Social Facilitation and Test Anxiety in Flight Simulation Training

Heidi Manijeh Mehrzad
Embry-Riddle Aeronautical University - Daytona Beach

Follow this and additional works at: https://commons.erau.edu/edt
Part of the Applied Behavior Analysis Commons, and the Cognitive Psychology Commons

Scholarly Commons Citation
https://commons.erau.edu/edt/105

This Thesis - Open Access is brought to you for free and open access by Scholarly Commons. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu.
SOCIAL FACILITATION AND TEST ANXIETY
IN FLIGHT SIMULATION TRAINING

by
Heidi Manijeh Mehrzad
B.S. Embry-Riddle Aeronautical University, 2006

A Thesis Submitted to the
Department of Human Factors & Systems
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Human Factors & Systems

Embry-Riddle Aeronautical University
Daytona Beach, Florida
Spring 2012
SOCIAL FACILITATION AND TEST ANXIETY
IN FLIGHT SIMULATION TRAINING

By
Heidi Manijeh Mehrzad

This thesis was prepared under the direction of the candidate’s thesis committee chair, Albert J. Boquet, Ph.D., Department of Human Factors & Systems, and has been approved by the members of this thesis committee. It was submitted to the Department of Human Factors and has been accepted in partial fulfillment of the requirements for the degree of Master of Science in Human Factors & Systems.

THESIS COMMITTEE:

Albert J. Boquet, Ph.D., Chair
Shawn Doherty, Ph.D., Member

Nickolas Macchiarella, Ph.D., Member

MS HFS Program Coordinator

Department Chair, Human Factors & Systems

Associate Vice President for Academics
First and foremost I would like to thank my parents. Mom, thank you for always believing in me. Your support throughout my academic career was the glue that held me together. Without you and our million phone conversations at all hours I would have never made it. Your love, words and dedication to my success throughout my life have kept me going and pushing myself to achieve better and greater things. You have been my idol, role model and inspiration to be everything that I want to be. Dad, thank you for always guiding me to the light at the end of the tunnel. Your never-ending optimism and faith in my capabilities make me want to be the best me I can be. Your humor and motivational speeches lifted me when I needed it. You always reminded me of my dreams and goals in life. Thank you both for your unwavering love and faith in me.

I would also like to give special thanks to my thesis committee. Dr. Boquet, Dr. Doherty and Dr. Macchiarella. Dr. Boquet, thank you for your guidance, wit, and of course, for always having the final joke to lighten up anything and everything. You have not only been my thesis chair, professor and mentor, you have been a friend throughout my journey. You have not only taught me the ins and outs on human factors, you have shown me how to build a career out of it and for that I will be forever grateful. I would not be where I am today without your ‘words of wise-dom’. And by the way: Thanks for always meeting me in the atrium and of course – thank you for giving me chirpees. Dr. Doherty, thank you for your continuous support and endless patience. Your dedication to your students is not only admirable; it is simply one of a kind. Thank you for answering every one of my e-mails and phone calls whenever I was across the country or across the globe. Without your meticulous comments and exceptional advice throughout the program I would have never made it through to the finish line. Dr. Macchiarella,
thank you for showing your support in my skills and abilities throughout my entire academic career – undergraduate to graduate. Your support is much appreciated and holds special meaning to me. Thank you for being a part of the beginning and the end of my flight at Embry-Riddle Aeronautical University. Special thanks also to Dr. French. You simply rock. Thanks to you I gave science another shot and ended up loving it. Only you can transform a former ‘F’ student in biology to an ‘A’ student in physiology and a certified EMT. Sorry about the fatigue though.

I would also like to thank Terry, Dave, Dane and the entire CRDM Human Factors Team. Not only did you guys make my year at Medtronic, Inc. a blast, you always lend me a helping hand when I needed it. Whether it was hours of review, dozens of pep talks, funny video clips or all of the above – whenever I asked, you were there to help, teach, and entertain. A thousand thanks for that. P.S.: Dane, if all fails, we can always work for clothes. Jodie aka Bertha, you deserve an award. You made Minneapolis everything I thought it would never be in -26 degrees.

Finally, I would like to thank all my friends and family – special thanks to Ande, Lauren aka Bob, Ben, Brian, Rafael, Adam aka Lt. Army, Angela, Nick, Pat, Natalie, Lally, Hilary aka Homeskillet and Martina. Our countless conversations in which you consistently reassured my belief in myself and showed your support and faith in me were the final pieces to my thesis puzzle. Friends are the family you chose, and I have to admit – I chose wisely. I could have not picked a better group of people to love and support me. Thank you everyone for never allowing me to give up on myself and always guiding me along the way when I needed it.
Abstract

The power of an individual’s presence to influence another’s behavior is apparent in many aspects of human interaction. Performance competitions, team working environments, as well as instructor-student relationships, to name a few, have demonstrated the potential to produce social facilitation and interference effects on the participating individuals behavior and performance (Hazel, 1978). These exact effects could support or hinder learning and affect training outcomes in a training environment. The goal of the present study was to examine the effects of the presence of an instructor (audience) on task performance and test anxiety levels of an individual during the early stages of flight simulation training. While trends indicated that a stronger performance was achieved by test participants in the no audience, low test anxiety and simple task groups, results revealed that only task complexity had a direct relationship with errors and time, causing test participants to commit more mistakes and require more time as complexity increased.
# Table of Contents

Abstract .......................................................................................................................................... vi

Table of Contents .......................................................................................................................... vii

List of Figures ................................................................................................................................ ix

List of Tables .................................................................................................................................. x

Introduction ..................................................................................................................................... 1

Social Facilitation ....................................................................................................................... 4

Audience Effects ......................................................................................................................... 7

Test Anxiety .............................................................................................................................. 15

Research Questions ................................................................................................................... 24

Methods......................................................................................................................................... 26

Participants ................................................................................................................................ 26

Materials ................................................................................................................................... 27

Design ....................................................................................................................................... 30

Procedure .................................................................................................................................. 33

Results ........................................................................................................................................... 35

Exploratory Analysis ................................................................................................................ 36

MANOVA................................................................................................................................. 42

MANCOVA.............................................................................................................................. 44

Discussion .................................................................................................................................... 46

Limitations ................................................................................................................................ 48

Future Research ........................................................................................................................ 49

Conclusion .................................................................................................................................... 50
List of Figures

Figure 1. Screenshot of the Cessna 172 with G1000 cockpit in Flight Simulator X.............. 28
Figure 2. Flight Simulator X workstation (including flight controls)................................. 29
Figure 3. FS Recorder add-on module within MS Flight Simulator X.................................. 29
Figure 4. FDR 1.0 command line tool converting .frc files to .txt files................................ 30
Figure 5. Flight pattern of a rectangular course maneuver................................................ 32
Figure 6. Means for the dependent measure of total errors............................................... 38
Figure 7. Means for the dependent measure of time......................................................... 40
Figure 8. Means for the dependent measure of test anxiety change value scores.................. 42
List of Tables

Table 1.  Descriptive Statistics – total errors. ................................................................. 37
Table 2.  Descriptive Statistics – time. ......................................................................... 39
Table 3.  Descriptive Statistics – test anxiety change value scores. ............................. 41
Introduction

A simulator is generally a device or software that attempts to create a real world environment by recreating characteristics of such (Beaubien & Baker, 2004). Flight simulation enables students to learn aircraft performance, experience aerodynamic effects, and perform flight maneuvers immediately and without risk (Embry-Riddle Aeronautical University, 2011). Simulation is a powerful training tool because it allows the trainer to systematically control the schedule of practice, presentation of feedback, and introduction or suppression of environmental distractions within a safe, controlled learning environment (Beaubien & Baker, 2004). Furthermore, the trainer can freeze and reposition each device instantly, as well as vary traffic loads within a flight scenario. These features allow 100 percent time-on-task in the simulator that is not possible in actual flight (Embry-Riddle Aeronautical University, 2011).

Aviation simulation has been of interest for psychologists, engineers, and computer scientist for decades. With the advancements in technology and engineering, simulators have evolved from simple input machines to being able to virtually replicate any real-world artifact. Advanced technologies, such as virtual environments, have enabled the aviation industry to immerse the pilot into advanced flying experiences and training scenarios. In parallel with the advancements in technology and the development of high fidelity virtual environment flight simulators, psychologists, cognitive engineers and behavioral scientists have focused on understanding the acquisition of knowledge, skills, and attitudes (KSAs) in the aviation community. Research in the areas of training, practice, feedback and performance measurement has also progressed at a growing rate. However, there seems to be a disconnect in applying the knowledge of these findings to the aviation simulation training environment. In essence, aviation training has not evolved, simulators and their technology have (Salas, Bowers, & Rhodenzier,
1998). The need to close this gap between technology and new findings in the behavioral sciences is now more apparent than ever. The best simulation in the world does not guarantee learning if it is not used and applied effectively; a simulator is just a tool (Salas & Cannon-Bowers, 2001). More advanced simulation does not ensure transfer of learning. Simulators have often been developed without the consideration of the learner or how they learn, including their struggles during the learning process (Salas, Bowers, & Rhodenzier, 1998).

Simulators range in cost, fidelity and functionality. It is often believed that the higher the equipment and environment fidelity, the better the training will be. Beaubien and Baker (2004) discussed one of various typologies of simulation fidelity in simulation training (adapted from Rehmann et al, 1995) by emphasizing that while all three simulation fidelity components: environment fidelity, equipment fidelity, and psychological fidelity are essential to the successful training process and outcome, it should be recognized that the psychological component should be considered a most critical requirement. Without temporarily suspending disbelief, trainees are unlikely to behave in the simulation as they would in the real world and as such the training could result in having minimal application value post training completion (Beaubien & Baker, 2004). Therefore ensuring that the optimal psychological conditions are provided is most essential.

During a flight lesson in the simulator a student typically sits in the left seat of the ‘cockpit’ and the instructor sits in the right seat, just as in the actual aircraft. In some instances the flight instructor does not sit in the ‘aircraft’ with the student, instead is positioned behind the aircraft within the flight training device controlling the software, overseeing flight tasks, route, and terrain. Sitting with the student in the flight simulator ‘cockpit’ is generally considered standard procedure, as this allows the instructor to observe the student during training. However,
observation from such close proximity could represent a psychological burden on the student, and as such could potentially alter the student’s performance, hindering the learning process and altering the training outcome.

The power of another individual’s presence influencing another’s behavior is apparent in many aspects of human interaction. Performance competitions, team working environments, as well as instructor-student relationships, to name a few, have demonstrated the potential to produce social facilitation and interference effects on the participating individuals behavior and performance (Hazel, 1978). Indeed, the presence of another has been shown to be sufficient enough to produce effects on an individual – interaction between individual and observer in any kind of behavior is not needed (Zajonc, 1965). Such effects could hinder learning and affect training outcome. Ensuring optimal psychological fidelity in the simulator is therefore highly critical, as the aim of the individual lessons in the initial stages of flight simulation training are mostly to introduce new and unrehearsed tasks and not typically designed to measure performance or aimed at the evaluation of such.

Considering the initial stages of flight simulation training, it is to be questioned whether such tasks should be ‘practiced’ and ‘learned’ in the presence of an instructor, as such unavoidably with the status of an authoritative standing, facilitates an atmosphere of evaluation and increases performance anxiety (Yantz & McCaffrey, 2005). In addition, ‘in the early stages of learning the most dominant responses may be incorrect ones. An audience may well raise the arousal level of an individual performer. When this happens, the degree of proficiency in the skill is important. If the performer is competent and has full mastery of the activity, the dominant responses will be enhanced and a subsequent improvement in performance should result. Conversely, if the skill is at a low level then the emission of dominant responses in this instance
may well result in impairment of skills rather than improvement (Underwood, 1976). Given that the aim of the flight simulator lesson is to establish familiarity with new tasks in the cockpit in order to increase the level of confidence and mastery with such, rather than create a performance evaluation situation for the student pilot; one must question whether the presence of an instructor in the flight simulator is beneficial or impairing to the learning (and learning outcome) of the new student pilot. In view of the question of instructor presence affecting student pilot learning and performance, an investigation into the optimal flight simulator environment (with respect to instructor presence or absence) is needed.

Social Facilitation

For over a hundred years researchers have been interested in the study of social influences. Focus has been put on the impact of the performance of an individual or group in the presence of an audience. Later referred to as social facilitation, the study of social presence on the performance of another is considered to be one of the oldest experimental paradigms of social psychology (Zajonc, 1965). The term facilitation refers to the early observations that performance was enhanced when others were present (Aiello & Douthitt, 2001). The paradigm itself stemmed from the original experiments of Triplett (1898), who investigated the consequences of individuals’ behavior and performance during the presence of others. Facilitating task performance through the mere presence of others demonstrated that the ‘bodily presence of another contestant participating simultaneously […] serves to liberate latent energy not ordinarily available’, enhancing such individual’s performance on the task executed (Triplett, 1898). In his first experiment Triplett demonstrated that cyclists completed their races faster when they were racing amongst other cyclists, and even faster when raced against each other in competition. Triplett then conducted a follow-up study in which he investigated the times of
children turning a fishing reel, when alone, and when in presence of other children. Again, Triplett found support that performance was enhanced when in the presence of others, as the children who were reeling alongside others reeled the fishing reel fastest. Triplett concluded that the mere presence of others allowed enhanced performance due to two major concepts: the presence of others simply evoked a competitive instinct, and that the sight of another performing the exact same task provoked the individual to perform the task faster (Aiello & Douthitt, 2001).

Triplett’s first observations of social facilitation in 1898 sparked interest for other researchers to explore the phenomenon of facilitation leading to performance output changes by separating competition (as such was an obvious condition in Triplett’s experiments) from social facilitation. In Allport’s 1924 publishing of *Social Psychology* he states that social psychology focuses especially on social situations, as individuals react to and serve as stimuli for others. Nonetheless it grounds itself on the assumption that all social behavior can be explained in terms of the principles of individual psychological functioning (Allport, 1924). Previous to his publishing in 1924, Allport already explained that ‘our first task has been to study those aspects of the individual which are destined to direct and control his behavior within the social sphere’ (Allport, 1920). In his study, a comparison between the mental processes of individuals when alone with their reactions to similar and equivalent stimuli when a member of a "co-working or co-feeling" group was investigated (Allport, 1920). His findings demonstrated that the presence of a co-working group is distinctly favorable to the speed of process of free association, and that not only are there individual differences in susceptibility to the influence of the group upon association, but that there is a tendency for slow individuals to be more favorably affected in speed of task performance by the group co-activity. Interestingly he also found that in its temporal distribution the benefit of the group effect was at its highest in the beginning of a task.
and at its lowest towards the end of the task performed. While these results mentioned here are by far not a complete representation of all of Allport’s findings and discoveries, these conclusions of the experiment conducted in 1920 provided some of the fundamental drivers in social psychology and the phenomenon of social facilitation.

While scientists such as Allport focused on the fundamental basics of social psychology investigating behavior and conciseness and its influences on performance in social settings, scientists such as Gates (1923) focused on the effects of ‘small’ and ‘large’ audiences on the ability of individuals to perform in test environments, specifically on various different cognitive tests presented during the experiment. Although Gates stated that the findings in score differences on the cognitive test performed are too small to be statistically reliable, the results suggested that better performing individuals were slightly more disturbed than the others by the presence of spectators. Such early findings in the theories of social facilitation demonstrated that the influence of facilitation on performance could be strongly correlated to the level of expertise of the individual performing – novice versus expert – as well as the complexity of a task and the needed concentration and mental arousal level of the individual to perform the task at an inferior, satisfactory or superior level. Taking different task complexity levels into consideration, Pessin (1933) studied serial learning and recall of nonsense syllables under different conditions in relation to social facilitation. In his study all subjects learned three lists of syllables under three different conditions: (1) entirely alone, (2) with an audience (spectator) present, and (3) in the presence of combined visual and auditory stimuli (visual and auditory stimuli consisted of the simultaneous flashing of a 150-watt light and the sounding of a buzzer, introduced by means of a mercury metronome that beat 54 time per min). They all returned in intervals of 1, 2, and 3 days and relearned the original list under the same conditions. The results were scored in terms of
errors (number of errors made) and repetitions (required to reach the point of one perfect reproduction). Results indicated that subjects learned best when alone, next best when in presence of an audience, and worst when in the presence of visual and auditory stimuli. However, retention of material learned was greater amongst participants who were in the presence of either an audience or combined visual and auditory stimuli, and poorest in the alone condition (Pessin, 1933). Pessin’s results indicated that while an individual may perform greater with an audience (of either kind) when performing a task already familiar with, their learning abilities were greater when alone, unsupervised and unobserved. Such findings allowed scientists to take their research on social facilitation, more concrete on audience effects, to a new territory. It enabled them to investigate deeper into the relations between audience effects on performance and different task complexities, levels, and individuals’ expertise level.

**Audience Effects**

Research into the phenomenon of social facilitation can be divided into two major components – audience effects and co-action effects. While the audience effect is the impact that a passive audience has on a subject performing a task, the co-action effect explains the impact caused by several individuals – in the presence of one another – engaging in the same task. For the purposes of this research the focus will be on the experimental paradigm of audience effects. As stated by Ganzer (1968) social facilitation refers to the effect upon the behavior of an individual, derived from the presence of other individuals (Allport, 1920). One aspect of social facilitation concerns the effects that the presence of passive spectators have on the behavior of the person being observed, and the learning or performance situation they are experiencing (audience affects). In 1965 Zajonc conducted a review of the social-facilitation effects: audience effects and co-action effects, in an attempt to organize previous studies with the assumption that
the presence of others did not necessarily only enhance or impair task performance due to the
mere presence of others, instead it affected the performer’s level of general drive in an enhanced
manner. This suggested that higher drive levels allowed the enhanced emission of well-learned
responses, and impaired the emission of newly acquired or unrehearsed responses. Zajonc stated
that up until then most studies contradicted one another, in that some stated enhanced
performance due to audience effects and others showed performance decrease when tasks were
performed in the presence of spectators. No focus was made on arousal states or differences
between learning and performance. He generalized that the mere presence of others, whether as
an audience or as co-actors in the task performed at the time, demonstrated the emission of
dominant responses, as when in an increased arousal or higher drive level. He noted specifically
that the mere presence of others seem to impair the task of learning, but in contrast enhances the
performance in responses that are already well established and are performed at a mastery level
(Zajonc, 1965). The theory of social facilitation, up until Zajonc’s review, stated that the mere
observation of an audience on an individual’s task performance, not solely the individual’s
abilities, could alter the performance outcome of the task attempted, as the mere presence of
another is a source of nonspecific and nondirective arousal that enhances the dominant responses
of the performer. After Zajonc’s review, a correlation between social facilitation and arousal
states and the emission of dominant responses was accepted. The theory moved forward with the
implication that task performance under observation would allow individuals to perform on an
increased level for tasks already mastered at an expert level to begin with, and perform on a
decreased level on tasks that were complex and challenging to the individual (Ganzer, 1968). In
1982 Guerin and Innes took a ‘new look’ at Zajonc’s hypothesis as he originally stated that only
‘indirect and scanty’ evidence supported the increased arousal levels due to the mere presence of
others. Guerin and Innes retained Zajonc’s formulation (whereby the direction of any effect upon performance may be predicted from a knowledge of the task characteristics) with the addition that while evaluation apprehension and a need for social comparison will play an important role in maintaining a drive state in a performer before an audience, such formulation reintroduces the role of a fundamental process in the understanding of behavior in the presence of others to the study of social facilitation (Guerin & Innes, 1982).

This addition allowed the study of social facilitation to be expanded in correlation to drive theory. It demonstrated a clear involvement of evaluation apprehension, monitoring, and a need for social comparison in maintaining a drive state in an individual performing in the presence of either an audience or co-actors (Guerin & Innes, 1982). While the term social facilitation has been defined in numerous ways, and since its inception in 1898 by Triplett has evolved from the simple definition of the ‘effect of audiences influencing individual task performance on a positive level’. Social facilitation no longer is ‘just’ considered an enhancement feature in task performance, instead can be clearly defined as the tendency of an individual exhibiting enhanced performance on simple tasks and inhibit performance on complex tasks in the presence of observers (Constantinou, Ashendorf, & McCaffrey, 2002).

**Audience effects in learning and testing environments.** A lot of progress in the research of social facilitation has been made since the early experiments of Triplett. More and more tangents are explored by researchers in an attempt to fine-tune and pinpoint the exact impact of audiences, spectators, co-actors, and third party observers during task performance and learning. In recent years, focus has been given to audience and third party observers and the effects of such in test environments (Horwitz & McCaffrey, 2008). Social facilitation denotes that the presence of an audience creates an atmosphere of evaluation, giving the individual
performing the task a feeling of elevated pressure and performance anxiety. It has also been found that the mere presence of others can create a source of arousal. While scientists such as Yerkes and Dodson (1908) (Yerkes-Dodson Law) focused on the differences in physiological and mental arousal levels for optimal task performance, Zajonc focused on the skill and expertise level an individual performed a task with; concluding that if a task was well mastered and rehearsed an audience would facilitate performance (and vice versa for unrehearsed tasks impairing performance and learning), as such would enhance the emission of dominant responses. Following Zajonc’s in-depth look at arousal levels in relation to social facilitation, Martens (1969) later conducted an experiment in which he investigated the effects of an audience on learning and performance of a coincident timing skill in relation to anxiety, and the differences between subjects of high and low anxiety. In accordance with Zajonc’s drive theory of social facilitation, Martens’ results indicated that error rate amongst subjects learning in the presence of an audience was increased, and that such subjects required significantly more time to learn the task given. In addition, Martens was able to introduce another variable into the picture by differentiating between subjects of low and high anxiety, as he found that subjects of high anxiety were faster in learning compared to subjects of low anxiety (Martens, 1969).

Underwood (1976) explains that tension levels, caused by the presence of another in a competitive framework, are dependent on each individual, as well as the level of the skill. The clear distinction between a learning task and a performance task is made in his study. Underwood clarifies that in a learning task an increased rate could be the result when subjects approached the task as a competitive challenge, and a decreased rate of learning occurred when subjects were too anxious to complete the task. In relation to performance he concluded that such an inverse behavior could result to the performance level of individuals. He expressed that low
anxiety level individuals would be more likely to perform on an increased level when in the presence of an audience, facilitating a competitive level, while subjects of high anxiety levels would tend to demonstrate no improvement in the same conditions (Underwood, 1976).

In 1958, Henry A. Landsberger, performed a data analysis on experiments performed between 1924 and 1932 by Elton Mayo, at the Hawthorne Works near Chicago. Here audience effects were discovered by chance, as the study was originally commissioned to determine if increasing or decreasing the amount of light workers received increased or decreased worker productivity. The researchers found that productivity increased due to attention from the research team and not because of changes to the experimental variable (light). Later coined as the *Hawthorne Effect*, the phenomenon was defined as ‘the process where human subjects of an experiment change their behavior, simply because they are being studied’. Individual behaviors of workers may be altered by the mere fact of being observed itself, rather than the effects of the environmental conditions or other external variables being changed (Landsberger, 1958). Proponents of the Hawthorne effect say that people who are singled out for a study of any kind may alter their performance or behavior not because of any specific condition being tested, but simply because of all the attention they receive through the simple observation of an audience (Hansson & Wigblad, 2006). While opinions are split on the Hawthorne effect, the phenomenon itself is yet another supportive argument that human performance, when simply being ‘observed’ by another during the execution of a task given, can show effects of social facilitation or interference. Concentrating on the subject of social facilitation in learning and testing environments, one study tackles the problem of the effects of the presence of an instructor in a learning environment and how such drives the student behavior and response. Kushnir (1986) analyzed, in terms of social facilitation and stress theories, the effects of instructor presence on
student nurses’ behavior. He rationalizes that social facilitation theories imply the existence of a paradox in the training and learning environment, as the aim of teaching is to reduce error, yet such error rate is negatively affected when students (student nurses) are in the presence of an audience (instructors). He explains that the presence of an authoritative figure only creates more tension and cause for increased emotional reactions in the student nurses’ behavior. This could lead to negative consequences in the students’ performance. As a student in most cases is still in the learning phase of most tasks during observation, (especially at the beginning of training) Kushnir elucidates that the presence of an authoritative audience is most likely to impair the student nurses’ learning. Looking at different theoretical processes, which mediate audience effects on performance, Kushnir considered evaluation apprehension, objective self-awareness, self-presentation, feed-back, and instructors’ attitudes toward mistakes. He explains that in such a critical learning environment as healthcare, in particular the training of nurses where room for errors is small, an increased atmosphere of evaluation and pressure is present. In his investigation he surveyed 28 second-year female nurses and asked them to provide a self-written report, describing in detail an interpersonal encounter, which they found to be stressful, involving themselves and a person of authority. His results were categorized into three categories: (1) cognitive processes, (2) behavior and actions, and (3) emotional reactions. Overall a significance of the instructor’s role in nurses’ training was found. The findings revealed that 71% of all participants referred to incidents with their instructors as the most stressful encounters during their training, suggesting that perceived evaluation through the presence of an instructor could be a source of emotional stress. Further, it was demonstrated that most of the stressful encounters occurred in novel situations, during initial phases of their learning and training. Kushnir’s study signifies that social facilitation effects may be most intense during the early stages of learning,
and involve cognitive, emotional, and behavioral levels; ultimately leading to an increase, rather than a decrease, in error rate when learning is performed in a perceived evaluative and observed environment. He states that his findings are consistent with two hypotheses: evaluation apprehension and self-presentation, as the first poses fear of failure and the second fear of embarrassment as the intervening variables responsible for audience effects (Kushnir, 1986). This study stands out as it investigates social facilitation and interference effects in a learning environment, just as the present study is by investigating the effects of observation and test anxiety in a flight simulation learning environment.

In a similar study, conducted by Rosenbloom et. al in 2007, the success rates on a practical driver’s license test with and without the presence of an audience (another testee) during the examination were investigated. While Kushnir concentrated on the audience effects caused by instructors being present, Rosenbloom et. al concentrated on the audience effect caused by the presence of another testee. Their study focused on the aspects of social facilitation, gender differences in driving performances, and differences in testing procedure. Their experiment involved the use of archival data of 1151 individuals, in which 802 persons performed their practical driver’s license test in pairs (in the presence of another testee), and 349 were tested alone. All students had passed medical and theoretical tests prior, and the success rates were analyzed with the assistance of the results of the practical driving test portion, which were obtained from the Ministry of Transport in Jerusalem. Looking at the audience effect on performance and results in relation to the effect of the presence of another testee during the practical test portion, their findings showed that both females and males had a higher success rate of passing the practical test when the student was tested without the presence of another testee (Rosenbloom, Shahar, Perlman, Estreich, & Kirzner, 2007). This study expresses the negative
effects of audiences in testing environments, as students’ task performance is significantly impaired due to the mere presence of another testee. In addition, it implies that the mere presence of another person can affect the performance level of an examinee and not necessarily the ‘status’ of the audience or person observing the individual performing the task.

While it is understood that students, in general, need to develop mechanisms to cope with the presence of evaluating individuals; results as such shown in these studies suggest a strong need for ‘unobserved’ study lessons in the initial stages of a student’s training and learning curriculum. It allows the student familiarization with the task, develop a higher confidence level, and allow the needed mastery with it to support the emission of a positive and correct dominant response when faced with an instructor or examiner in an advance learning or test environment such as that found in the cockpit training environment. These audience effects are not limited only to real world environments; they have also been researched and supported in virtual environments, such as simulation environments. Application of previous social facilitation findings to modern technology enabled the discovery of performance differences in virtual learning environments. In a computer study performed by Hall and Henningsen (2008) individuals were tested to determine whether the presence of a computer icon during a typing task could produce social facilitation effects. Task difficulty and the presence and absence of a computer icon, such as the Microsoft Word’s Clip were considered. Results provided some support that the mere presence of a computer icon can produce the perception of performance observation and therefore creates an atmosphere of social facilitation and therefore may influence task performance (Hall & Henningsen, 2008).

As performance anxiety increases while participants are observed and evaluated during their performance and completion of various different tasks and processes, the effects of social
facilitation on performance are often connected to tension levels in individuals (Underwood, 1976) and anxiety levels (Martens, 1969). As Martens (1969) already discovered, differences in error rates and learning times were demonstrated between subjects of high and low anxiety. Therefore anxiety levels should be weighed when conducting a social facilitation study, as such could represent an influencing factor as well. Considering this interconnection, the question arises what differences should be considered when investigating the effects of observation and mere presence on another individual’s performance if such individual already suffers from performance influential factors such as anxiety. As the present study will focus exclusively on the testing and learning environment (of flight simulation training), a closer look at the effects of anxiety, specifically test anxiety, is needed to understand the relationship between social facilitation and test anxiety and their resulting effects on performance. As these summations of previous findings here demonstrate, research on the effects of social facilitation on task performance has been conducted from many different approach angles. Some studies focused on task difficulty, some on skill level, some on the difference between learning and performance, some focused on the learning environment itself (real-world vs. virtual), and yet some investigated the sole presence of an audience on task performance in general; yet all studies showed strong support that effects of social facilitation cause inhibited task performance and are present in many different aspects and cannot be underestimated, especially when investigating learning and testing environments.

**Test Anxiety**

Spielberger (1983) defined anxiety as an emotional state characterized by subjective feelings of tension, apprehension, nervousness and worry, as well as its effects on the nervous system. He differentiates between state and trait anxiety. State anxiety refers to the emotional
reaction to a stressful situation and the feelings associated with such at that moment; it can vary in intensity and fluctuates over time. Trait anxiety refers to the relatively stable individual differences in anxiety proneness, more specifically, to differences in the person’s tendency to respond to situations of perceived threat with increased levels of state anxiety. Trait anxiety is found to have direct influence on achievement through its influence on state anxiety (Glanzmann & Laux, 1978). Heightened state-anxiety levels are evoked by the stress of an event or a specific situation, but the magnitude depends on the amount of stress and the individual’s interpretation of the situation as personally dangerous or threatening.

When the characteristics of anxiety are linked to academic or evaluative situations, this identifies the concept of test anxiety. As our society is more than ever evaluated by performance, individuals have become more aware of the implications and consequences of their test performance. Test anxiety has become an omnipresent issue in schools as well as on college campuses. Test anxiety is considered to be a situation specific form of trait anxiety, which heavily depends on the student’s perception of a particular test or evaluation situation as personally threatening (Sarason, 1978). Persons high in test anxiety tend to perceive examinations and situations of evaluation as highly threatening and disturbing. Concern, respecting the evaluative event, disturbs their mind and keeps it in a state of painful uneasiness. Fear of failure becomes the most prevalent concern of the highly anxious person. Whereas the less anxious person can focus on the task he or she is being evaluated on, the highly anxious person tends to focus on worrying about their performance and the failure of such (Sarason, 1978).

Repeatedly research has demonstrated that individuals high in test anxiety experience performance decrements when in evaluative situations. Such evaluative situations evoke task-
irrelevant and self-centered worry that hinders optimal performance (Spielberger, Gonzalez, Taylor, Algaze, & Anton, 1978). Individuals with high test anxiety demonstrate physiological reaction patterns similar to worry. Worry is a cognitively demanding activity marked by self-preoccupation, self-depreciation, and concern over the consequences of poor performance (Sarason, 1978). Liebert and Morris (1967) conceptualized test anxiety as consisting of two major components, one cognitive and the other emotional. The cognitive component, labeled as Worry, is best conceptualized as the cognitive activity associated with negative thoughts and concerns about self-evaluation, performance, test taking and the potential consequences relating to such. The emotional component, labeled as Emotionality (also referred to as the physiological component), represents the autonomic arousal and somatic reactions to testing. It manifests itself in discomfort, muscle tension, sweating, elevated heart rate, and nervousness. The distinction between these two major anxiety components was supported by their findings that Worry was negatively related to performance expectancy while Emotionality was unrelated. Following up this study Morris and Liebert (1970) tested the implications of the Worry-Emotionality distinction for students’ reactions to an actual classroom examination. They did this by investigating the relationship of Worry and Emotionality components of test anxiety to pulse rate, performance expectancy, and actual test grades. Most critical to the present study were their findings on the relationship between the two anxiety components and test performance. As in the previous research of Morris and Liebert (1969), their study revealed that Worry was more highly negatively related to grade than Emotionality, elucidating that the Worry component is mostly responsible for ‘anxiety effects’ on academic or intellectual performance. In their discussion of results Morris and Liebert stated that ‘the critical point is that the distinction between these components should be kept in mind when designing studies to investigate the relationship
between “anxiety” and performance’ (Morris & Liebert, 1970). Spielberger (1983) supported this point as he stated that ‘an adequate factor study of a test anxiety scale should include measures of emotionality and worry’. In the present study this distinction is ensured, as the test anxiety inventory (TAI) includes both subscales: Emotionality (E) and Worry (W). This will allow the comparison of the Worry component and participant performance to analyze the weight of the ‘anxiety effects’ in the overall performance during the study, if required.

**Test anxiety in learning environments.** Early research into test anxiety in the academic environment showed that college students with high test anxiety performed more poorly than students with low test anxiety when in evaluative situations. Sarason and Mandler (1952) concluded from a series of studies with college students that decrements in performance could be attributed to the arousal of task-irrelevant responses (in student during test-situations) and the students’ reaction to test associated stress caused the emission of negative self-centered responses, and that such was higher in student of high test anxiety. The researchers discovered that such anxiety-mediating task-irrelevant responses could not align with optimal performance and therefore were incompatible with performance facilitation. As subjects high in test anxiety demonstrated a higher rate of such anxiety-mediating and task-irrelevant thoughts and responses than subjects low in test anxiety it was further concluded that subjects high in test anxiety presented poorer performance on intelligence test and learning tasks than subjects low in test anxiety (Sarason & Mandler, 1952).

But poorer performance is not necessarily automatic just because a person is test anxious. Interactive effects of test anxiety, test difficulty, and feedback have been investigated and found to be valid, suggesting that test difficulty and feedback, in addition to test anxiety, play a critical role in performance outcome. In a study conducted by Rocklin and Thompson in 1985 the
relationship and effects of item difficulty and test anxiety were examined. In addition, the
researchers examined the effects of item-by-item feedback. In their study participants completed
test anxiety measures establishing their Test Anxiety Scale (TAS) score and were then randomly
assigned to receive either a hard or an easy test. Each group (low, moderate, and high TAS) was
divided into subgroups in which one group of participants received feedback on an item-by-item
case throughout the test they were taking (easy or hard). The results indicated that students
differing in anxiety respond differently to verbal ability tests depending on test difficulty and
feedback given. The researchers’ findings showed that low anxious students did significantly
better on the hard test than they did on the easy test. Just the opposite was true for the moderate
anxious students. Finally, while the high anxious students seem to have poor performance on
both tests compared to low anxious students, they demonstrated slightly better performances on
the easy test than they did on the hard one. In all cases feedback increased performance (Rocklin
& Thompson, 1985). These findings allow the assumption that a high anxious person will display
superior performance when presented with a simple task in a learning environment, as opposed
to a complex task. Further, it allows the implication that with increase in task complexity in a
learning situation a high anxious person would demonstrate a decrease in performance,
compared to a person low in anxiety, who would demonstrate an increase in performance.

Test difficulty has also been closely related to an increase on the Worry scale. In a study
by Hong (1999) data revealed that students who perceived a test as more difficult the higher the
measurement of Worry and Emotionality was (compared to their peers who thought of the test as
less difficult). While perceived test difficulty or task complexity does not seem to directly affect
test performance in this study, the indirect effect through mediation of Worry was most certainly
existent. As students’ perception of test difficulty increased their arousal of Worry increased,
which in turn affected their performance (Hong, 1999). Regardless of test difficulty and feedback during an examination, anxiety is one of the major predictors of academic performance. Subjects high in test anxiety are more susceptible to disruptions of learning and more prone to poorer performance than subjects of low test anxiety. These students have a reduced memory span, lose concentration, lack confidence, and have poor reasoning power (Vitasari, Wahab, Othman, Herawan, & Sinnadurai, 2010).

While under neutral conditions high and low test anxious persons perform comparably, when in a learning or evaluative situation, high test anxious persons experience a problem of intrusive and interfering thoughts which diminish attention to and efficient execution of a task. When an evaluative component, such as an audience, is introduced, a red flag goes up in highly test anxious persons and they in turn perceive the situation as highly threatening and personally salient (Sarason, 1984). For these students Worry increases and thoughts of failure and self-doubt take over and dominate their thought process, interfering with their concentration on the task.

Students who encounter failure most often experience self-devaluation due to the high value placed on success and achievements in their lives. Such devaluation can deter future efforts as these students will internalize the failure experience with their level of abilities. As these encounters with failure multiply and the learner is faced with more situations in which attribution of failure are internally based, the likelihood that future tests will be considered threatening rather than simply challenging also increases (Schwarzer & Jerusalem, 1992). Results from a study conducted by J. C. Cassady in 2004 further support these findings as the data elucidated that students with high-cognitive test anxiety (worry component of test anxiety) had low-self efficacies related to testing. This association was validated by their perception of tests as
threatening and their attributions of helplessness over controlling performance outcomes due to their anxiety state (Cassady, 2004).

**Test anxiety and audience effects.** It seems probable that if an audience can be perceived as an evaluative element, persons high in test anxiety would perform poorer in the presence of an audience than persons low in test anxiety. This exact relationship was investigated in a study by Victor J. Ganzer in 1968. His study was directed at clarifying the relationship between test anxiety and audience presence in connection to learning and retention in a serial learning situation. Using a modified version of the Test Anxiety Scale (TAS) developed in 1962 he randomly assigned subjects of low, middle, and high TAS values to either the no-observer condition (NO) or the observer condition (O) – high TAS values representing persons high in test anxiety and low TAS values representing persons low in anxiety. Over the course of his 2-day study he would instruct subjects of both groups to learn and recall random lists of words. His results suggested that those subjects attempting to learn new material when in the presence of an audience performed poorer than subjects who learned alone. In addition, audience presence was more detrimental to high anxious subjects than low anxious subjects. For subjects high in anxiety this effect was most detrimental during the initial stages of learning (as well as the later stages leading up to performance evaluations). Another factor in relation to high anxious test participants and task performance was revealed when interpreting error rates from his data, as Ganzer found that high text anxiety subjects in the observer condition made significantly fewer correct responses than any other group of subjects. This finding provides support that persons high in test anxiety will demonstrate task performance with an increased error rate than persons low in test anxiety. Ganzer stated that ‘high scorers respond to threat with habitual, personalized responses of a self-deprecatory, critical nature. These self-preoccupations
are essentially task irrelevant and interfere with efficient learning and performance. […] low
scorers do not respond in this manner and may be expected to react to threat or stress with
increased effort and attention’ (Ganzer, 1968). This is the case if one assumes that an observer or
an audience constitutes a somewhat threatening situation, especially when the observer or
audience is perceived by the subject as someone of an evaluative nature, such as an instructor or
superior. Finally, even subjects in the high TAS category showed a more efficient and correct
performance in the no-observer condition than high TAS subjects in the observer condition.

In another study Horwitz and McCaffrey (2008) investigated whether performance on
tests of executive function is similarly impaired by the presence of an audience, as it is on tests
of effort, attention, concentration, learning, and memory. In their study each participant was
administered five tests of executive functioning in the following order: (1) Fear of Negative
Evaluation scale (FNE), (2) phonemic (letters F, A, and S) and semantic (animal naming) verbal
fluency tests, (3) Trail Making Test (TMT), parts A and B (4) Tactual Performance Test (TPT),
and (5) State–Trait Anxiety Inventory (STAI). Their findings revealed that the presence of a
third party observer is associated with poorer performance on certain tests of executive
functioning, adding to previous literature on the negative effects of audience presence on test
performance. In addition Horwitz and McCaffrey set out to examine the interactions among
observation and state and trait anxiety. Although there were no significant findings associated
with state anxiety, they discovered an association between performance and trait anxiety,
suggesting that the presence of an audience differentially affects performance among individuals
high and low in trait anxiety. When alone, high anxiety subjects performed over a half-standard
deviation above the normed mean while high anxiety subjects being observed performed over a
half-standard deviation below this mean. These results suggest that individuals high in trait
anxiety may be especially vulnerable to the potential negative impact of an added stressor such as an audience. While these results were derived in a setting investigating the effects of audiences on subjects performing tests on executive function and test anxiety is considered to be a form of trait anxiety, these findings again provide support that an audience seems to negatively affect performance in testing environments, most significantly when the test subject is high in test anxiety.

**Measuring test anxiety.** As a self-reporting psychometric scale the test anxiety inventory (TAI) was initially developed by Spielberger, Gonzalez, Taylor, Algaze, and Anton in 1978 to measure differences in test anxiety (TA) as a situation-specific trait. Later revised and published by Spielberger in 1980, the TAI is a brief one page instrument that consists of 20 statements on which respondents indicate how often they experience the feeling described in each statement listed. On a 4-point scale: (1) almost never, (2) sometimes, (3) often, (4) almost always; respondents indicate their answers. The TAI provides a measure of total TA (TAI-T) as measures of two TA components worry (W) and emotionality (E). Total scores of the (TAI) range from 20 (low test anxiety) to 80 (high test anxiety). Two major components of test anxiety are: (1) worry - cognitive concerns about consequences of failure and, (2) emotionality - reactions of the autonomic nervous system that are evoked by evaluative stress. The TAI also provides measures on these two TA components – worry (W) and emotionality (E). These subscales are derived from eight statements each out of the total 20 of the entire TAI. The total score of the test anxiety inventory test (TAI-T) is based on all 20 statements. Four statements that are on both subscales contribute to the TAI-T score but are not scored on either the W or E subscales. Percentile ranks are calculated from the raw scores. The relationship between the TAI and its subscales with other anxiety measures (e.g., Sarason’s Test Anxiety Scale (TAS), Liebert & Morris’ Worry and
Emotionality Questionnaire (WEQ), the STAI State and Trait Anxiety scales, and the STAI State Anxiety scale administered under examination stress conditions) all provide evidence of convergent validity. The correlation between the TAI-T score and the TAS was sufficiently high (.82 to .83) to suggest that the two scales measure essentially the same construct.

Given the existing research on social facilitation, test anxiety, and behavioral differences in simulated training environments, the question arises if these findings would apply to the environment of flight simulator training. Understanding the effects of observation due to the presence of another individual on another’s performance in simulation training could allow us to enhance our knowledge of the learning environment in the simulator and close the gap between technology and behavioral research.

**Research Questions**

Since the aim in flight simulation training is to establish familiarity in the cockpit and gain a level of mastery with new tasks and not to evaluate the student pilot’s performance; the question is whether the presence of a flight instructor (audience) in the flight simulator in the initial stages of flight training is recommended for optimal learning and flight training progression. The goal of the present study was to examine the effects of the presence of an instructor (audience) on the task performance and test anxiety levels of an individual during the early stages of flight simulation training. Further, the study was designed to explore the possibility that the presence of an instructor can affect task performance directly, as well as indirectly by influencing (through presence of an audience) test anxiety levels in the test subject. Based on the existing literature the following five research questions arose. Does the mere presence of an observer have an effect on another individual’s task performance? Second, when in presence of an audience or observer does an individual increase the tendency of exhibiting
enhanced performance on simple tasks and inhibit performance on complex tasks? Third, does the mere presence of an audience create an atmosphere of evaluation, giving the individual performing a task a feeling of elevated pressure and test anxiety, leading to subjects high in test anxiety demonstrating poorer performance than subjects low in test anxiety? Fourth, do subjects high in test anxiety demonstrate an increased error rate and require significantly more time to learn a task in the presence of an audience compared to subjects low in test anxiety? Finally, do subjects high in test anxiety perform better on simple tasks when in the presence of an audience than on complex tasks?
Methods

Participants

Thirty undergraduate and graduate students (29 men, 1 woman) enrolled in the private pilot flight training program at Embry-Riddle Aeronautical University (ERAU) participated in the study. Mean age of all participants was 21.77 (SD = 5.34).

All participants were private pilot students in the beginning phase of their flight training and had not yet received their license. Fifty-three percent of all participants reported to have had less than 20 hours of pilot in command time logged in the aircraft and 46% reported more than 20 hours. Reported simulator hours showed 90% of all participants having had less than 20 hours and 10% with having had more than 20 hours logged in a simulator. Thirteen participants held a Class I, nine a Class II and four a Class III Federal Aviation Administration (FAA) Medical Certificate. Four participants indicated not holding a medical certificate.

Experience levels for the maneuvers straight and level flight, heading change and traffic pattern were collected through self-report questions. Forty-three percent of all participants reported to consider their experience level with the straight and level flight maneuver as novice, 40% as advanced and 17% as expert. For the flight maneuver heading change 37% reported their experience level as the one of a novice, 43% as advanced and 10% as expert. Experience levels for the flight maneuver traffic pattern 67% of all participants reported an experience level of novice, 27% advanced and 6% as expert. All demographics were self-reported.

All participants were volunteers. Recruitment was conducted through verbal and electronic announcements, e-mail notifications through the ERAU flight scheduling system, as well as in-class recruitment in the human factors department of the university. While some students received extra credit if recruited through in-class announcements, no monetary
compensation was given. Extra credit assignment was at the discretion of the respective professor of each class.

Materials

Consent form. All participants completed a consent form (see Appendix A) prior to the study. Each participant received a briefing to the general scientific purposes of the study, as well as a brief outline of the tasks to be completed during such.

Test Anxiety Inventory (TAI). Each participant completed the TAI before and after the study (see Appendix B). This ensured capture and measurement of each participant’s test anxiety levels before and after the study. As a self-reporting psychometric scale the TAI was developed to measure differences in test anxiety (TA) as a situation-specific trait. The TAI is a brief one page instrument that consists of 20 statements on which respondents indicate how often they experience the feeling described in each statement listed. On a 4-point scale: (1) almost never, (2) sometimes, (3) often, (4) almost always; respondents indicate their answers. Total scores of the test anxiety inventory (TAI) range from 20 (low test anxiety) to 80 (high test anxiety). Although it consists of only 20 items, the TAI has high internal consistency and is highly correlated to other widely used test anxiety measures (Spielberger, Gonzalez, Taylor, Algaze, & Anton, 1978). Research has demonstrated that validity and reliability of the test is relatively equivalent to such of the Test Anxiety Scale (TAS), as the correlation between the TAI-T score and the TAS was sufficiently high (0.82 to 0.83) to suggest that the two scales measure essentially the same construct’ (Spielberger, 1980).

Two major components of test anxiety are: (1) worry - cognitive concerns about consequences of failure and, (2) emotionality - reactions of the autonomic nervous system that are evoked by evaluative stress. The TAI also provides measures on these two TA components –
worry (W) and emotionality (E). These subscales are derived from eight statements each out of the total 20 of the entire TAI. The total score of the test anxiety inventory test (TAI-T) is based on all 20 statements. Four statements that are on both subscales contribute to the TAI-T score but are not scored on either the W or E subscales. Median splits between high and low test anxiety were calculated from the raw scores for all test participants. Test administration took between 8-10 and minutes.

**Flight task and simulator.** Participants performed one of two task scenarios (Simple or Complex) on the flight simulator (see Appendix C). The flight simulator used during this study was the Microsoft Flight Simulator X software.

*Figure 1.* Screenshot of the Cessna 172 with G1000 cockpit in Flight Simulator X.
Prior to the study the MS Flight Simulator X software was installed on four PC workstations in the ERAU human factors laboratory. Each flight simulator work station was set up with a monitor, flight controls, and rudder pedal equipment (see Figure 2).

Figure 2. Flight Simulator X workstation (including flight controls).

In addition, to capture and record the entire flight path and pilot performance of each participant, Flight Simulator Recorder (FS Recorder) was installed on each workstation. FS Recorder is an add-on module for Microsoft Flight Simulator. It allows recording and playback of flights (see Figure 3).

Figure 3. FS Recorder add-on module within MS Flight Simulator X.
FS Recorder allows saving specific flight data in binary file format (.frc files). With the help of a command line conversion tool (FDR 1.0) all files are then converted from .frc format into .txt files (see Figure 4), allowing further data analysis in other software applications later.

Figure 4. FDR 1.0 command line tool converting .frc files to .txt files.

The FS Recorder software was set up to record data points of flight performance of each participant in a two-second-interval. Recorded data in these .frc files included: altitude (ft), heading, groundspeed and time elapsed (secs). Prior to and immediately after completion of a flight task by each participant the experiment facilitator ensured start and stop of the FS Recorder add-on module.

Design

A 2 (test anxiety) x 2 (audience) x 2 (task complexity) factorial design was employed. All 30 participants were randomly assigned to either condition of the audience variable: absence or presence of audience, as well as either condition of the task complexity variable: simple or complex. Seventeen test participants were assigned to the audience group and thirteen test participants were assigned to the no audience group. Participants only completed one of the two scenarios: simple or complex. Thirteen test participants were assigned to the simple task scenario group and seventeen test participants were assigned to the complex task scenario group. Test
anxiety categorization of ‘high test anxiety’ or ‘low test anxiety’ of each participant was derived from the calculated raw test scores from the TAI each participant completed before and after completion of the assigned flight task.

**Independent variables.** Independent variables were (1) test anxiety, (2) audience presence and (3) task complexity. All independent variables were between-subject variables. For the Test Anxiety (TA) variable the participant pool was categorized into two groups: high and low test anxiety. This categorization was performed through the calculated raw scores of the test anxiety inventory (TAI) each participant completed prior to the performed flight task and immediately after. Total scores of the test anxiety inventory (TAI) range from 20 (low test anxiety) to 80 (high test anxiety). Test Anxiety categorization was executed by performing a median-split in SPSS on the pre-TAI test scores, assigning high and low test anxiety levels. Test participants with TAI scores above the median were labeled as high in test anxiety and test participants with scores below the median were labeled as low in test anxiety.

The independent variable of audience presence had two levels: presence and absence of audience. For the purpose of this study audience was defined as the presence of another individual immediately positioned next to the participant while the test participant is completing the flight task scenario in the flight simulator. The experiment facilitator served as the ‘audience’ in the same named group. The task complexity variable also employed two levels: simple and complex task. Through the restriction of each participant to only one flight task scenario, training effects and/or familiarization with tasks, training environment, or flight simulator were eliminated.

In the simple task scenario, participants completed a heading change while maintaining straight and level flight ±200 feet (60 meters) in altitude, airspeed of ±10 knots, and a heading of
±20°. As the Federal Aviation Administration (FAA) Airplane Flying Handbook (FAA-H-8083-3A) states ‘straight and level flight’ and ‘turns’ are two of the four basic flight maneuvers, the task of completing a heading change while maintaining straight and level flight was therefore categorized as the ‘simple task’ (Federal Aviation Administration, 2004).

According to FAA Practical Test Standards (FAA-S-8081-14A) satisfactory performance to meet the requirements for certification is based on the applicant’s ability to safely perform tasks within the approved standards. Approved standards for turns to headings are maintaining an altitude of ±200 feet (60 meters), airspeed of ±10 knots, and maintaining a standard rate turn, rolling out on the assigned heading ±10° (Federal Aviation Administration, 2002). Therefore, exceeding these tolerances (altitude, heading and groundspeed) was defined as errors.

In the complex task scenario participants completed a rectangular course. In this maneuver, the airplane is flown in a traffic-pattern-like maneuver (rectangular course) while maintaining ±100 feet (30 meters) in altitude and an airspeed of ±10 knots (Federal Aviation Administration, 2004).

![Figure 5. Flight pattern of a rectangular course maneuver.](image)
As Federal Aviation Administration (FAA) Practical Test Standards (PTS) categorizes a ‘rectangular course’ as an ‘advanced flight maneuver’ this task was therefore categorized as the ‘complex task’. FAA Practical Test Standards for this maneuver require the pilot to plan the maneuver so as to enter a left or right pattern at 600 to 1,000 feet (180 to 300 meters) above ground level (AGL), at an appropriate distance from the selected reference area, 45° to the downwind leg maintaining an altitude of ±100 feet (30 meters) and airspeed of ±10 knots (Federal Aviation Administration, 2002). Exceeding these tolerances (altitude, heading and groundspeed) was defined as an error.

**Dependent variables.** The dependent variables were (1) numbers of errors performed by each participant during a flight scenario, (2) time (seconds) taken to complete the flight task scenario and (3) the change in test anxiety levels pre-scenario to post-scenario. Dependent variables 1 and 2 quantified task performance in each scenario completed in the flight simulator.

**Procedure**

Prior to the study all participants were asked to complete the consent form (Appendix A). Participants were given a basic description of the purpose of the study and the tasks they were asked to complete during such. Study and task descriptions were read from a study script in order to ensure all participants were given the same information (see Appendix D). Participant demographics were collected through self-report questions asked by the facilitator after the description of the study in order to determine participant eligibility (see Appendix D). Participants were then randomly assigned to condition ‘audience’ or ‘absent audience’. All participants received a detailed description of the flight simulator and a flight task scenario sheet outlining the task they were asked to complete in the simulator during the study. Tolerances of the task scenarios and error definition were provided on the same sheet. Each participant was
asked to complete the TAI twice: before and after the study. This allowed the recording of
differences in test anxiety levels before and after the study. The resulting scores were used to
analyze the interactions between the different conditions of all the independent variables and
their effects on the dependent measures. All scores of the TAI were tabulated and calculated
based on the TAI manual for each participant. Participants were asked to complete one of two
task scenarios (simple, complex) during the study. Task performance data (errors = deviations
from tolerances, time = seconds per scenario) for each scenario in the flight simulator were
captured by the Microsoft Flight Simulator X software and extracted for data analysis after the
study.
Results

The present study was intended to examine audience effects on task performance in student pilots during the early stages of flight simulation training. Further, the design explored the possibility of such audience presence affecting test anxiety levels indirectly.

The data from all 30 student pilots was computed and analyzed using Statistical Package for the Social Sciences (SPSS) version 20.0 software. Task complexity levels (simple and complex) assigned to each student pilot and the condition of audience presence were entered into SPSS. Flight performance data, such as time, heading, altitude, and groundspeed was extracted from the FS Recorder software and also entered into SPSS.

The dependent variable of test anxiety level change scores from pre to post-task completion was derived from the raw scores of the before and after TAI questionnaires each test participant completed. To obtain a change value score for this variable, pre-scenario TAI raw scores were subtracted from the post-scenario raw scores.

Due to non-normality of the dependent measures, a log transformation was performed on each of the dependent measures. Data transformation is a viable option for improving normality of a variable in cases of considerable non-normality due to outliers, or in cases where the researcher chooses not to remove such outliers (Osborne, 2002). Calculating the logarithm of each data point helps to normalize more strongly skewed datasets ("log-normal" distribution). ‘Working with the logarithms of data rather than the data themselves may have several advantages. Multiplicative relationships may become additive, skewed distributions may become symmetrical, and curves may become straight lines’, (Bland & Altman, 1996). Osborne (2002) explains, that ‘[…] data transformations are the application of a mathematical modification to the values of a variable. There are a great variety of possible data transformations, from adding
constants to multiplying, squaring, or raising to a power, converting to logarithmic scales, inverting and reflecting, taking the square root of the values, and even applying trigonometric transformations such as sine wave transformations. [...] Many statistical procedures assume that the variables are normally distributed. A significant violation of the assumption of normality can seriously increase the chances of the researcher committing either a Type I (overestimation) or Type II (underestimation) error, depending on the nature of the analysis and the non-normality. [...] Thus, one reason (although not the only reason) that researchers utilize data transformations is to improve the normality of variables’, (Osborne, 2002).

**Exploratory Analysis**

An exploratory data analysis was conducted on the data to investigate how audience presence and task complexity influenced task completion time, total errors made and the TAI change scores. The raw data is listed by dependent measure in tables 1, 2 and 3. Correlating figures can be viewed in figures 6, 7 and 8.
Table 1

*Descriptive Statistics – total errors.*

<table>
<thead>
<tr>
<th>Dependent</th>
<th>AUD</th>
<th>TASK</th>
<th>TAI</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Errors</td>
<td>No Audience</td>
<td>Simple</td>
<td>Low</td>
<td>1.5170</td>
<td>.28323</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>1.6740</td>
<td>.13820</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1.6067</td>
<td>.20815</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.1862</td>
<td>.21987</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.1133</td>
<td>.16068</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.1570</td>
<td>.19204</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Low</td>
<td>1.9631</td>
<td>.40275</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>1.8937</td>
<td>.27277</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1.9304</td>
<td>.33902</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Audience</td>
<td>Simple</td>
<td>Low</td>
<td>1.7775</td>
<td>.63723</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>1.5909</td>
<td>.54702</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1.7153</td>
<td>.55926</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.0508</td>
<td>.28225</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.3566</td>
<td>.29332</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.2692</td>
<td>.30478</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Low</td>
<td>1.8686</td>
<td>.52866</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.1378</td>
<td>.49680</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.0136</td>
<td>.50928</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simple</td>
<td>Low</td>
<td>1.6659</td>
<td>.49916</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>1.6463</td>
<td>.27045</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1.6568</td>
<td>.39390</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.1523</td>
<td>.22324</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.2485</td>
<td>.26293</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.2032</td>
<td>.24252</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Low</td>
<td>1.9253</td>
<td>.44136</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.0076</td>
<td>.39859</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1.9665</td>
<td>.41532</td>
<td>30</td>
</tr>
</tbody>
</table>
Figure 6. Means for the dependent measure of total errors by groups: audience, task complexity and test anxiety levels.
Table 2

*Descriptive Statistics – time.*

<table>
<thead>
<tr>
<th>Dependent</th>
<th>AUD</th>
<th>TASK</th>
<th>TAI</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Time (Log10)</td>
<td>No Audience</td>
<td>Simple</td>
<td>Low</td>
<td>2.3178</td>
<td>.17194</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.0992</td>
<td>.16959</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.1929</td>
<td>.19464</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.4039</td>
<td>.05759</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.4213</td>
<td>.06766</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.4109</td>
<td>.05873</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Low</td>
<td>2.3752</td>
<td>.10638</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.2603</td>
<td>.20960</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.3211</td>
<td>.16845</td>
<td>17</td>
</tr>
<tr>
<td>Aud ience</td>
<td>No Audience</td>
<td>Simple</td>
<td>Low</td>
<td>2.0907</td>
<td>.31266</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.1850</td>
<td>.13535</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.1221</td>
<td>.25434</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.3456</td>
<td>.12836</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.4537</td>
<td>.30238</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.4228</td>
<td>.25785</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Low</td>
<td>2.1757</td>
<td>.28157</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.3769</td>
<td>.28496</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.2840</td>
<td>.29076</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Simple</td>
<td>Low</td>
<td>2.1880</td>
<td>.27106</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.1278</td>
<td>.15127</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.1602</td>
<td>.21736</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.3893</td>
<td>.07383</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.4393</td>
<td>.21846</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.4158</td>
<td>.16404</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Low</td>
<td>2.2954</td>
<td>.21218</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.3147</td>
<td>.24576</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.3050</td>
<td>.22581</td>
<td>30</td>
</tr>
</tbody>
</table>
Figure 7. Means for the dependent measure of time by groups: audience, task complexity and test anxiety levels.
Table 3

Descriptive Statistics – test anxiety change value scores.

<table>
<thead>
<tr>
<th>Dependent</th>
<th>AUD</th>
<th>TASK</th>
<th>TAI</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAI Difference</td>
<td>No Audience</td>
<td>Simple</td>
<td>Low</td>
<td>2.0118</td>
<td>.03586</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>2.0082</td>
<td>.02153</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.0098</td>
<td>.02577</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.0022</td>
<td>.02704</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>1.9874</td>
<td>.04321</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1.9963</td>
<td>.03297</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Low</td>
<td>2.0054</td>
<td>.02832</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High</td>
<td>1.9978</td>
<td>.03350</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>2.0018</td>
<td>.03012</td>
<td>17</td>
</tr>
<tr>
<td>Audience</td>
<td>Simple</td>
<td>Low</td>
<td>1.9953</td>
<td>.02046</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>1.9910</td>
<td>.01881</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1.9939</td>
<td>.01808</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.0000</td>
<td>.00614</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>1.9977</td>
<td>.02562</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1.9983</td>
<td>.02110</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Low</td>
<td>1.9968</td>
<td>.01627</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>1.9958</td>
<td>.02252</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1.9963</td>
<td>.01908</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Simple</td>
<td>Low</td>
<td>2.0024</td>
<td>.02677</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>2.0025</td>
<td>.02068</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>2.0024</td>
<td>.02316</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Complex</td>
<td>Low</td>
<td>2.0016</td>
<td>.02300</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>1.9931</td>
<td>.03252</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1.9971</td>
<td>.02792</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Low</td>
<td>2.0020</td>
<td>.02391</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>1.9969</td>
<td>.02792</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>1.9994</td>
<td>.02567</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>
Figure 8. Means for the dependent measure of test anxiety change value scores by groups: audience, task complexity and test anxiety levels.

MANOVA

In order to assess the potential significance and interactions between the independent variables: (1) test anxiety, (2) audience presence and (3) task complexity; and the dependent measures: (1) numbers of errors, (2) time and (3) test anxiety change score, a multivariate analysis of variance was initially employed. As this study’s experiment design includes more than one dependent measure and various different independent variables a multivariate analysis of variance was applied instead of a simple analysis of variance. A MANOVA takes into account the intercorrelations among all dependent variables and therefore explores how independent variables influence some pattern of response on the dependent variables. Multivariate test results
in this study are reported with Pillai’s trace, as such criterion should be reported when sample sizes decrease, unequal sample sizes appear, or homogeneity of covariance is violated such as in the case of this study where unequal sample sizes exist (Tabachnick & Fidell, 2001). Pillai’s trace seems to be the most robust of all multivariate tests, especially when the assumption of similar variance-covariance matrices might be violated (Johnson & Field, 1993).

A significant three-way interaction between audience presence, task complexity and test anxiety values was discovered, Pillai’s trace = .363, $F$ (3, 20) = 3.807; $p$ = .026, partial eta squared = .363. However, this result did not manifest itself in the univariate tests, as these showed no significance. This lack of significance in the univariate tests most likely resulted from the small sample sizes within the cell distributions. As a MANOVA uses the sum of all dependent measures and creates a linear combination of such and then tests for differences in this new variable using the analysis of variances methods. The independent variables are used to group these cases categorically. MANOVA then tests whether the categorical variable explains a significant amount of variability in the newly created dependent variable. Reasons for using a MANOVA are that it generally is considered to be more robust against Type I and II errors. Additionally, through this combination of all sums of differences from the individual dependent measures it allows for more power in the analyses. However, a MANOVA can also be highly sensitive to a false positive in significance. While a multivariate significance through this heightened power and the combination of sums can be achieved it is possible that these same results will not be replicated once the MANOVA is pulled apart and analyzed in individual univariate tests where each dependent variable is assessed singularly. Therefore, a multivariate significance can be achieved, as was in this study, but once the univariate analysis is employed it can fall apart through limitations, such as unequal cell distribution and small sample sizes, as
was the case in this study. Further analysis investigating the cause of the interactions was not considered.

The MANOVA revealed a significant multivariate main effect for task complexity, Pillai’s trace = .446, $F(3, 20) = 5.373; p = .007$, partial eta squared = .446. The univariate tests of between-subjects effects demonstrated that this significance resulted from the main effect of task complexity and total errors, $F(1, 22) = 16.71; p = .000$, as well as task complexity and time, $F(1, 22) = 8.97; p = .007$. As the task became more complex test participants demonstrated an increase in errors and time. Therefore, it can be concluded that task complexity has a direct positive relationship with errors and time, as test participants required more time to complete the complex task scenarios as well as performed more mistakes during such.

Because of the potential differences in responses to the manipulation as a function of experience, a MANCOVA, where experience was covaried was performed using the same IVs and DVs.

**MANCOVA**

As the previous literature review has shown audience effects and test anxiety levels to be pre-dominantly affected by the experience level an individual has with the task to be performed, a follow-up multivariate analysis of covariance was performed in relation to flight experience. This MANCOVA was conducted to investigate the effects of flight hours logged in an aircraft on the MANOVA. The MANCOVA revealed that there was a significant three-way interaction between audience, task complexity and test anxiety level, Pillai’s trace = .346, $F(3, 19) = 3.345; p = .041$, partial eta squared = .346, however, the univariate tests were not significant.

In addition to the significant three-way interaction between audience, task complexity and test anxiety, a multivariate main effect was also found for task complexity. Significance for
task complexity was revealed, Pillai’s trace = .380, $F (3, 19) = 3.885; p = .025$, partial eta squared = .380. No significance was found for the covariate of flight hours. The univariate tests indicated that the significance for task complexity resulted from the main effect of task complexity and total errors, $F (1, 21) = 12.273; p = .002$; and task complexity and time, $F (1, 21) = 6.267; p = .021$. These results replicate the results of the previous employed MANOVA, as in test participants demonstrated higher error rates and required more time as task complexity increased.

Although it was believed that flight hours would co-vary between the groups, results of the preceding MANCOVA revealed they did not. Therefore, only the results of the MANOVA will be discussed.
Discussion

While research in the areas of social facilitation has demonstrated that the mere presence of an observer has an effect on another individual’s task performance (Triplett, 1898), the current study could only partially support these findings. In the present study, test participants did demonstrate increased means in the audience groups compared to the no audience groups in all three dependent measures. While means for the no audience group showed lower values in total errors, time and test anxiety change value score compared to the audience group; these results were only partially supported in the MANOVA. Although a significant interaction between audience, task complexity and test anxiety was detected in the MANOVA, potentially supporting such research findings as the ones from Ganzer (1968), Underwood (1976) and Triplett (1898); the univariate tests were not able to detect the interaction. Therefore, while means indicated that the presence of an audience lead to increased error rates, time and test anxiety change scores, these differences in means did not achieve statistical significance.

Additionally, as the previous literature indicated, observers and audiences often can create an atmosphere of evaluation, giving the individual performing a task a feeling of elevated pressure and test anxiety. As this can potentially lead to subjects high in test anxiety (TA) exhibiting inferior performance than test subjects low in test anxiety, the present study also investigated the influences of TA on task performance. While these findings were also reflected in the means of the present study, statistical significance could not be achieved in the multivariate or univariate analysis of variance. Test participants in the low TA group showed lower means for total errors than the high TA group. This rising behavior in means between the low and high TA groups was also recorded in the time and TAI change score values, as well as change value scores. As Martens (1969) discovered, differences in error rates and learning times
were demonstrated between subjects of high and low anxiety; however, these results could not be replicated in the present study.

As stated previously by Zajonc (1965) an audience may well raise the arousal level of an individual performer. When this happens, the degree of proficiency in the skill is important. If the performer is competent and has full mastery of the activity, the dominant responses will be enhanced and a subsequent improvement in performance should result. Conversely, if the skill is at a low level then the emission of dominant responses in this instance may well result in impairment of skills rather than improvement (Underwood, 1976). This would implicate that novice student pilot would demonstrate higher error rates in states of elevated arousal than the more experienced student pilot who’s had more flight training hours. Results in this study however, did not indicate a correlation between performance and flight training hours.

Social facilitation can be clearly defined as the tendency of an individual exhibiting enhanced performance on simple tasks and inhibiting performance on complex tasks in the presence of observers (Constantinou, Ashendorf, & McCaffrey, 2002). The effects of test anxiety on task performance present very similar patterns as a high anxious person will display superior performance when presented with a simple task in a learning environment, opposed to a complex task. This would mean that with an increase in task complexity in a learning situation a person high in TA would demonstrate a decrease in performance, compared to a person low in TA, who would demonstrate an increase in performance.

While no statistically significant results could be achieved to support these findings in relation to audiences and test anxiety, significant support was found for the effects of task complexity of the performing individual. Means for the groups’ simple and complex task also showed a trending increase. In the simple task group the means for total errors, time and TA
change value score were lower than in the complex task group. The trends of means demonstrate that increased errors, time, and change value scores were recorded for the audience, complex task and high TA groups. The multivariate analyses conducted in the present study supported these findings and revealed that as the task became more complex test participants demonstrated an increase in errors and time. Therefore, it can be concluded that task complexity has a direct positive relationship with errors and time, causing test participants to perform more mistakes and require more time the more complex the task becomes.

While the present findings could not support the hypothesized assumptions with statistical significance on all levels, trends of the means recorded allow support in further investigating the aspects of social facilitation and test anxiety in flight simulation training in future research.

**Limitations**

While the present study utilized a modest and straightforward experimental design, some of the measurements and manipulations of variables left the experiment vulnerable for several limitations. Although most limitations stemmed from measurements and manipulations, the execution of the study as well as resources also played a major role in it. While the study employed a 2 x 2 x 2 factorial design only 30 test participants were run. This small sample size likely restricted power, thus limiting the findings.

Another potential limitation was the physical fidelity of the flight simulator. A higher fidelity simulator could have enhanced the experience of the test participant and allowed for stronger motivation. Authority was another aspect and limitation of the study design. While the facilitator served as the audience, employing actual flight instructors in uniform could have enhanced the authority of the facilitator. Higher authority could have enhanced the perception of
an evaluative situation. Finally, there was a time lag between running the first test participants and the last. This time lag could have allowed test participants to spread the word about the study, tools and tasks. This limitation could have caused different perception levels of the study in test participants and lessened the effects of manipulations.

**Future Research**

As the trends of means and partial results from the multivariate analysis in this study indicate that the presumed direction of assumptions was positive, further research into the area of social facilitation and test anxiety effecting flight simulation training and learning outcomes could show of benefit. As technology rapidly evolves investigating the optimal flight simulator environment (with respect to instructor presence or absence) would not only assist in developing better training programs, it could also potentially benefit the students and organizations in better skill acquisition and possibly condensed programs periods. Additional potential benefits could be derived from applying this study to multi-trainee teams to investigate co-action effects, the other major component of social facilitation. Finally, future research looking at gender differences with different forms of authority in the audience role could also serve of potential benefit, as these results could possibly allow for more optimal and efficient instructor and student pilot pairing and enhance learning outcomes.
Conclusion

Understanding the effects of observation due to the presence of others on a persons performance, as well as test anxiety in simulation training could allow us to enhance our knowledge of the learning environment in the simulator and close the gap between technology and behavioral research. The goal of the present study was to examine the effects of the presence of an instructor (audience) on task performance and test anxiety levels of an individual during the early stages of flight simulation training. While trends indicated that a stronger performance was achieved by test participants in the no audiences, low test anxiety and simple tasks groups, only partial significant support could be achieved. Therefore, it can be concluded that task complexity has a direct positive relationship with errors and time, causing test participants to commit more mistakes and require more time as complexity increases.
References


a practical driver's license test with and without the presence of another testee. *Accident Analysis and Prevention*, 39, 1296-1301.


Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.

Appendix A

Informed Participant Consent Form
Social Facilitation and Test Anxiety in Flight Simulation Training

Albert Boquet, Ph.D.
Heidi M. Mehrzad, B.S.
Embry Riddle Aeronautical University
Human Factors Research Laboratory
ERAU, Daytona Beach, FL 32114-3977

The purpose of this experiment is to examine the effects of social facilitation and test attitude on task performance in flight simulation training environments. The experiment consists of one flight task scenario to be completed using Microsoft Flight Simulator X. You will be asked to complete a Test Attitude Inventory (TAI) questionnaire before and after the study to capture your test anxiety levels. For the experiment, you will be assigned a number so that your responses and actions remain anonymous and your name will not be associated with your data nor published. Based on past research with computer-based simulations, a small percentage of participants may experience eye strain, headaches, or slight nausea because of the simulation. If these symptoms develop during the study, you will be asked to stop the task given and take a break until these symptoms subside. If the symptoms continue you will be asked to terminate your participation. You may terminate participation or withdraw from the experiment at any time.

Thank you for your participation. If you have any questions, please feel free to contact Albert Boquet at Albert.Boquet@erau.edu or Heidi Mehrzad at mmanijeh@hotmail.com.

Statement of Consent

I acknowledge that my participation in this experiment is entirely voluntary and that I am free to withdraw at any time. I have been informed as to the general scientific purposes of the study.

Participant’s name (please print):

Signature of Participant: ______________________________ Date: __________

Experimenter: ______________________________ Date: __________

57
Appendix B

Test Inventory (TAI) & Score Key
Test Attitude Inventory

Please provide the following information:

Name________________________ Date________________

Gender (please circle): Male  Female

Directions

A number of statements which people have used to describe themselves are given on the following page. Read each statement and then circle the appropriate number to the right of the statement to indicate how you generally feel:

1 = Almost Never, 2 = Sometimes, 3 = Often, 4 = Almost Always.

There are no wrong or right answers. Do not spend too much time on one statement but give the answer which seems to describe how you generally feel.

Please answer every statement.

Please turn the page for the statements.

Do not write below this line.

Score: T_________ W_________ E___________

© 1980 Charles D. Sleisenger. All Rights Reserved.
Published by Mind Garden, Inc. www.mindgarden.com
Test Attitude Inventory

1. I feel confident and relaxed while taking tests ........................................... 1 2 3 4
2. While taking examinations I have an uneasy, upset feeling ......................... 1 2 3 4
3. Thinking about my grade in a course interferes with my work on tests ... 1 2 3 4
4. I freeze up on important exams .................................................................... 1 2 3 4
5. During exams I find myself thinking about whether I'll ever get through school ................................................................. 1 2 3 4
6. The harder I work at taking a test, the more confused I get ......................... 1 2 3 4
7. Thoughts of doing poorly interfere with my concentration on tests .... 1 2 3 4
8. I feel very jittery when taking an important test ........................................... 1 2 3 4
9. Even when I'm well prepared for a test, I feel very nervous about it .... 1 2 3 4
10. I start feeling very uneasy just before getting a test paper back .............. 1 2 3 4
11. During tests I feel very tense ..................................................................... 1 2 3 4
12. I wish examinations did not bother me so much ........................................ 1 2 3 4
13. During important tests I am so tense that my stomach gets upset ....... 1 2 3 4
14. I seem to defeat myself while working on important tests ..................... 1 2 3 4
15. I feel very panicky when I take an important test ..................................... 1 2 3 4
16. I worry a great deal before taking an important examination ................ 1 2 3 4
17. During tests I find myself thinking about the consequences of failing ... 1 2 3 4
18. I feel my heart beating very fast during important tests ......................... 1 2 3 4
19. After an exam is over I try to stop worrying about it, but I can't ............ 1 2 3 4
20. During examinations I get so nervous that I forget facts I really know .... 1 2 3 4

© 1980 Charles D. Spielberger. All Rights Reserved.
Published by Mind Garden, Inc., www.mindgarden.com
TAI Emotionality Subscale (E)
Items # 2, 8, 9, 10, 11, 15, 16, 18. Minimum E Score: 8, Maximum: 32.
Add the circled values (1, 2, 3, or 4) marked for items # 2, 8, 9, 10, 11, 15, 16, and 18. Enter the sum on the appropriate line on the answer sheet. (If 1 or more items are omitted, see the scoring instructions in the Manual.)

TAI Worry Subscale (W)
Items # 3, 4, 5, 6, 7, 14, 17, 20 Minimum W Score: 8, Maximum: 32.
Add the circled values (1, 2, 3, and 4) for items # 3, 4, 5, 6, 7, 14, 17, and 20. Enter the sum on the appropriate line on the answer sheet.

TAI Total Score (T)
Items # 1 (values of item 1 are reversed), 12, 13, and 19 Minimum T Score: 20; Maximum: 80
To obtain the TAI Total Score, add the values (1, 2, 3, and 4) marked for item 1 (reversed values), 12, 13, and 19 and add the sum to the scores obtained for W and E. The grand total is the TAI Total Score and should be entered on the appropriate line of the answer sheet. NOTE THAT THE VALUES OF RESPONSES TO ITEM 1 ARE REVERSED: i.e., “almost never” is 4 instead of 1, “sometimes” is 3 instead of 2, “often” is 2 instead of 3, and “always” is 1 instead of 4.
If you have not computed W and E scores, the total score may be obtained by adding the circled values of responses to all 20 items; be sure to reverse the face values of responses to item 1.
Appendix C

Flight Task Scenarios ‘Simple’ & ‘Complex’
FLIGHT TASK SCENARIO ‘S’

TASK: HEADING CHANGE DURING STRAIGHT AND LEVEL FLIGHT

1. Fly aircraft straight and level at heading 290° at an altitude of 1,200 feet.
2. When ready, change flight course to heading 240° while maintaining original altitude of 1,200 feet.
3. Maintain a standard rate turn, rolling out on the assigned heading ±10°.
4. Maintain altitude ±200 feet (60 meters), airspeed ±10 knots, and heading ± 10°.
FLIGHT TASK SCENARIO ‘C’

TASK: RECTANGULAR COURSE

1. Select a suitable ground reference area (rectangular course).
2. Plan the maneuver so as to enter a left or right pattern at 600 to 1,000 feet (180 to 300 meters) AGL, at an appropriate distance from the selected reference area, 45° to the downwind leg.
3. Apply adequate wind-drift correction during straight-and-turning flight to maintain a constant ground track around the rectangular reference area.
4. Divide attention between airplane control and the ground track while maintaining coordinated flight.
5. Maintain altitude, ±100 feet (30 meters) and airspeed, ±10 knots.
Appendix D

Study Script & Participant Questionnaire
Introduction

Hi, I am [facilitator’s name] and I will be facilitating the experiment today. To ensure that everyone participating in this study receives the same information, I will be reading from a script (so please bear with me).

The goal of this experiment is determine examine the effects of social facilitation and test attitude on task performance in flight simulation training environments. The experiment will not evaluate you or your abilities, instead just capture your performance in a given simulated environment. All collected data will be anonymous and kept confidential.

The experiment consists of one flight task scenario to be completed using Microsoft Flight Simulator X. You will be asked to complete a Test Attitude Inventory (TAI) questionnaire before and after the study to capture your test anxiety levels.

There are three parts of this study.
1. First you will be asked to complete the Test Attitude Inventory (TAI) questionnaire.
2. Then you are asked to perform a given flight scenario on the flight simulator.
   a. Heading change
   b. Rectangular course
3. Finally, you will be asked to complete the Test Attitude Inventory (TAI) questionnaire again.

Thank you for your participation.

Do you have any questions?

Before we get started, I have a couple of initial questions about you and your flight experience.
Date: ____________________________________________________________

Participant Number: __________________________ Age: ____________________________

Facilitator(s): _________________________________________________________________

1. Gender  
   - [ ] Female  
   - [ ] Male

2. Do you hold an FAA Medical Certificate? If so, which one?  
   - [ ] Class I  
   - [ ] Class II  
   - [ ] Class III  
   - [ ] Do not hold one

3. How many hours of flight instruction have you logged in the aircraft?  
   - [ ] < 10  
   - [ ] 10-15  
   - [ ] 15-20  
   - [ ] > 20

4. How many hours of flight instruction have you logged in a flight simulator?  
   - [ ] < 10  
   - [ ] 10-15  
   - [ ] 15-20  
   - [ ] > 20

5. Are you currently a student pilot enrolled in a private pilot training course on or off campus?  
   - [ ] Yes  
   - [ ] No

6. Please rate your experience level for the maneuver: straight & level flight?  
   - [ ] Beginner  
   - [ ] Advanced  
   - [ ] Expert

7. Please rate your experience level for the maneuver: heading change?  
   - [ ] Beginner  
   - [ ] Advanced  
   - [ ] Expert

8. Please rate your experience level for the maneuver: rectangular course (traffic pattern)?  
   - [ ] Beginner  
   - [ ] Advanced  
   - [ ] Expert
First, let’s review the flight simulator and clarify any questions you may have. For this experiment we will be using the Microsoft Flight Simulator X.

Now I will give you a flight task scenario. The scenario sheet outlines (in a checklist-type-manner) the process and different steps of the maneuver. Just as in an actual flight lesson I ask you to complete the scenario as per checklist, as you would during a lesson. FAA practical test standards will apply and have been listed on the task scenario sheet. During the experiment, if you are unsure of what action you are supposed to perform, please ask and I will clarify.

Now, let’s get started.