be calculated. We can also see any deviation from the ideal flight state on axes that need to be maintained during a task. For example, if a task requires the student to maintain an altitude of 3000 feet, any deviations from 3000 feet (whether the pilot is too high or too low) contribute to their altitude deviation score.

Root mean squared deviations were calculated based on the anchor points every 10 milliseconds for the three main axes (Airspeed, Heading, and Altitude). The deviation values were averaged for tasks with 1, 2, and 3 axis maneuvers. Furthermore, the deviation values for 1, 2, and 3 axis maneuvers were broken into categories based on whether a student had to change a particular axis or not. Sorting the data this way allows us to compare the average root mean squared deviations when a student pilot has to maneuver the plane on a specific axis and when they have to maintain the plane on that same axis.

It is important to point out that the deviations can only be calculated for the simulator flights since the simulator records and stores airspeed, heading, and altitude every 10 milliseconds during each flight. No similar device was available to record and store these variables during the airplane flights.

For airspeed, the data in Table 14 shows that in 1 and 2 axis maneuver tasks the root mean squared deviation values for maneuvers that require an airspeed change are smaller than in 1 and 2 axis maneuver tasks were the student does not have to change the airspeed. The finding may be due to a cross-coupling of axes involved in tasks where the student has to change heading or altitude. During a climb or descent, the change in altitude causes a change in airspeed that must be actively compensated for. To a lesser extent, the same problem arises during turns: the airplane experiences increasing drag during turns, necessitating active compensation to maintain airspeed. Thus, the results are not as odd as they seem.

Table 14

The Deviation in Airspeed for Tasks With 1, 2, or 3 Axes of Change Maneuvers That Require a Change in Airspeed are Shown Apart From Those That do not Require a Change in Airspeed

Axis/Axes	RMS Deviation (in knots) for tasks that require a change in airspeed	RMS Deviation (in knots) for tasks that do not require a change in airspeed
1	7.4	8.4
2	7.5	10.3
3	8.3	

Also note that the difference in deviation between 1 and 2 axis maneuver tasks that does require a change in airspeed is small (.1 increase in deviation) compared to the difference in deviation between 1 and 2 axis maneuver tasks for no change in airspeed

(1.9 increase in deviation). This could be because manipulating airspeed is easier to do than maintaining heading despite the complexity of the task.

Table 15 shows a similar breakdown of deviation scores for heading. The root mean squared deviation values in this case are larger for 1 and 2 axis maneuver tasks that require a change in heading compared to 1 and 2 axis maneuver tasks that do not require a change in heading. The difference in deviation between 1 and 2 axis maneuver tasks requiring a change in heading is larger (2.6 degrees) compared to the difference in deviation between 1 and 2 axis maneuver tasks that do not require a change in heading (-2.2 degrees). It is worth noting that, within the first column, as we shift from 1 to 2 to 3 axes, heading control gets worse (deviations become larger) as expected. Curiously, on the right column, we find that the deviation for the 2-axis task (changing airspeed and altitude) is smaller than the average deviation for the two single-axis tasks (changing airspeed; changing altitude). The reduction in error as the task becomes more complex suggests students are learning to maintain heading even as other axes change.

We also find that the deviation for a 2-axes maneuver that does not require a change in heading is smaller than the deviation for a 1 axis maneuver that also does not require a change in heading. The result may suggest an improvement on behalf of the students in maintaining the heading as the task gets more complicated.

The data for the deviation in altitude (Table 16) shows that for 1 and 2 axis maneuver tasks, the root mean squared deviation for maneuvers that require an altitude change are larger compared with 1 and 2 axis maneuver tasks where the student does not change the altitude. Moreover, the difference in deviation between 1 and 2 axis maneuver tasks for a change in altitude is 47.7 feet compared to the difference in deviation between 1 and 2 axis maneuver tasks for no change in altitude, 18.2 feet. Thus, manipulating altitude as a task gets more complex is more difficult than maintaining altitude as a task gets more complex. The finding could suggest that altitude is a difficult axis to manipulate, especially in combination with other axis changes.

Table 15

The Deviation in Heading for Tasks With 1, 2, or 3 Axes of Change Maneuvers That Require a Change in Heading are Shown Apart From Those That do not Require a Change in Heading

Axis/Axes	RMS Deviation (in degrees) for tasks that require a change in heading	RMS Deviation (in degrees) for tasks that do not require a change in heading
1	14.8	5.7
2	17.4	3.5
3	20.4	

To summarize the deviations observed on the 3 axes, we see a greater deviation in airspeed from the ideal flight state in tasks that do not require the student to change the airspeed (likely due to cross-coupling of axes for some tasks). However, for the heading and altitude we see a greater deviation from the ideal flight state when the student is required to change those axes. The finding might suggest that airspeed is an easier axis to manipulate compared to heading and altitude.

Table 16

The Deviation in Altitude for Tasks With 1, 2, or 3 Axes of Change Maneuvers That Require a Change in Altitude are Shown Apart From Those That do not Require a Change in Altitude

Axis/Axes	RMS Deviation (in feet) for tasks that require a change in altitude	RMS Deviation (in feet) for tasks that do not require a change in altitude
1	67.9	38.2
2	115.6	56.4
3	190.8	

Note the complications of interpreting the previous data because the tasks differ in complexity, however, we do have one opportunity to examine the effect of learning in a more controlled fashion. We can compare the average root mean squared deviation from the ideal flight state that occurs between two consecutive single axis flight maneuvers of equal difficulty. The goal is to see whether we see an improvement in the student's flight skills by looking at the amount of deviation on all three axes (Airspeed, Heading, and Altitude).

We selected Task 2 and 3 because both require the pilots to execute similar maneuvers. In Task 2 the pilots have to turn left at standard rate (3 degrees per second) to a heading of 180 degrees while maintaining an airspeed of 90 knots and altitude of 3000 feet. Task 3 requires the pilots to maintain the same specifications as Task 2 but to turn right instead of left. Table 17 shows the average root mean squared deviation across all pilots for Task 2 and Task 3.

Table 17

The Root Mean Squared Deviation in Instrument Values Across all Pilots in the Simulator for Two Tasks of Equal Difficulty (Task 2 and 3)

Tasks	Airspeed (Knots)	Altitude (Feet)	Heading (Degrees)	Bank (Degrees per Second)
2	8.0	36.9	12.7	9.2
3	9.1	31.8	15.2	8.2

We anticipate that the error scores in Task 3 will be smaller than those in Task 2 of Table 17. Turning first to the heading axis, the data show a greater degree of deviation in heading for Task 3 than for Task 2. However, we find a smaller amount of deviation in bank for Task 3 than for Task 2. The finding suggests that students deviated less in the subsequent task while in the process of achieving the desired heading (banking) but deviated more once the heading was achieved in order to maintain it. Task 2 and 3 also required the student to maintain a specific airspeed and altitude while turning. We find a greater deviation in airspeed in Task 3 but a smaller of deviation in altitude.

To summarize we found several patterns in the tutoring dialog during the flight sessions in the simulator. First, the root mean squared deviation for tasks that do not require a change in airspeed were larger than tasks that do require a change in airspeed. However, the root mean squared deviation for tasks that require a change in heading and altitude were larger than tasks that do not require a change in heading and altitude. We also found the deviation for 2 axes maneuver tasks that do not require a change in heading were smaller than the deviation for 1 axis maneuver tasks that do not require a change in heading.

When we looked at the average root mean squared deviation (for airspeed, heading, altitude, and bank) across two tasks of similar complexity we found a greater degree of deviation in heading for the latter task. However, we find a smaller amount of deviation in banking for the former task.

Overall, the data for objective 5 shows no unambiguous evidence of improvement across tasks. However, this quest is made difficult by the experimental protocol, where more difficult tasks occur later in the sequence. Thus, in order truly know whether learning occurs or not, task difficulty needs to be given to a large pool of participants in a random fashion in order to avoid confounding.

As we review the five objectives discussed in this chapter, we find that 16% of our instructor's utterances generalize to the utterances employed by the instructors in McArthur's model. We also find that 15% of our instructor's utterances generalize to the utterances employed by the instructors in Graesser's model.

The frequency of utterances employed by our instructor varied for pilot's skill level, flight setting, and task complexity. However, we find that the instructor did not vary or varied slightly the frequency of certain utterance types based on pilot's skill level, flight settings, and task complexity.

In the last objective of this chapter we find little evidence suggesting that the student improves during the tutoring session. However, due to the experimental design of this study we are not able to fully know whether the student's skill level improves after each task.

CHAPTER IV

DISCUSSION

Hypothesis and Results

The main goal of this study is to examine the patterns of instructor-to-student communication in instrument flight and how it relates to two representative models of one-on-one tutoring proposed by McArthur et al. (1990) and Graesser et al. (1995). How well does each tutoring model generalize to tutoring student pilots how to perform instrument flight? Previous investigations of tutoring have examined problem-solving tasks that follow an explicit set of principles and do not change spontaneously over time.

We hypothesized that tasks requiring manual control which are not governed by an explicit set of principles, and are performed in unpredictable domains (where the problem does change over time) may reveal new features and characteristics of tutoring not evident in either McArthur’s or Graesser’s model of tutoring. To test the generality of both models we examined how a certified flight instructor tutors five student pilots to perform nine instrument flight maneuvers in a simulator and in an airplane. We organized our investigation into five objectives that examine the patterns of communication between the instructor pilot and the student pilots. The results will be discussed on an objective-by-objective basis.

Objective 1

Objective one focuses on how much of the models proposed by McArthur et al. (1990) and Graesser et al. (1995) apply to tutoring pilots how to fly. To answer this question we recorded one certified instructor pilot tutoring five student pilots in two flight settings performing nine tasks. The data were transcribed and a coding system was created to define the different utterances made by the instructor and the student during the 10 flight sessions.

The first step was to examine the overall proportion of utterances employed by the instructor and the students during all the flight sessions. We found that the instructor dominates the dialog during the flight sessions. We also found that student pilots produce only 21 utterances in 90 student tasks that reveal anything about their knowledge of flight.

The overall pattern of the instructor dominating the tutoring dialog is similar to the overall pattern we see the tutors in McArthur et al.'s 1990 study. The similarity may be due to the tutor in our study and the tutors in McArthur's study being professionals. In contrast, the tutors applying Graesser's model were either paraprofessionals or peers, which explain why the tutors applying Graesser's model engaged in a collaborative conversation with the student rather than dominating the conversation.

Perhaps there is a philosophical difference in the approach to tutoring based on different goals: deep understanding in Graesser's situation and equipping students with algebra problem-solving skills in McArthur's situation. We don't really know why these different approaches exist or when one is applicable or preferred over the other. In our study the instructor is concerned about getting the student through the task rather than encouraging deep learning on behalf of the student.

We then examined the proportion of various utterance types employed by our instructor to see how they would fit into McArthur and Graesser's models. We found that more than half the utterances (55%) employed by our tutor were commands. Commands were not employed by the tutors in either McArthur or Graesser's models. Commands are short directive statements that tell the student what immediate action they need to take. One plausible explanation why the instructor employs a large proportion of commands is that errors in a volatile and dynamic task can build up over time and cause significant problems. If the airplane is in a slow descent, for example, it will eventually reach an unsafe altitude. Commands provide the fastest way to prevent a developing problem from becoming worse. The predominance of commands in our corpus suggests that the flight instructor feels it important to help the student keep the airplane close to the target flight path. Most commands given are control commands, demanding immediate manipulation of the airplane's controls. These issues do not arise in the static problem-solving tasks studied by McArthur or Graesser.

Instructions were the second highest proportion of utterances employed by the tutor during instrument flight (18%). We found that four of the six subcategories of instructions defined in our coding system are similar to the utterances the tutors employed in McArthur's model. The two exceptions were Action outcomes (1%) and plans (2%) that had no similarity to the utterances employed by the tutors in McArthur's model. Thus, 16% of the utterances employed by our instructor resemble the utterances employed by the tutors in McArthur's model.

Warnings make up 15% of all utterances produced by our flight instructor. Warnings in our coding system resemble hints/prompts employed by the tutors in Graesser's model. In both cases the instructor is directing the student's attention to a problem without telling the student how to solve it. However, the way the instructor directs the student's attention to a problem in our study is different from how the instructors using Graesser's model direct the student's attention. The instructor in our study might say, "Watch your heading" as a means to direct the student attention to a problem that needs to be fixed. The instructors in Graesser's model ask a question to direct the student's attention.

Comments on status make up 12% of the utterances employed by the instructor in our study. On the surface comments on status seem similar to comments on a student's performance in both models (e.g., "Yes", "Okay", "No"). However, comments employed by our instructor are lengthy and employed at the end of a task to signal to the student that they have achieved the task objective. For example, the flight instructor might say, "We have reached our altitude of 4000 feet, we are at 90 knots, and flying at 180 degrees, all is looking good." In both McArthur and Graesser's models the instructors employ comments on the student's performance after every response given by the student. Furthermore, comments are usually short one word statements that act as cues telling the student if they are on the right line of reasoning or need to change their line of reasoning. Thus, comments on status in our study do not resemble comments on the student's performance in both McArthur and Graesser's models.

Thus, we find that 16% of the utterances employed by our instructor resemble utterances employed by the instructors in McArthur's model. We also find that 15% of the utterances employed by our instructor resemble utterances employed by the instructors in Graesser's model. Thus, we see a low percentage of utterances (10-20%) employed by the instructors in both models that generalize to tutoring students instrument flight instruction. The poor fit of previous accounts of tutoring to our situation suggest that tutoring differs significantly from problem-solving tasks to manual-control tasks, and we need a better account of the task influences on tutoring.

Objective 2

The second objective looks at whether the instructor changes his method of tutoring with student pilots of different skill levels. First we examined the overall frequency of utterances given by the flight instructor to all student pilots across both flight settings and across all nine tasks. We found a negative linear relationship between the number of hours flown by the student pilot and the total frequency of utterances given by the instructor. Thus, the instructor gave a progressively higher frequency of utterances to pilots with lower flight hours. The first piece of evidence suggests that the instructor might be more sensitive to pilots of different skill levels and exercises more control over the plane with student pilots that have less flight experience.

Examining the proportions of the four main categories of utterances and the fifteen subcategories of utterances yielded a significant difference. To find which utterances yielded the greatest difference in proportions a difference score was calculated. Significant score differences were found for six subcategories of utterances: control commands, instrument commands, target values, direct warnings, comments on status, and pilot critiques. The differences in score were not pronounced enough to suggest that the instructor pilot tutors students differently. Additionally, the score differences were between pilots of comparable levels of skills rather than pilots with distinctly different skill levels. Finally, nine of the 15 subcategory of utterances had a score difference of 1% or less.

Thus, we see pronounced differences in the frequency of utterances employed by the instructor across pilots of different skill levels. However, we do not see pronounced differences in the proportion of various utterance types employed by instructor to pilots of different skill levels.

Objective 3

Objective three explores the difference in how the instructor tutors students in the simulator versus in the airplane. The data reveals that the instructor produces more utterances in the simulator than in the airplane. No significant difference was found in the proportions of the four broad utterance types (commands, instructions, warnings, and comments) employed by the instructor in the simulator and in the airplane. However, upon examining the 15 sub-categories of utterances, we found a significant difference in the proportion of utterances. The data show that control commands and instrument commands differed significantly and were employed in higher proportions in the simulator while task reminders, target values, and comments on status were also significantly different and employed in higher proportions in the airplane. Thus, four of the five utterance types employed by the instructor in both settings: control commands, instrument commands, task reminders, and target values fall under the category of commands.

Objective 4

The fourth objective examines how the complexity of the task (based on the number of maneuvers needed to complete the task) affects tutoring. We found a significant positive relationship between the number of axis maneuvers required to complete the task and the total average number of utterances given by the instructor to the student. Thus, the instructor progressively produced more utterances based on the number of axis maneuvers required by the task. We also found no significant difference in the proportions of the four main utterance types (commands, instructions, warnings, and comments on status) given by the instructor based on the number of maneuvers needed to complete the task (1, 2, or 3 maneuvers). Although the instructor gives a higher number of utterances in more difficult tasks, there is little variation in the kinds of messages produced by the instructor as he tutors students across tasks of varying difficulty. Thus, the complexity of the task has no bearing on how the instructor tutors student pilots.

Objective 5

Objective five focuses on whether the flight skills of the students improve during the tutoring session. An anchor point was set at the beginning of each task to track any changes made on all axes. Thus, we were able to calculate any deviations from the ideal flight state whether a student had to maneuver the plane on an axis/axes or to maintain the plane on an axis/axes. We found less average deviation in airspeed in tasks that require the student to change their airspeed to a given target value compared with tasks that require the student to maintain their airspeed at a constant value. However, we found more deviation in heading and altitude in tasks that require the student to change their heading and altitude to a target value versus tasks that require the student to maintain their heading and altitude at a constant value. The average deviation values increases as the number of maneuvers required to complete the task increases. To explore this last issue further we compared the root-mean-square deviation values in two consecutive tasks with similar difficulties. We found that the average deviation across airspeed and heading increased instead of decreased, suggesting little or no improvement in the pilots performance.

Implications

In the first objective we found that 16% of McArthur's model and 15% of Graesser's model generalize to instrument flight. We also found that commands, which made up more than half the utterances produced by our instructor, are unique to flight tutoring. The finding suggests that the instructor's primary concern is to keep the plane under control while teaching the student. It is not clear whether the student understands why the instructor employs certain commands. It could be the case that the student understands the reason why a certain command was employed by the instructor, and thus the student might learn a new rule of flight. However, the frequency of utterances employed by the instructor may not give the student time to formulate a reason why a command was employed.

In objective 2, 3, and 4 we see a distinct difference in the overall frequency of utterances employed by the instructor. However, the differences in the proportion of utterance types employed by the instructor are not as pronounced. The finding may suggest that the instructor produces more utterances as a means to control the state of the plane rather than specific utterance types.

Limitations of Current Study

In the first objective we notice that the instructor does not ask the student questions in order to diagnose what the student knows and does not know about instrument flight. The instructor seems to be observing the instruments and how the student is manipulating the controls in order to gain information about their knowledge of instrument flight. However, by observing the instruments and the student’s behavior, the instructor has at best, an imperfect understanding of the student’s level of knowledge. It could be that the student knows how to perform a particular maneuver, but due to other factors, such as an unexpected gust of wind or multiple maneuvers, the student appears to have a lack of knowledge. Thus, by observing the instruments or the student’s behavior the instructor only gains a surface level understanding of what a student knows and does not know. In contrast, the instructors using both McArthur and Graesser's models explicitly elicit knowledge from the student in order to address any misunderstandings or lack of knowledge the student may have.

To better understand how the instructor gains information about the student's level of knowledge an eye-tracking device needs to be used by the instructor during instrument flight. The current study did not use such a device and is thus, one of the limitations of our study.

A second limitation is that all participants had to perform the same nine tasks in the same order. Moreover, each participant had to fly in a simulator before flying in a real airplane. The order of the flight setting was due to the safety concerns set by the IRB for this study. Because the two research variables did not differ across participants, the treatment effects cannot be separated from the practice effects. Thus, the study is not a true experiment but instead an observation of how the variables affect the tutoring process.

Another limitation of the current study is the lack of inter-rater reliability measures of the coding system. A second rater was not available to code the transcript to test the reliability of the coding system in this study.

The final limitation is the small number of student participants in the study (N = 5). Furthermore, the range of experience the student pilots had was also small (between 28-130 hours of flight experience). To gain a better understanding of how one-on-one tutoring works during flight a larger number of participants with a wider range of experience are needed.

Future Research

Future research should focus on using a larger sample size of student pilots that have a greater range of experience. The two research variables, flight setting (simulator versus airplane) and task difficulty should be experimentally counterbalanced across participants to minimize history effects.

REFERENCES

Anderson, J. R., Boyle, C. F., & Reiser, B. J. (1985). Intelligent tutoring systems. *Science, 228,* 456-462.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13,* 3-16.

Chappell, A. R., Crowther, E. G., Mitchell, C. M. & Govindaraj, T. (1997). The VNAV Tutor: Addressing a mode awareness difficulty for pilots of glass cockpit aircraft. *IEEE Transactions on System, Man, and Cybernetics, Part A*, 327-385.

Chi, M. T. H. (1996). Constructing self-explanations and scaffolded explanations in tutoring. *Applied Cognitive Psychology, 10*, 33-49.

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science, 25*, 471-533.

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, *19*, 237-248.

Fitz-Gibbon, C. T. (1977). An analysis of the literature of cross-age tutoring. Washington, DC: National Institute of Education, (ERIC Document Reproduction Service No. ED 148 807).

Fox, B. A. (1991). Cognitive and interactional aspects of correction in tutoring. In P. Goodyear (Ed.), *Teaching Knowledge and Intelligent Tutoring* (pp. 149-172). Norwood, NJ: Ablex Publishing Corporation.

Fox, B.A. (1993). *The human tutorial dialogue project*. Hillsdale, NJ: Erlbaum.

Gertner, A. S., & VanLehn, K. (2000). Andes: A coached problem solving environment for physics. In G. Gautheier, C. Frasson & K. VanLehn (Eds.), *Intelligent Tutoring Systems: 5th international Conference, ITS 2000* (pp. 133-142). New York: Springer.

Graesser, A. C. (1992). Questioning mechanisms during complex learning. Memphis State University, Memphis, TGN (ERIC Document Reproduction Services ED No. 350 360).

Graesser, A. C. (1993a). Dialogue patterns and feedback mechanisms during naturalistic tutoring. *Proceedings of the 15th Annual Cognitive Science Society* (pp. 126-130). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Graesser, A. C. (1993b). *Questioning mechanisms during tutoring, conversation, and human-computer interaction*. Memphis State University, Memphis, TN (ERIC Document Reproduction Service No. TM 020 505)

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, *31*, 104-107.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialog patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychology, 9,* 359-387.

Juel, C. (1991). Beginning reading. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research, Vol. 2* (pp.759-788). White Plains, NY: Longman.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30-43.

McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction*, *7*, 197–244.

Merrill, D. C., Reiser, B. J., & Landes, S. (1992). Human tutoring: Pedagogical strategies and learning outcomes. *Presented at the annual meeting of the American Educational Research Association*.

Merrill, D. C., Reiser, B. J., Merrill, S. K., & Landes, S. (1995). Tutoring: Guided learning by doing. *Cognition and Instruction*, *13*, 315-372.

Person, N. K., Graesser, A. C., Magliano, J. P., & Kreuz, R. J. (1994). Inferring what the student knows in one-to-one tutoring: The role of student questions and answers. *Learning and Individual Differences*, *6*, 205-229.

Remolina, E., Ramachandran, S., Fu, D., Stottler, R., & Howse, R. W. (2004) *Intelligent Simulation Based Tutor for Flight Training*, Proceedings of the Industry/Interservice, Training, Simulation & Education Conference.

Robyn, A., Stasz, C., Ormseth, T., Lewis, M. W., & McArthur, D. (1991). Implementing a novel computer-integrated algebra course. RAND N-3326-ED.

Schoenfeld, A., Gamoran, M., Kessel, C., Leonard, M., Or-Bach, R., & Arcavi, A. (1992) Toward a comprehensive model of human tutoring in complex subject matter domains. *Journal of Mathematical Behavior, 11*, 293-319

Slavin, R. E. (1987). Making Chapter 1 make a difference. *Phi Delta Kappan, 69*(2), 110–119.

VanLehn, K. (1996). Conceptual and meta learning during coached problem solving. In C. Frasson, G. Gauthier, & A. Lesgold (Eds.), *ITS96: Proceeding of the Third International conference on Intelligent Tutoring Systems.* New York: Springer-Verlag.

Wickens, C. D. & Hollands, J. (2000). *Engineering Psychology and Human Performance (3rd Ed.).* Upper Saddle River, NJ: Prentice Hall.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

APPENDIX A

PREFLIGHT QUESTIONNAIRE

QUESTIONNAIRE

Dynamics of Instructor-Student Interactions

The purpose of this questionnaire is to ascertain your aviation background. We are interested in your current flight status as well as your previous aviation experience. Please answer the questions as accurately as possible. No identity information gathered in this study will be disclosed to anyone other than Dr. Doane and her collaborators.

Note- this page will be destroyed when you complete (or withdraw from) the experiment and at that time your name will not be linked with your subject id.

Subject number:

1. Name ______________________

2. Local phone number ________________

Subject number:

3. Age __________

4. Gender: Male_____ Female_____

5. Education High School _____

2 Year College_____

4 Year College BA or BS_____

Post Graduate Degree______

6. Certificates and ratings (**Airplanes, Helicopters, Gliders, others**). Please check certificates and ratings you have obtained. The default is assumed to be airplanes, please indicate if certificates/ratings are for helicopters, gliders, etc.

Certificates	Ratings	Approx. Month/Year
STUDENT_______		________
PRIVATE________	Single Engine	
	Land _______	_________
	Sea _______	_________
	Multi Engine	
	Land _______	_________
	Sea _______	_________
	Instrument	
	_______	_________

Certificates	Ratings	Approx. Month/Year
COMMERCIAL________	Single Engine	
	Land ________	_________
	Sea ________	_________
	Multi Engine	
	Land ________	_________
	Sea ________	_________
	Instrument	
	________	_________
	Multi Engine – Center line (Military)	
	Land ________	_________
	Sea ________	_________
FLIGHT INSTRUCTORS_______	Single Engine	
	Land ________	_________
	Sea ________	_________
	Multi Engine	
	Land ________	_________
	Sea ________	_________

Instrument

____

Certificates	Ratings	Approx. Month/Year
		Multi Engine
	Center line (Military)	
	Land ________	_________
	Sea ________	_________
ATP_______	Single Engine	
	Land ________	_________
	Sea ________	_________
	Multi Engine	
	Land ________	_________
	Sea ________	_________
	Instrument	
	________	_________
	Multi Engine – Center line (Military)	
	Land ________	_________
	Sea ________	_________
Certificates	Ratings	Approx. Month/Year

7. Fill in the appropriate flight hours.

a) Total time ________

b) Total instrument time (hood and actual instrument) _________

c) Total actual instrument time ________

d) Total ground trainer and simulator time __________

e) Total instrument time, last 90 days (hood and actual instrument) _________

f) Total time, last 90 days ________

8. Please indicate below, with an "X" the type of aircraft you have flown. If you have flown any aircraft type that is not listed below, please list the aircraft in the "other" category.

AIRCRAFT TYPES

A. **SINGLE ENGINE - TRAINER**

Archer_________ C-182_________ Cherokee________ Tomahawk________

BE-19/23______ C-150/152______ Tampico_________ Warrior___________

Other____

B. SINGLE ENGINE – COMPLEX TRAINER (RETRACTABLE)

Tobago_________ Arrow________

Other____

C. SINGE ENGINE – ADVANCE COMPLEX TRAINER (RETRACTABLE – HIGH POWER)

Bonanza_________ C-120________ Lance_______ Malibu________

Other____

D. SINGE ENGINE – TAIL DRAGGER

A75N-1__________ Cub_________ Decathlon________ Citabaria_________

Cap 10_________ Other____

E. MULTI ENGINE – LIGHT

BE-76__________ Other____

F. MULTI ENGINE – MEDIUM

Navajo__________ Baron__________ Seneca_________ Aztec_____________

C-310___________ Other____

G. TURBO PROP

King Air_________ Other____

H. MULTI ENGINE – HEAVY

____

I. SINGLE ENGINE - JET

F-16D___________ Other____

J. MULTI ENGINE – JET

T-41A_________ T-41B_________ T-41C_________ YF-4E___________

T-37A_________ T-37B_________ T-4C___________

Other____

9. Please indicate below, with an “X”, the number of different airports you have flown into.

a) Large (i.e., O’Hare, Birmingham International, St. Louis Lambert, etc.)

0	________	5-8	________
1	________	8-10	________
2	________	10-15	________
3-5	________	More than 15	____

b) Small (i.e., Starkville, Columbus, etc.)

0	________	8-10	_________	30-40	______
1	________	10-15	_________	40-50	______
2	________	15-20	_________	50+	______
3-5	________	20-25	_________		
5-8	________	25-30	_________		

10. Have you had and experience flying in mountainous areas?

No ________

Yes ________

11. Do you know the instructor/student pilot?

No ________

Yes ________

12. Have you received/given any flight instruction from/to the instructor/student pilot?

No ________

Yes ________ (How many hours? ________)

APPENDIX B

STUDENT CONSENT FORM

CONSENT FORM – Student Pilot
Verbal Protocol

You are invited to participate in a study entitled "Verbal Protocol" conducted by Dr. Stephanie Doane of the Institute for Neurocognitive Science and Technology of Mississippi State University. The study involves interacting with a flight instructor while using a flight simulator to complete flight maneuvers. During the experiment, control movement and performance data will be recorded. In addition, your interaction with the flight instructor will be recorded on video. There are no expected discomforts or risks involved in your participation. Equally, there are no anticipated direct benefits to you other than monetary payment. We do ask that you keep your interactions with the flight instructor in confidence.

This experiment will require a 2-hour time commitment. For your participation, you will be paid $50 in cash at the end of the study. You are free to withdraw your consent and discontinue participation in the study at any time, in which case your data will not be used in the study. Because the session is video recorded, it is impossible to mask your identity. However, in any publication of the study, all information will be anonymous.

This project is funded by a federal agency. Therefore, in the event of an audit, funding agency personnel will have access to all information. Additionally, because the information is held by a state entity, it is subject to disclosure if required by law. Except for the circumstances described above, only Dr. Doane and her collaborators will have access to the information.

The investigator will be glad to answer any questions regarding the study procedures. However, answers that may influence your performance will be deferred until the end of the simulation time. For further questions, you may contact Dr. Doane at (662) 325-4718 or Mark Jodlowski at (662) 325-2481. For any additional information regarding human participation in research, please contact the Mississippi State Regulatory Compliance Office at (662) 325-3294.

I understand the above information and voluntarily consent to participate in the experiment entitled "Verbal Protocol." I have been given a copy of this consent form to keep for my own records.

Signature ______________________________ Date _______________

APPENDIX C

CODING SYSTEM FOR FLIGHT COMMUNICATION

Coaching

Commands

Demand immediate action on the part of the pilot

Coaching - Control Command - Specific	C-CC-S	A command to manipulate a specific control. Example: “Pull back on the yoke” (the control is mentioned by name)
Coaching - Control Command - Abstract	C-CC-A	A command that does not explicitly mention a control. Example: “Make a power reduction” Example: “Pull up on the nose”
Coaching - Control Command- Abstract - with a Reason	C-CC-A+R	A command given with a reason immediately after it. Example: “Lower the nose to gain some airspeed”
Coaching - Control Command - Abstract - with a Target	C-CC-A+T	A command given with a target/goal to achieve. Example: “Pull that nose on up to get to that 500 foot per minute rate of climb”
Coaching – Control Command - Abstract - Forewarning	C-CC-A+FW	A command given with a forewarning that does not explicitly mention a control Example: “make a standard rate turn, not more than standard rate”
Coaching - Control Command- Abstract - with a Reason followed by a Target	C-CC-A+R+T	A command given with a reason then a target. Example: “Lower the nose to gain some airspeed because we want 80 knots”
Coaching - Control Command- Abstract - with a Target followed by a Reason	C-CC-A+T+R	A command given with a target then a reason. Example: “Lower the nose to get 80 knots so we can gain some airspeed”
Coaching - Control Command- Abstract - with a Reason followed by a forewarning	C-CC-A+T+FW	A command given with a target then a forewarning. Example: “so you could make a 50 RPM reduction. Probably no more than that.”
Coaching - Instrument	C-IC-S	A command to look at a specific instrument.

Command - Specific		Example: "Look at the attitude indicator"
Coaching - Instrument Command - Abstract	C-IC-A	A command that does not explicitly mention an instrument to look at. Example: "Watch your heading" or "Check your airspeed"
Coaching - Direct Warning	C-DW	A direct message calling the pilot's attention to a problem with the current airplane status. Example: "Your airspeed is low"
Coaching - Direct Warning with an explanation	C-DW+Expl	the airspeed is slightly high because we are still running the power setting that will eventually get us up to 90 knots.
Coaching - Indirect Warning	C-IW	An indirect message calling the pilot's attention to a problem area with the current airplane status. Example: "Your airspeed is too high, I think you know how to correct that"
Coaching – Forewarning	C-FW	Forewarning a pilot about a future situation that needs to be avoided. Example: "Don't just pitch up. You've got to add that power at the same time."
Coaching - Pilot Style	C-PS	A comment about the general style of flight control. Example: "Make small corrections"
Coaching – Comment on Status	C-COS	Making an observation about the status of the airplane. Example: "Matter of fact we are correcting downward slightly and our airspeed is slightly low"
Coaching – Task Reminder	C-TRem	A reminder given to the student about the objective of the task. Example: "We want to be at 80 knots at 4000 feet, heading 1-8-0"
Coaching – Instrument – Target	C-I-Targ	A target instrument value given to the student to achieve. Example: "It is gonna take almost to the red line with power"
Coaching – Target	C-Targ	A value given to the student to achieve. Example (no context): "80 knots, 3-6-0"

Instructions

Information that a student needs to have to perform the task properly

Instruction – Experiment	I-Expr	An instruction or a set of instructions given to the pilot that are part of the experimental procedure. Example: “When I give you instructions I want you to say them back to me as is I were a controller” Example: “If you were on the radio you would always acknowledge who was turning left to 180.”
Instruction – Flight Procedure	I-FP	Instructing students about standard flight procedures. Example: I: “ TIGER 01XRAY decrease your airspeed to 80 knots descend to 2500 at 500 foot per minute” S: “Decrease airspeed to 80 knots descend to 2500 at 500 feet per second for TIGER 01XRAY” I: “ Lets make that 500 feet per minute”
Instruction- Flight	I-Fly	Instructions about how to fly the airplane. Example: “you could use your pitch on the airspeed and your power to adjust your vertical velocity either one of those combinations will work, its just, you have to kind of control one with the other, divide and conquer so to speak they are actually both interrelated”
Instruction - Flight - Instrument	I-Fly+I	Instructions on how to scan the instruments. Example: “so you have to speed up your scan and include that attitude indicator”
Instruction – Action Rule	I-AR	A statement conveying a situation-action rule: In this situation perform that action. Example: “When you are 100 feet low, ascend by 200 feet per minute.”

Instruction – Action Outcome	I-ActOut	A statement that explains the outcome of a particular course of action. Example: “At a standard rate turn it takes one minute to change course by 180 degrees.”
Instruction - Plan – Abstract (1 step)	I-P-A+1	Plan given to the pilot on how to maneuver the plane. It is abstract and does not explicitly mention controls. Example: “Once you got the new airspeed you readjust your power to maintain airspeed, heading and altitude”
Instruction - Plan – Abstract (2 steps)	I-P-A+2	Plan given to the pilot on how to maneuver the plane. It is abstract and does not explicitly mention controls. Example: “As you roll in use the attitude indicator to roll towards what you know is the standard rate of turn and then look at the turn coordinator to stop at standard rate.”
Instruction - Plan – Abstract (3 steps)	I-P-A+3	Plan given to the pilot on how to maneuver the plane. It is abstract and does not explicitly mention controls. Example: “As you roll in use the attitude indicator to roll towards what you know is the standard rate of turn and then look at the turn coordinator to stop at standard rate.”
Instruction - Plan – Abstract (4 steps)	I-P-A+4	Plan given to the pilot on how to maneuver the plane. It is abstract and does not explicitly mention controls. Example: “You want to start that roll-out, anticipate that it is gonna balloon a little bit. You need to lean forward over, back off on that power to keep the airspeed within, where you want it. You want to over shot a little bit.”
Instruction - Plan - Specific (1 Step)	I-P-S+1	Plan given to the pilot on how to maneuver the plane. It mentions a specific control. Example: “Once you got the new altitude you push forward on the yoke to level off.”

Instruction - Plan – Specific (2 Steps)	I-P-S+2	Plan given to the pilot on how to maneuver the plane. It mentions specific controls. Example: "Once you got the new altitude you push forward on the yoke to level off then you pull back on the throttle to get the airspeed you need"
Instruction - Plan – Specific (3 Steps)	I-P-S+3	Plan given to the pilot on how to maneuver the plane. It mentions specific controls. Example: "Once you got the new altitude you push forward on the yoke to level off then you pull back on the throttle to get the airspeed you need and once you've done that you press on the right rudder to make a slight turn"
Instruction - Plan – Specific (4 Steps)	I-P-S+4	Plan given to the pilot on how to maneuver the plane. It mentions specific controls. Example: "Once you got the new altitude you push forward on the yoke to level off then you pull back on the throttle to get the airspeed you need and once you've done that you press on the right rudder to make a slight turn and then pull back on the throttle to slow down
Instruction - Explanation - Abstract	I-E-A	An explanation that conveys a cause-and-effect relationship that does not mention a specific control(s). Example: "If you roll too quickly it stops the turn all at once"
Instruction - Explanation - Specific	I-E-S	An explanation that conveys a cause-and-effect relationship that does mention a specific control(s). Example: "pushing in on the throttle even a tad too quickly causes a sudden increase in airspeed"
Instruction – Pilot Critique	I-PCrit	A critique of the pilot's flying performance. Example: "But when you went to pull down, you pulled back and you started climbing."

Instruction – Pilot Critique – Target	I-PCrit+T	A critique of the pilot's performance with a target. Example: "At your static condition your letting it come back to level when it needs to be about a bar width below the horizon."
Instruction – Pilot Critique – Explanation	I-PCrit+Expl	A critique with an explanation. Example: "So during the climb there when you meant to make a little bit of a power reduction, you actually reduced the power too much and while you were concentrating on turning back to the heading, you pulled the pitch up too much which did two thing initially it slowed you down and increased drag coupled with your lower power setting made your climb rate go away. But for the most part about three quarters of that climb was completed with everything right where they were supposed to be so"
Instruction – Pilot Critique – Pilot Style	I-PCrit+PS	A critique of the style of flying the airplane. Example: "You keep overcorrecting your pitch."
	Qualifiers for Instructions	
Specific/Abstract	-A or -S	An abstract instruction does not mention a specific control or instrument. A specific instruction mentions a particular control or instrument.
Explanation	+Expl	An explanation given that conveys a cause-and-effect relationship for an instruction.
Target	+T	An instruction with a target value.
Steps	1,2,3,4	Indicates the number of steps a plan contains. This qualifier is specific to plans

Other

A statement that does not pertain to the flight task or instructions

Other – Instructor – Flight Comment	O-I-FC	A general comment about flying conditions or flying the airplane made by the instructor. Example: “It is a little bit windy today but it is not too bad”
Other – Student – Flight Comment	O-S-FC	A general comment about flying conditions or flying the airplane made by the student. Example: “Because I was following your directions”
Other – Instructor – Task Objective	O-I-TO	The task objective given to the student by the pilot. Example: “801 XRAY decrease your airspeed to 80 knots”
Other – Student – Task Objective	O-S-TO	The task objective repeated back to the pilot by the student. Example: “801 XRAY decreasing airspeed to 80 knots”
Other-Instructor-Reminder	O-I-Rem	A reminder to the student. Example: “We learned that a while ago.”
Other – Instructor – General Comment	O-I-GC	A comment made by the instructor that does not pertain directly to instructing or coaching. Example: “Easier said than done”
Other – Student – General Comment	O-S-GC	A comment made by the student that does not pertain directly to instructing or coaching. Example: “It is pretty sunny today”
Other – Student Comment – Request for an Explanation	O-SC-ReqE	A question that requests an explanation Example “Why am I losing altitude?”
Other – Instructor – Response to an Explanation	O-I-ResE	An explanation given to a question Example “because you didn’t add enough power”

Other – Instructor – Request for an Acknowledgement	O-I-ReqA	A question that evokes an acknowledgement Example: “Okay?” or “Are you focused on the turn coordinator momentarily to set the pitch?”
Other – Student – Response to an Acknowledgement	O-S-ResA	An acknowledgement usually in response to a request for an acknowledgement. Example: “Yes sir” or “Okay”

Global Qualifiers

Axis

Indicates the number of axes involved and applies to messages in all three categories (Coaching, Instructing, and Other)

Single Axis	+SA	A single axis
Double Axis	+DA	A double axis
Triple Axis	+TA	A triple axis

Delivery

Indicates the way a message is delivered applies to messages in all three categories (Coaching, Instructing, and Other)

Repeat	(R)	A statement in any of the above categories (command, plan, observation, etc…) that is repeated immediately after the initial statement is made. The statement does not have to be repeated word for word. Example: The instructor pilot telling the student pilot to lower the nose for a second time in a row. “Lower the nose to gain some airspeed”, “lower the nose”

APPENDIX D

A CODED FLIGHT TRANSCRIPTION

Code	Start of Utterance	Utterance Time (Sec)	Word Count	Speaker	Message
Admin	4:31				Reset for task 1
Admin	5:37				anchor 1 for task 1
O-I-TO	5:30	6	10	Instructor	CESSNA 3-4 FOXTROT reduce your airspeed to eight zero knots,
O-I-TO	5:37	3	12	Instructor	so you say it back, 3-4 FOXTROT reduce to eight zero knots
O-S-TO	5:41	3	7	Student	3-4 FOXTROT reducing to eight zero knots
C-CC-A+T	5:44	24	89	Instructor	And so you pull the power back slightly maybe to about 2000 RPM or so
C-FW+Expl	5:49			Instructor	and don't just pull the nose up abruptly because if you do you'll start climbing
C-CC-A+T	5:53			Instructor	so instead you just pull that power back to about what you know you're going to need maybe a little less
C-CC-A	5:58			Instructor	Stop that vertical velocity,
C-CC-A	5:59			Instructor	push down
C-COS	6:01			Instructor	you're almost 100 foot off
C-CC-A	6:02			Instructor	so you want to correct down to about 200 feet per minute.
C-FW	6:04			Instructor	No more than 200,
C-CC-A	6:05			Instructor	pull it back up
C-FW	6:07			Instructor	no more than 200 feet a minute
C-COS	6:13	9	34	Instructor	You're almost on your altitude now
C-CC-A	6:14			Instructor	so stabilize it

I-Impl	6:16			Instructor	leave the power back at about 2000 RPM initially until the airspeed goes down to the 80 knots
C-CC-A+T	6:21			Instructor	pull that power back to about 2000
C-IC-A	6:27	6	19	Instructor	Concentrate on trying to stay on your altitude
C-CC-A	6:30			Instructor	make a little correction downward
C-DW	6:31			Instructor	not much just barely correct down
I-PCrit	6:34	7	29	Instructor	staying on our heading fairly well you've even corrected to the right just a little bit, airspeed is just about where you want it,
C-COS				Instructor	so you're at 80 knots
C-CC-A	6:45	5	13	Instructor	Stop that turn,
C-CC-A (R)	6:47			Instructor	stop that turn,
C-CC-A (R)	6:48			Instructor	level those wings and stop that turn.
O-I-GC	6:52	12	52	Instructor	It is important that when you go back onto heading,
C-CC-A	6:53			Instructor	turn back to the left now
C-COS	6:54			Instructor	you're back on your heading
C-CC-A	6:56			Instructor	and add that pitch,
I-PCrit+Expl	6:57			Instructor	notice while you were distracted with your turn you lost your altitude because you let your pitch down,
C-CC-A	7:02			Instructor	start banking to the left,
C-CC-A (R)	7:03			Instructor	turn back to the left

I-Fly	7:06	9	26	Instructor	You don't need to use the rudders, I know, even in a real airplane for small turns like this you wouldn't do much with the rudders
C-CC-A+T	7:20	6	15	Instructor	Set your power to about 2050 RPM
C-COS	7:23			Instructor	half way, there you go.
C-COS	7:24			Instructor	Just under 2100.
C-CC-A+T	7:27	2	7	Instructor	Try to hold it stable on north.
O-I-GC	7:31	2	7	Instructor	Okay we are going to reset. Relax
Admin	7:32				End of Task 1

APPENDIX E

IRB APPROVAL LETTER

Mississippi State University

April 27, 2005

Stephanie Doane
INST
Mailstop 9630

Re: IRB Docket #05-050: Dynamics of instructor-student interactions

Dear Dr. Doane:

On April 27, 2005, the Mississippi State University Institutional Review Board for the Protection of Human Subjects in Research voted to approve your application titled "Dynamics of instructor-student interactions".

Please note the expiration date for approval of this project is April 15, 2006. If additional time is needed to complete the project, you will need to submit a Continuing Review Request form 30 days prior to the date of expiration. Additionally, any modification to the project must be reviewed and approved by the IRB prior to implementation. Any failure to adhere to the approved protocol could result in suspension or termination of your project. Please note that the IRB reserves the right, at anytime, to observe you and any associated researchers as they conduct the project, as well as audit research records associated with this project.

Please refer to your IRB number (05-050) when contacting our office regarding this project.

We wish you the very best of luck in your research and look forward to working with you again. If you have any questions, please feel free to contact me.

Sincerely,

[Unsigned version for electronic submissions -- printed 6/9/08]

Tracy S. Arwood
Director

Office for Regulatory Compliance
P. O. Box 6223 • 8A Morgan Street • Mailstop 9563 • Mississippi State, MS 39762 • (662) 325-3294 • FAX (662) 325-8776