AN APPLICATION OF COGNITIVE LOAD THEORY: ASSESSMENT OF STUDENT PILOT PERFORMANCE

by

Brian G. Dillman

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Educational Studies

West Lafayette, Indiana

May 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Aman Yadav, Co-Chair

Department of Educational Psychology and Educational Technology,
Michigan State University

Dr. Ala Samarapungavan, Co-Chair

Department of Educational Psychology and Research Methodology

Dr. Richard Fanjoy

School of Aviation and Transportation Technology

Dr. Richard Olenchak

Department of Educational Psychology and Research Methodology

Approved by:

Dr. Richard Olenchak

Head of the Graduate Program

I would like to dedicate this finished project to my lovely wife, Mary. She has been my partner on many great adventures and she has been a vital aspect of my success throughout this journey. During the years that we have been working on this PhD there have been significant events and many transitions within our family. Her efforts were greatly needed at critical moments along the way and she not only delivered, but shined!! We have discussed many times that she should be allowed to walk across that stage during graduation as much as me. It is certainly true that without her support, encouragement, nudging, and personal dedication to this journey it would not have happened. For that I am forever thankful and words cannot express how blessed I am to consider her my partner in life. We are both extremely grateful that this chapter is coming to a close and I look forward to our next adventure, whatever that may entail.

ACKNOWLEDGMENTS

I would like to acknowledge those that have encouraged and supported me through this journey. I couldn’t have asked for a better Committee and I am thankful for their efforts as a group as well as individually.

Dr. Richard Fanjoy was a dogged presence throughout this journey. He was the one that strongly encouraged me to start the PhD and without his persistence I don’t know that I would have pursued this endeavor. His support along the way was always welcomed and his encouragement toward the end helped to complete this effort. I will be forever grateful to the energy he expended on my behalf. I can only hope to be as great of a mentor and encouragement to others as Dr. Fanjoy was to me.

Dr. Aman Yadav was a tremendous asset and encouragement and I couldn’t have asked for a better Committee Chair while he was at Purdue and his continued support as a Co-Chair once he transitioned to Michigan State University. I was especially thankful for his recognition of my situation as a Faculty member at Purdue myself and his graceful consideration of my progress, albeit as slow as it was. I am also thankful for his considerate and well thought out remarks and recommendations during the editing process and the effort that he dedicated to assist me in achieving this finished project. I thoroughly enjoyed working with Dr. Yadav and am pleased to call him a colleague and friend.

Dr. Ala Samarapungavan was instrumental in keeping me engaged in the educational cycle at Purdue and ensuring that I was able to meet all of the educational outcomes required by the Graduate School. I am especially grateful for her willingness to offer an EDPS 69900 course even when she was on sabbatical and her willingness to accept the Co-Chair position to satisfy the University’s requirements when Dr. Yadav transitioned to MSU. Her dedication to me as a student was truly welcomed and I am thankful for her guidance, correction, and leadership. Her quick responses to my last minute scheduling requests and the multitude of Form 23 signatures were always received with gratitude and thankfulness. I will always be especially grateful for Dr. Samarapungavan’s efforts.

Dr. Richard Olenchak was added to my committee as the final member in my PhD efforts. Although he was the last to join, he was heartily welcomed for this final requirement. He is tremendously busy and I am thankful that he was willing to add an additional workload to his plate in order that I could complete the requirements of my studies. His editorial comments were especially attentive and I believe the end product is stronger because of his efforts on my behalf. I appreciate Dr. Olenchak’s efforts and his flexibility during the final coordination efforts for the completion of the committee requirements.

All of my committee is extremely busy and I am grateful for the time that they have devoted on my behalf. This is an arduous process and I appreciate their willingness to stick with me through the lengthy timeline toward completion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACS	Airmen Certification Standards
CFI	Certified Flight Instructor
CLT	Cognitive Load Theory
FAA	Federal Aviation Administration
NAFI	National Association of Flight Instructors
PTS	Practical Test Standards
SAFE	Society of Aviation Flight Educators
TEL	Technology Enhanced Learning
UAA	University Aviation Administration

ABSTRACT

Author: Dillman, Brian G. PhD
Institution: Purdue University
Degree Received: May 2019
Title: An Application of Cognitive Load Theory: Assessment of Student Pilot Performance
Committee Chair: Aman Yadav and Ala Samarapungavan

Pilot training and certification have largely remained the same since the Practical Test Standards (PTS) were issued more than twenty years ago by the Federal Aviation Administration (FAA). Within the last several years, the general aviation training sector has acquired the capability to collect and analyze digital data from certain training aircraft. With the implementation of digital information analysis, a more accurate picture of the capabilities of student pilots is possible. These advancements could be used by flight instructors in the assessment process of flight students. With the inclusion of digital data from the aircraft, the cognitive load necessary to make an accurate assessment of a student's performance could be affected, ideally in a positive manner. Cognitive load researchers typically focus on three aspects to enhance the likelihood of success in learning or task completion. There are three techniques to reduce cognitive load: (a) reduce extraneous load, (b) manage intrinsic load, and (c) optimize germane load (Young, Cate, O'Sullivan, & Irby, 2016). The current research project focused on the impact to the cognitive load of flight instructors who were presented with digital information retrieved from an airplane during their assessment of a student pilot's aircraft landing competence, endorsement readiness for initial solo, the willingness of the instructor to mentor the student, and how well they liked the student pilot. The study found that a digital condition, when presented alone, created extraneous cognitive load and did not enable flight instructors to accurately rate student landing performance. Additionally, flight instructors were not able to use a combined digital + traditional condition to accurately assess student landing performance. When student performance was on the extreme (i.e. 'poor' and 'good'), flight instructors were better able to determine whether or not a student was ready for a solo endorsement, but instructors did have difficulty distinguishing an 'average' student from a 'good' performing student. Lastly, all of the conditions presented failed to provide the proper visualizations to allow participants to make assessments of their willingness to mentor the

students, and participants indicated that they did not like the students presented with the digital condition. Digital visualizations from aircraft data will require careful development in order to limit the extraneous load and reduce the intrinsic load for student flight assessment, and should be developed in collaboration with flight instructors to provide information to assist the analysis of student flight performance.

CHAPTER 1 INTRODUCTION

1.1 Background

Certified Flight Instructors have a tremendous amount of responsibility when conducting flight training. They are required to manage the requirements of the Federal Aviation Administration (FAA) as well as ensure the safety and operation of an aircraft by individuals who have minimal experience. To compound matters, the cost of flight training can easily exceed $200 per hour for the airplane as well as $50 per hour for the flight instructor. It is expected by students that CFIs would maximize the educational opportunity in order to minimize the amount of flight time that is required to achieve the various instructional milestones. There is an acceptance of the minimum experience that the FAA requires and a reluctance to go beyond that amount. Pilot training and certification have been largely unchanged since the Practical Test Standards (PTS) were issued more than twenty years ago. Within the last four years the FAA has been upgrading the PTS to incorporate explicit aspects of risk management and have renamed the PTS as Airmen Certification Standards (ACS). The ACS is a document that flight instructors can use to determine readiness for pilot certification, but it relies heavily on CFI expertise. Within the last several years, the general aviation training sector has acquired the capability to collect and analyze digital data from certain training aircraft. With the implementation of this analysis there are potential advancements to create a more accurate picture of pilot capabilities. These advancements could be used by flight instructors to maximize the potential benefits of the utilization of digital data in the assessment process, but it is unknown as to whether the instructors will readily embrace this new technology. Additionally, if this information is displayed incorrectly it is possible that the amount of effort, or cognitive load, necessary to make an accurate assessment of a student’s performance will be increased. The current research project focused on the impact to the cognitive load of an instructor when presented with digital information retrieved from an airplane during their assessment of a student’s readiness for initial solo flight in four areas; landing competence, endorsement readiness, mentorability, and likeability.

1.2 Research Questions

1. What is the influence of the use of digital flight data on instructor assessments of student competence for initial solo flight by a flight student?
2. What is the influence of the use of digital flight data on instructor assessments of student readiness for an endorsement for initial solo flight?
3. What is the influence of the use of digital flight data on instructor willingness to mentor students who are preparing for initial solo flight?
4. What is the influence of the use of digital flight data on instructor assessments of student likeability who are preparing for initial solo flight?
5. What is the cognitive load when determining the readiness for solo flight when using varying types of information for analysis?

1.3 Assumptions

1. Participants held an active Flight Instructor Certificate issued by the Federal Aviation Administration
2. Participants did not complete the survey more than once
3. Participants provided responses in an honest manner.
4. Participants did not discuss the survey with other participants

1.4 Limitations

1. For the group that has access to the digital + traditional condition, it is possible that a participant will disregard the digital condition and only use the traditional condition to make their assessments, and vice versa.
2. The data for some assessment groups contained outliers.

1.5 Delimitations

1. Flight instructors with certificates issued by the United States were the only population assessed. Generalizability to instructors certified in other countries is limited.

2. Flight instructors are geographically dispersed and the analysis of participant responses does not account for regional impact across the United States.
3. The assessment will not take into consideration the impact of flight instructor age on the use of digital versus traditional information concerning a flight student.

CHAPTER 2 REVIEW OF LITERATURE

2.1 Cognitive Load Theory

Decades of research has informed the understanding of human working memory, which has been found to have limited capacity. Miller (1956) identified the limited nature of short-term human memory as being limited to 7 ± 2 items. Limited working memory has been a central focus of research within Cognitive Load Theory (Baddeley, 1992; Crapo et al., 2000; Green et al., 2009; Miller & Kintsch, 1994; Van Merrienboer & Sweller, 2005). The prior work on Cognitive Load Theory has commonly classified three types of load that affect working memory, which limit cognitive capacity during learning and problem-solving tasks (Anderson, Potter, Matzen, Shepherd, Preston, & Silva, 2011). These three types of cognitive load are: intrinsic, extraneous, and germane (Carlson et al., 2003; Paas et al., 2003; Sweller & Chandler, 1994; Van Merrienboer & Sweller, 2005; Young et al., 2016).

2.1.1 Extraneous cognitive load

Extraneous cognitive load is comprised of items that are non-essential to the educational goal or task; these items may have been introduced by the instructional technique, but they can still hinder or prevent learning (Young et al., 2016). Extraneous cognitive load could be present due to the nature of the situation, or it could be inadvertently introduced by well-intentioned educators in the development of instructional resources. Researchers have spent much of their time creating educational designs and procedures that reduce extraneous cognitive load compared to conventional efforts (Paas, Renkl, and Sweller, 2003). This enables the individual to focus their limited working memory resources on difficult aspects of a task and, hopefully, to develop patterns to be used in future tasks. All three forms of cognitive load combine together for a total load, so limiting the amount of extraneous cognitive load is especially necessary when intrinsic cognitive load is high (Paas et al., 2003; Van Merrienboer & Sweller, 2005).

2.1.2 Intrinsic cognitive load

Intrinsic cognitive load is comprised of items that are essential for the completion of an educational assignment or task (Young et al., 2016). The intrinsic load will vary for different tasks, but it can be stated that simple tasks will have lower intrinsic load, and difficult tasks will have higher intrinsic load. When the intrinsic cognitive load of educational tasks is too high, new educational methods are needed to manage the cognitive load (Van Merrienboer & Sweller, 2005). The total cognitive load is a balancing act that must be carefully mitigated. In an attempt to develop educational materials to reduce the intrinsic load for a given learner, it is possible to increase the extraneous load, which would continue to hinder or prevent learning.

2.1.3 Germane cognitive load

Germane load is comprised of the effort put forth to recall information already learned and stored in long-term memory (Young et al., 2016). Whereas extraneous cognitive load created barriers to learning, germane cognitive load assists learning (Paas et al., 2003). Germane load focuses on the utilization of schema, which includes general knowledge that embodies numerous items of information stored within long-term memory into a single element (Carlson, Chandler, and Sweller, 2003). These schema will increase in complexity as a person gains experience within a subject and the robustness of their knowledge increases, which can then be accessed and utilized during future tasks or educational experiences. Long-term memory allows individuals to tremendously increase this processing ability by storing large numbers of schemas that include a variety of informational elements into a single element for a dedicated purpose (Paas et al., 2003).

2.2 Research in Cognitive Load Theory

Cognitive load researchers typically focus on three aspects to enhance the likelihood of success in learning or task completion. These techniques are to: (a) reduce extraneous load, (b) manage intrinsic load, and (c) optimize germane load (Young et al., 2016). A research study that particularly highlighted the effect of cognitive load in a task very similar to the assessment of student flight performance is that of Young, et al. (2016); the researchers assessed the techniques utilized when medical personnel hand off patient information from one shift to another within the context of cognitive load. Both medical patient assessments as well as flight student assessments rely on the expertise of the individual making an assessment. The professional’s background knowledge on a particular individual will impact their ability to make accurate future decisions. Flight students occasionally change instructors for a variety of reasons and this change could affect the cognitive load of the new instructor as they transition into a position of responsibility. In the context of medical patient hand-offs, the authors found that the extraneous cognitive load was increased due to the need to access multiple sources of information. Specifically, they identified that when a learner is required to access information in multiple locations or during multiple timeframes, the limited working memory that is available is dedicated to collecting information and combining the sources rather than interpreting and understanding the information (Young, et al., 2016, p. 91). In other words, having to access multiple sources of information at the start of a new shift increased the extraneous cognitive load of the incoming medical personnel; this reduced the ability for the medical personnel to utilize past experiences and knowledge of the tasks to address the demands of the current hand-off task for the successful completion of medical patient hand-off. Consequently, best practices for patient hand-off accuracy tend to reduce interruptions, employ communication standardization, and locate all information sources in a readily accessible location in order to reduce extraneous load (Young et al., 2016). Similarly, locating information in a single physical space directly applies to the task of student assessment by flight instructors as a determinant of the likelihood of future performance. Pilot logbooks do not adequately provide information in a readily accessible format for instructor assessment and force the instructors to rely on individual memory concerning student capability. Young et al. (2016) proposed a method of minimizing extraneous load by incorporating an electronic record of the medical information that would automatically populate all necessary information from the multiple sources into a single electronic portal to be used as a checklist for each step of the hand-off, and would decrease searching for information. Through this visualization, medical professionals (and potentially flight instructors) could more effectively complete the required tasks before them in their daily efforts. Young et al. (2016) also highlighted the issue of experience and the impact of presented material on cognitive load. They identified that current techniques do not consider the level of expertise of those involved in the hand-off or the complexity of the cases into account and that they tend to have a "one-size-fits-all" approach. This is also a dynamic that is important to assess concerning the impact of information presentation for a flight instructor during student assessment and whether or not the visual presentation of information will assist in the assessment process when determining student readiness for initial solo flight. The measurement of flight instructor expertise to be used in the present study's data collection phase was the number of hours of 'Dual Given' that each participant has delivered to flight students. Dual given is recorded in a student pilot and flight instructor's logbooks for every hour of instruction recorded in an airplane and is the primary mechanism by which flight instructor experience is measured.

2.2.1 Visualizations and Analysis

Visualization is relatively new as a field of study, and an exact understanding of what it entails does not exist (Yau, 2013). Visualization is a term that has a wide range of interpretations from the use of a diagram by a student to answer a basic question to the visuals used by a researcher to evaluate complex data. The choice of the best visualization technique for a particular data set is difficult to make (Anderson et al., 2011; Kosara et al., 2003; Van Wijk, 2005; Yau, 2013), and the choice will need to satisfy the specific needs of the individual user. When data is not clearly understood by the observer, the use of multiple views is the one most likely visualization to lead to spontaneous insight (Buja et al., 1991; Green et al., 2009; Kosara et al., 2003; Tufte, 2001, 2006). Van Wijk (2005) and Yau (2013) cautioned users that visualizations should not be used to verify the final truth but rather to inspire new hypotheses to be checked afterwards, which is particularly significant when inspecting complex data (Ahlberg et al., 1992; Keim et al., 2008; Shneiderman, 2014; Van Wijk, 2005). Additionally, the maturity of the research area within which visualizations are desired to be used may factor into the type and use of visualizations.

There are, however, limits to the amount of opportunity for insight that visualizations can provide in the context of data inquiry, but Morse, Lewis, and Olsen (2000) argued that visualizations for supporting information retrieval activities have significant utility. Additionally, Keim, Andrienko, Fekete, Görg, Kohlhammer, and Melançon (2008) made a distinction between visualization (creating a visual representation of information) and analysis (using a visualization to garner information about the data). Thereby, the process of visualization is creating an image from information without any level of interpretation, whereas analysis is the process of utilizing that image to draw a conclusion or to understand more about a particular topic. Keim et al. (2008) further suggested that visual analytics are more than just the creation of a visualization and that visual analytics combine the strength of human and electronic data processing. The user has the ability to both guide the creation of a visualization as well as utilize that visualization in the analysis of the information. There have been many research efforts concerning the solidification of the goals of visualizations and their analysis in education and the workplace (Green et al., 2009; Keim et al., 2008; Kosara, et al., 2003). Green, Ribarsky, and Fisher (2009) notably stated that one goal of visualizations should be to assist in the flow of reasoning. It is this aspect that was evaluated concerning flight instructor assessment of student performance in the present research project in regards to the assessment of student flight performance.

Visualizations that combine diagrams and text can prove to be a valuable tool in the reduction of extraneous load when completing a task (Chandler & Sweller, 1991). Visualizations reduce cognitive load by providing scaffolds and access to perceptual information that are more beneficial than having only access to written or verbal information alone (Larkin & Simon, 1987; Liu et al., 2008). In order for these scaffolds to positively influence participants and not create extraneous cognitive load, they must be intentionally developed in order to support the educational goal. Dastani (2002) determined that this is a vitally important step in data visualization so that there is not an increased cognitive load during the process of assessment as a user is evaluating a visualization created from data. The interpretation of the visualization should be directly linked with the intended structure of the data (Dastani, 2002). The ability to make the connection between the visualization and the structure of the data will maximize the use of germane load and prevent the introduction of extraneous load. Along the path toward this intended objective, there have been many realizations concerning best practices and guidelines for visualization development within the educational context as discussed in the following section.

2.2.2 Visualization in an Educational Context

In educational settings, there have been a variety of efforts to determine the benefit of using digital data as a tool for providing a better-rounded picture of the efforts that are being made toward learning a particular discipline. These efforts have largely focused on data distillation for human judgment in education for two key purposes: the identification of student abilities and the classification of their performance (Baker, 2010). The consumers of these visualizations are the individuals who possess the ability to change or modify the educational efforts made for the students (Dyckhoff et al., 2012; Luan, 2002). Thus, in order for visualizations to be used successfully, it is important for the investigator to be a domain expert who is attuned to the meaning of the data being visualized (Luan, 2002). Within flight training, there is limited opportunity to be attuned to the information that can be displayed from an aircraft, such as side loading during landing and the deviation from the runway centerline on rollout after touchdown. Current flight training programs do not rely on this flight information, especially when looking at longitudinal performance; rather, they rely largely on flight instructor assessment and personal memory.

In other educational contexts there have been efforts of implementing programs that will allow teachers and instructors to have a more well-rounded idea of a student's learning. Dyckhoff, Zielke, Bültmann, Chatti, and Schroeder (2012) utilized a concept called Technology Enhanced Learning (TEL) to combine data from a variety of sources into visualization displays for teachers' assessments of course design. They received positive feedback concerning the measures of student success, such as the activity of students, program and duration of study, adoption rate, and the "Top 10" resources that students were reviewing. A main aspect of Dyckhoff et al.'s (2012) method of information display was that they indicated certain aspects about the learning environment and tried to visualize the usage and properties through the use of indicators tied to a specific question.

In another study, Mazza and Dimitrova (2004) utilized a program called CourseVis to assess social, cognitive, and behavioral aspects through visualization techniques. This was accomplished by plotting data from discussion boards into a three-dimensional scatterplot. Within these interactions, the teachers were able to identify the depth and breadth of the discussions of each individual on a variety of topics. This allowed the teachers to influence the discussions if they were weighted too heavily regarding one particular topic. Ultimately, CourseVis provided the teachers with a better overall ability to assess their student's learning and to determine areas where they could focus on educational improvement.

In contrast to the more traditional methods of visualization illustrated above, an effort to visualize students' understanding of reading assignments through the use of reflective blog posts was conducted by Stover, Yearta, and Harris (2016). While a blog is not a traditional representation of data concerning the aspect of visualization, the teachers were able to use the blogs to gain a detailed understanding of what the students were thinking as they were reading the assigned material. At the very minimum, Stover et al. (2016) discovered that the blog served as a complete record of what the students' were thinking and provided a mechanism by which they could reflect on their work. In the midst of discussing a particular piece of writing, there can be value in the ability to return to previous assessments as an opportunity to reconsider earlier thoughts and ideas. From the teachers' standpoint, the blogs provided an opportunity for unique responses to individual students with further scaffolding opportunities for students who needed a little more assistance. Overall, the researchers found that as a result of the reflection guided by the teachers' ability to understand what students were thinking, the students were more thoughtful when discussing books online. In relation to student pilot education, digital flight data is not reflective in nature, but it does provide an opportunity for comparison and evaluation of the achievement of competencies rather than a few singular instances where a student may have had a good day.

In a similar study, Shelton, Smith, Wiebe, Behrle, Sirkin, and Lester (2016) determined that students' represented their level of understanding of magnetism differently when completing written versus drawn assignments. However, when both the written and drawn artifacts were combined the educators gained a better awareness of the students' understanding of the science topics. Contrary to the previous study, where the blog served as a visual representation of the students' understanding of the reading material, it was the combination of the written text and drawings (as a visualization of material understanding) that allowed the researchers to arrive at a true awareness of the students' level of knowledge concerning the scientific topic. The researchers did acknowledge that they were working with young students, who are generally much better at using illustrations rather than using the written word to articulate their understanding of a given topic.

Each of the educational visualization efforts highlighted above relied on information retrieved from databases and presented in a format that was readily accessible by the teacher. These visualization portals provided a lens by which the teachers could understand the interactions that the students were having with the material in a way that would not be possible otherwise.

2.2.3 Visualizations in Aviation

One resource for the development of data visualizations within flight training is the digital data that can be accessed from the training aircraft. This data is captured from the avionics and flight instrumentation of the aircraft with electronic flight displays. As an example, digital data is captured from the Garmin G1000 avionics platform for sixty-nine parameters at a rate of 1 Hz (i.e., once per second). The data can be accessed from the aircraft in the raw format of a comma separated values (.csv) file for each flight. In order for this digital information to be viewed visually, it would need to be processed and displayed in a visual format. With the access to digital data and proper analysis techniques, there are potential opportunities to create a more accurate picture of the capability and proficiency of pilots. The visualizations created from the raw digital data could be used by flight instructors to maximize the potential benefits of this new educational paradigm, including more efficient training platforms, less acquisition cost for initial certification, and better awareness for both students and instructors concerning strengths and weaknesses during training.

However, the amount of data coming from an aircraft varies among avionics and aircraft platforms. The manufacturer of the aircraft, the manufacturer of the avionics system, the model of the avionics system, and the type of engine all factor into the number and frequency of data points that are collected and available for assessment. In general aviation, the number of data points is relatively small in comparison to the commercial aircraft sector. For example, the Cirrus SR-20 aircraft that Purdue University operates collects data once per second (1 Hz) on sixty-nine parameters. In comparison, commercial aircraft collect at least 88 parameters and at a frequency of at least 8 Hz recorded on primary flight controls as required by the FAA (2008).

2.2.4 Issues with visualizations

A consistent issue with the deployment of digital data in all fields is that there is a need (at least at the outset) for a level of expertise within the discipline in order to utilize, incorporate, refute, or interpret the information that is produced from a digital data system. For instance, Rönkä, Tolvanen, Lehikoinen, von Numers, and Rutkari (2008) analyzed visualizations for bird management and conservation purposes and determined that solid background knowledge of the wildlife analyst concerning the habitat preferences of birds was necessary before the digital data could be used as a sole means of analysis. Even with this background expertise, the authors could not determine a ‘good’ model for a particular species of tufted duck. So, despite the best efforts of the team, a model for one particular species of bird was not able to be determined.

Data mining in education, as compared to other fields, emerged later and was due, in part, to the lack of availability of large datasets in usable formats (Baker, 2014). School records most commonly were in paper formats and the information from online learning systems was not user friendly and readily accessible (Baker, 2014). The author also determined that information that could directly indicate the level of knowledge by a student of a given topic was sometimes not recorded or available. Aviation finds itself in this state at the present, as traditional logbooks that are used to record flight time, routes flown, airports visited, the number of takeoffs and landings, and any additional comments by instructors have just started to transition into the digital arena in the last several years. Many flight operations still conduct everything by paper. Baker (2014) highlighted a challenge of scalability in developing visualizations in the educational environment and research. The author emphasized that a small number in the hundreds or thousands may be able to benefit from educational visualizations, but it may be impractical for large numbers to benefit, especially in longitudinal outcomes and assessments. The author explicates that this lack of information makes it difficult to pay attention to validation methods that inform developers about whether a model is representative for all learners. This issue can be addressed by the expert validation that was discussed earlier, but the risk of inter-observer errors comes into play until the amount of information available for analysis has reached a threshold that can be vetted with confidence. Once a significant amount of students' performance data has been uploaded into a digital database, it is reasonable to increase confidence that the resulting visualizations will be applicable for more than just the students from whom the visualizations were first developed (Baker, 2014).

2.2.5 Aviation privacy issues with visualizations

Bienkowski, Feng and Means (2012) also highlighted that, "education institutions must consider privacy, policy, and legal issues when collecting, storing, analyzing, and disclosing personally identifiable information from students' education records to third parties for data mining and analytics." (p. 42) When it comes to data from daily operational platforms, the FAA has taken a hands-off approach concerning operational digital flight data, but there has yet to be a formal decision concerning access, use, and control. This is in contrast to the clear-cut policies concerning the control of digital flight data when an aircraft is involved in an accident or an incident, contained in both NTSB 830 (2017) and NTSB 831 (2017).

To this end, there are some within the aviation industry who see the use of digital flight data much in the same way that the development of educational data was seen by Blumenstyk

(2016), who stated that colleges need to find ways to raise the money for the research and development of learning software so that for-profit companies did not end up owning the classroom delivery system of the future. Blumenstyk also highlighted an issue with digital data analysis being limited in the determination of how or why a student answered a problem incorrectly. Largely, the researcher indicated that the systems have difficulty determining if a student does not understand a given educational concept or simply made a mistake in the underlying process for determining an answer.

Moreover, an issue that Galyardt and Goldin (2015) brought up concerning the visualization of digital data for educational assessments is that that a handful of the most recent observations are a better summary of the learner's mastery of a skill than the student's entire history of practice. This is especially true of flight education from the beginning to the acquisition of a new flight capability. It cannot be expected that a private pilot applicant will be able to successfully land an airplane during the first flight lesson. As the pilot approaches their practical exam date, the assurance of a successful landing is more likely. The question, therefore, is at which point should the student be able to successfully demonstrate a given skill or aspect of knowledge. It becomes imperative that a digital data system have safeguards in place to ensure that the incorporation of older data not occur if it might make an accurate assessment less likely.

2.2.6 How to know it is working

With all the information available concerning best practices in visualizations as well as the impact of structure on cognitive load, it is important to determine whether or not the information is assisting or hindering the flight instructor. One mechanism for the evaluation of visualizations is through subject matter expert evaluations and a feedback loop process on the visualization itself as well as the assessment of student performance (Anderson et al., 2011). This is "low-hanging fruit" and can have value in the process of assessing visualizations and their potential impact on cognitive load. However, Anderson et al. (2011) note that if, "working memory is a central construct of the cognitive process, then the burden placed on working memory and cognitive load can be used as a means to measure the efficacy of visualizations." (p. 2) This assumes that the cognitive load of a task is known prior to the implementation of visualizations. One mechanism for measuring cognitive load is the Paas Scale (Sweller & Paas, 2017), which consists of a single item that asks participants to indicate how much effort they devoted to the task at hand, with responses commonly made on a nine-point scale that rates effort from extremely low to extremely high. The use of this scale and its approach for measuring cognitive load has an interesting dynamic for use in measuring the cognitive load needed when a flight instructor is using a digital portal for the assessment of student performance. This scale was used to provide insight as to the cognitive load impact of the information presented to flight instructors during the assessment of student performance at various performance levels.

2.3 Pilot Training

Pilot training and certification have remained largely the same since the Practical Test Standards (PTS) were issued more than twenty years ago by the Federal Aviation Administration (FAA). The PTS and, now, the Airmen Certification Standards (ACS) are used by Certified Flight Instructors (CFIs) as a guide for teaching pilots as well as by Designated Pilot Examiners (DPEs) to assess the abilities of pilots. This guidance offered by the PTS and ACS is very subjective in nature and is influenced by the experiences and knowledge of the individuals involved in the pilot training and certification process. The journey of pilot certification is one that generally takes a full calendar year to accomplish. It begins with a CFI providing a student pilot with flight training in an aircraft and ground training in a classroom with several hurdles to overcome. An initial solo flight is the first hurdle, and it is completed after the student pilot has met all of the requirements set forth in Part 61 of Title 14 of the Code of Federal Regulations. It is the CFI’s responsibility to ensure that the student pilot has met all of these requirements prior to the initial solo flight. The next hurdle to overcome is a cross-country flight. The student’s CFI completes at least three hours of flight time with their student to airports that are at least 50 nautical miles away from their home airport. Once the student’s CFI has ensured that their student has met all of the solo cross-country requirements in Part 61 of Title 14 of the Code of Federal Regulations, they then release the student to conduct five hours of solo cross-country flight. The last hurdle prior to Private Pilot Certification is to prepare for the practical exam with a DPE. The practical exam is comprised of a period of one-on-one oral questioning, typically in an office setting, followed by a flight in the aircraft in which the student has been practicing all of the requirements for certification leading up to this point. After successful demonstration of their knowledge during the oral/ground portion and the successful demonstration of aircraft control during the flight portion, the student will be awarded a private pilot certificate. During this entire process, the CFI has a tremendous amount of responsibility, authority, and liability; furthermore, the decisions they need to make concerning the safety of their student pilot can seem quite daunting. These decisions and an awareness of the responsibility that is placed on a CFI become easier with experience and time. For the purposes of this study, the CFI's assessment of a student pilot's readiness for initial solo flight and the cognitive load required to complete this assessment will be the main focus of the research.

2.3.1 Solo Flight Training Process

Preparing a student pilot for the initial solo flight requires the development and assessment of a complex set of knowledge, skills, and abilities. During the training process, the FAA requires that a student pilot demonstrate both knowledge (i.e., oral knowledge) and safe aircraft handling. Specifically, the process includes a demonstration of satisfactory aeronautical knowledge on applicable sections of flight regulations, airspace rules, and procedures for the airport where the solo flight will be performed as well as the flight characteristics and operational limitations for the make and model of the aircraft to be flown (FAA, 2017). The FAA also requires that the student pilot has received and logged flight training for the maneuvers and procedures appropriate to the make and model of the aircraft and demonstrated satisfactory proficiency and safety, as judged by an authorized instructor, during the maneuvers and procedures required by Part 61 of Title 14 of the Code of Federal Regulations in the make and model of the aircraft to be flown (FAA, 2017). Prior work has found that flight instructors could be influenced by their own background and experiences when evaluating whether a student can independently execute flight maneuvers prior to solo flight and that they place varying importance on different areas when making the final determination if a student is ready for initial solo flight (Thomas & Richards, 2015). Thomas and Richards (2015) identified that instructors largely focus on two aspects when determining readiness for initial solo flight—safety and student competence. However, these two aspects can be interpreted and assessed very differently by flight instructors.

Thomas and Richards (2015) note several key themes for the determination of safety of flight operations. One of which was situational awareness and the ability to adapt when the conditions vary from what the students have previously experienced (2015). One instructor expressed this as, “The biggest thing I think is when the student can identify when things don’t look right” (p. 118). However, other instructors viewed their own ability to identify situational awareness as whether the student themselves felt ready for solo flight or if they needed a few more practice laps around the traffic pattern. Another theme for the level of safety required dealt with the weather conditions present for the initial solo flight (Thomas & Richards, 2015). One instructor expressed that the weather conditions needed to be perfect with light wind conditions, but another instructor expressed that perfect weather conditions were not necessary as long as the student was able to handle the conditions on the given day that they intended to conduct their solo flight.

Furthermore, an additional key theme for the determination of student competence was the consistency of performance, but, like the determination of safety, instructors varied in the metrics that they utilized for determining whether a student had demonstrated consistent performance (Thomas & Richards, 2015). This difference was most apparent in comparing quotes from two instructors; while one instructor simply stated, “The process I use is I look for consistency. They do a minimum of three circuits, and I want to see the numbers the same on every circuit”, (p. 118) another instructor stated,

> Some instructors will say that if they get into the plane and the student does three perfect circuits in a row, they will send them solo. I am a little more conservative. I want to see more consistency than three. If I am assessing them, I should be able to sit there and say and do nothing. If I have to give them any prompting, they are not ready. (p. 118)

In comparing these two statements, it can be seen that some instructors have a metric to evaluate students that has worked for them in the past, one that they trust informs the outcome of whether an initial solo flight by a student will result in a successful flight outcome.

The discussion above shows that the process for determining readiness of a flight student for initial solo flight varies amongst flight instructors. The amount of information to process and the substantial risk involved with approving a student for initial solo flight can result in hesitancy in endorsing someone for the flight. The amount of cognitive load necessary for an assessment could influence a flight instructor’s capability to accurately gauge a students’ ability to fly a plane given the amount of information required to assess it.

2.4 Purpose of the Study

Flight instruction and aircraft utilization are expensive, and there is a substantial potential for reducing the amount of expenses that are utilized in the unnecessary duplication of maneuvers and training efforts. Without access to digital measures of flight proficiency, instructors are often left to deciphering poorly written comments in logbook records that often leave little guidance concerning the strengths and weaknesses of flight students. The resulting outcome is that flight instructors often duplicate previous lessons in order to ensure that they have comprehensively evaluated the performance of a student prior to allowing them to progress to the next step in their training sequence. Likewise, instructors are required to make high-stakes assessments of a pilot's proficiency and competency. Thus, the incorporation of digital data has the potential to reduce unnecessary expenditures and to allow instructors to make better informed decisions concerning pilot competency and readiness for certification. The goal of this study is to examine how aviation data and visualizations influence instructor evaluations of student readiness for solo flight and the amount of cognitive load that is necessary when conducting these assessments.

CHAPTER 3 METHODOLOGY

This chapter provides the methodology for this dissertation research project. A randomized Qualtrics Survey in a 3x3 design was utilized. The goal of this research was to determine the cognitive load of an instructor pilot during the assessment of a student pilot's readiness for initial solo flight while having access to a variety of information sources.

3.1 Research Design

The research design for this project was an experimental quantitative case study. The dynamics of the circumstances that surround the environment of flight instruction are such that it was necessary to create a manufactured flight student that could be evaluated by flight instructors. The design was experimental in such that the "portfolio" of both the traditional condition and digital condition was fictional but representative of a flight students training to be a private pilot. The design was quantitative in that the measurements of assessment were collected through a Qualtrics Survey seven-point Likert-type questions. Finally, the design was a case study in that the individuals providing the assessments of confidence evaluated a hypothetical individual flight student and assessed the performance and readiness of the individual for initial solo flight.

3.2 Participants

Pilots with current Certified Flight Instructor (CFI) certificates were recruited through the University Aviation Association, National Association of Flight Instructors (NAFI), and the Society of Aviation and Flight Educators (SAFE). According to their website, UAA is the voice of collegiate aviation to their members, the industry, government and the general public, and plays a pivotal role in the advancement of degree-granting aviation programs that represent all segments of aviation. NAFI is dedicated exclusively to "raising and maintaining the professional standing of the flight instructor in the aviation community." SAFE is a member-oriented organization of aviation educators fostering professionalism and excellence in aviation through continuing education, professional standards, and accreditation. As of November 3, 2017, the Federal Aviation Administration's Airmen Certification Branch indicates that there are 80,688

CFIs with valid medicals in the United States. Based upon the total CFI population with a significance level (α=0.05) and a confidence interval of 5% the calculated total sample size is 171 participants for an ANOVA analysis.

3.3 Procedures

Prospective participants received an email containing a link to a Qualtrics survey for the study. Upon clicking the link, the participants were randomly selected into one of the groups identified in Table 1.

Table 1 – Participant Breakdown

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="3">Student Performance</td><td colspan="2" rowspan="2">Totals</td></tr>
<tr><td>‘Poor’</td><td>‘Average’</td><td>‘Good’</td></tr>
<tr><td rowspan="3">Performance Data</td><td>Traditional</td><td>19</td><td>19</td><td>19</td><td>57</td><td rowspan="3">171</td></tr>
<tr><td>Traditional + Digital</td><td>19</td><td>19</td><td>19</td><td>57</td></tr>
<tr><td>Digital</td><td>19</td><td>19</td><td>19</td><td>57</td></tr>
<tr><td colspan="2" rowspan="2">Totals</td><td>57</td><td>57</td><td>57</td><td colspan="2" rowspan="2"></td></tr>
<tr><td colspan="3">171</td></tr>
</table>

After agreeing to participate in the study the participants were presented with a series of six questions to assess their demographics. Demographic information collected was the length of time (in years) that participant has been a flight instructor, how many total flight hours of experience the participant has accumulated, the total number of instructional flight hours the participant has accumulated, and how many students the participant has endorsed for initial solo.

As Young, et al. (2016) highlighted in assessing the impact of visualization tools in the handoff of a medical patient amongst health professionals, experience can play a role on the impact of such tools. The determination of flight instruction experience of each survey participant allowed the analysis of whether or not this aspect impacts the utilization of flight performance visualizations.

3.4 Student Profiles for Evaluation

There were three different levels of performance that were compiled for the purposes of the developed student profiles. Both traditional logbook entries (traditional condition) were developed as well as digital information about landing performance (digital condition). 'Good', 'average', and 'poor' student performance was determined based upon the ability of the student to keep the aircraft close enough to the centerline of the landing runway so that the main wheels remained on either side of the approach path. The measures of 'good', 'average' and 'poor' was based upon a typical progression sequence of a flight student toward learning the skill of landing an airplane. An 'average' student will historically spend seven to ten flight hours in the traffic pattern of an airport practicing landings with a flight instructor and will demonstrate steady progress toward the goal of an initial solo flight. The hypothetical 'average' student for the current study replicates that traditional footprint and shows steady improvement after 8.3 hours of flight time and by all indications will be ready for an initial solo flight after one or two more flights with an instructor. The hypothetical 'good' student for the current study shows solid aircraft control during landings after 8.3 hours of flight time and by all indications is ready for an initial solo flight. The hypothetical 'poor' student for the current study shows inconsistent aircraft control during landings and a lack of progress after 8.3 hours of flight time. The 'poor' student will require significantly more flight time in the airplane in order to be ready for an initial solo flight.

Traditional Condition: The traditional condition is intended to replicate a pilot logbook which is utilized by certified flight instructors and flight students on a regular basis. It is a booklet that captures information about the flight as required by the FAA concerning flight time logged for training or recent experience. The first three entries in the traditional condition are the same for all hypothetical student performance levels. The last three entries are unique to the 'good', 'average', and 'poor' conditions. This condition will include one of the representations of student performance from Tables 2, 3, or 4 only.

Digital Condition: The digital condition is a representation of the aircraft landing performance which specifically focuses on the amount of deviation from the centerline of the runway for each landing. Data was from actual flights recovered from aircraft flown within Purdue University's Professional Flight Program. This condition will include one of the representations of student performance from Figures 1&4, 2&5, or 3&6 only.

Traditional + Digital Condition: The traditional + digital condition will include both of the representations of student performance from the traditional and digital conditions. The potential combinations are: Table 2 with Figures 1&4, Table 3 with Figures 2&5, or Table 4 with Figures 3&6.

3.4.1 Survey Participant Introduction to Scenario

The following script was used in the Qualtrics survey as an introduction to the student that participants evaluated. The scenario that the script presents is one that, while not a weekly occurrence, does happen with regularity.

Jacob Smith has been working on his first solo. Jacob started his flight training at the beginning of the summer and has been flying 2 or 3 times per week for the last several weeks. His instructor, Steve, acquired the hours necessary to get a job with the Regional Airlines and left for his starting class date at the beginning of last week. The Chief Flight Instructor at Jacob's school contacted you and asked that you continue working with him toward his goal of becoming a Private Pilot. The Chief Pilot did mention that Jacob has been paying for his flight lessons out of his savings that he had acquired working a part time job. The Chief CFI asked that you not duplicate any flights that weren't necessary for fear that Jacob would run out of money prior to him completing his flight training. Look over the records that Jacob's previous instructor left and answer the questions concerning the student's readiness for initial solo.

3.4.2 Traditional Conditions

Table 2 – 'Poor' Student Performance

Flight	Dual Given	Number of Landings	Comments
Landing Lesson 1	1.3	8	Good attempt at first landings. Keep aileron correction in to the wind and increase deflection as you get slower as required especially upon touchdown.
Landing Lesson 2	1.5	8	Landing sequence is good, but still need work on airspeed control on final approach and flaring. Generally flaring too late and flat.
Landing Lesson 3	1.4	6	Improved but still landing firm occasionally. As aircraft slows over the runway, need to add increasing back pressure to prevent it from sinking into the runway with the nose dropping.
Landing Lesson 4	1.4	7	Flare was high. Make sure your eyes are at the end of the runway to see if flattening out. Runway width may cause visual illusion.
Landing Lesson 5	1.6	9	Flare a little aggressive, landing a little flat/sideways, remember xwind correction
Landing Lesson 6	1.1	6	Landing flare improving, but speed on final wasn't stable. Small corrections for approach & crosswind. Don't stop flare when adding crosswind correction. After touchdown continue adding aileron while stopping to correct for wind and keep wings level.

Table 3 – 'Average' Student Performance

Flight	Dual Given	Number of Landings	Comments
Landing Lesson 1	1.3	8	Good attempt at first landings. Keep aileron correction in to the wind and increase deflection as you get slower as required especially upon touchdown.
Landing Lesson 2	1.5	8	Landing sequence is good, but still need work on airspeed control on final approach and flaring. Generally flaring too late and flat.
Landing Lesson 3	1.4	6	Improved but still landing firm occasionally. As aircraft slows over the runway, need to add increasing back pressure to prevent it from sinking into the runway with the nose dropping.
Landing Lesson 4	1.4	7	Keep working on stabilizing the final approach
Landing Lesson 5	1.6	9	Good job all the way through the flare. Work on flaring lower but not diving toward the runway. Let speed bleed off during flare to get the nose up and not land flat
Landing Lesson 6	1.1	6	Better job with adding backpressure during landing flare. Keep focusing on close to full stall landing. Better airspeed control.

Table 4 – ‘Good’ Student Performance

Flight	Dual Given	Number of Landings	Comments
Landing Lesson 1	1.3	8	Good attempt at first landings. Keep aileron correction in to the wind and increase deflection as you get slower as required especially upon touchdown.
Landing Lesson 2	1.5	8	Landing sequence is ‘good’, but still need work on airspeed control on final approach and flaring. Generally flaring too late and flat.
Landing Lesson 3	1.4	6	Improved but still landing firm occasionally. As aircraft slows over the runway, need to add increasing back pressure to prevent it from sinking into the runway with the nose dropping.
Landing Lesson 4	1.4	7	Good job monitoring speed. Good improvement on flare. Keep the crosswind correction in and land in a sideslip
Landing Lesson 5	1.6	9	Your landings are looking (and feeling) good, good flares and touchdowns
Landing Lesson 6	1.1	6	Nice rudder for sideslip during flare. Nice flares. Good job adding backpressure after touchdown. Good job on brakes

3.4.3 Digital Conditions

Graphical Images of Landings

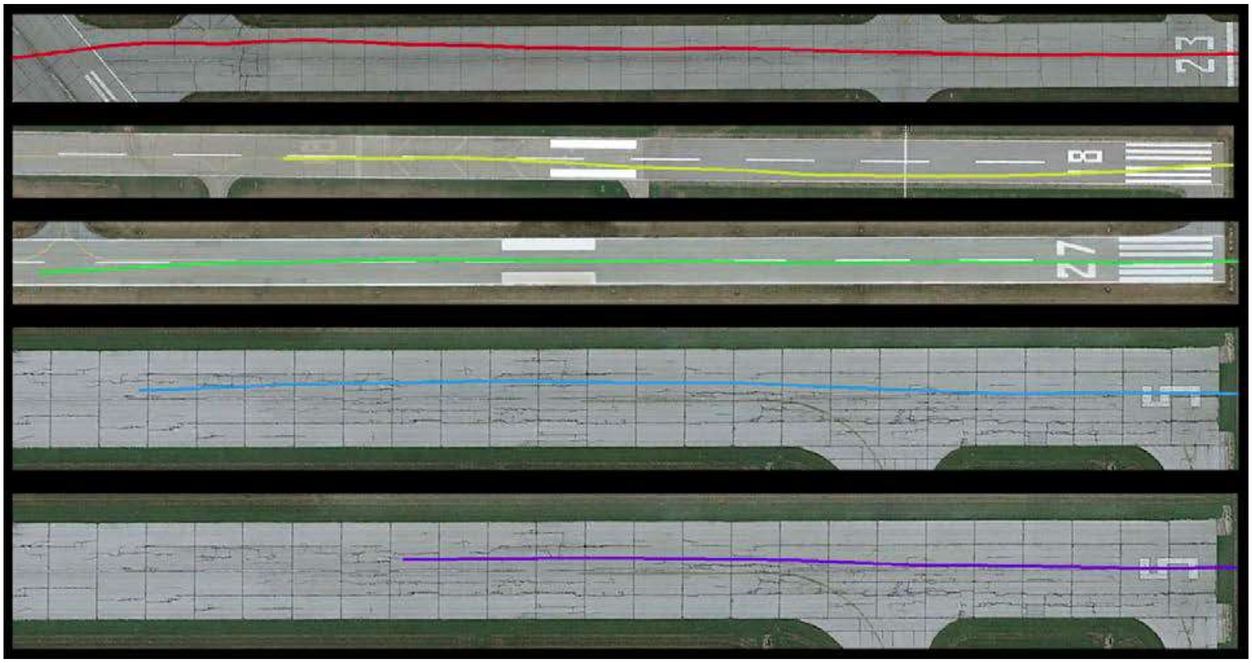


Figure 1 – Digital Landing Visualization – ‘Poor’ Performance

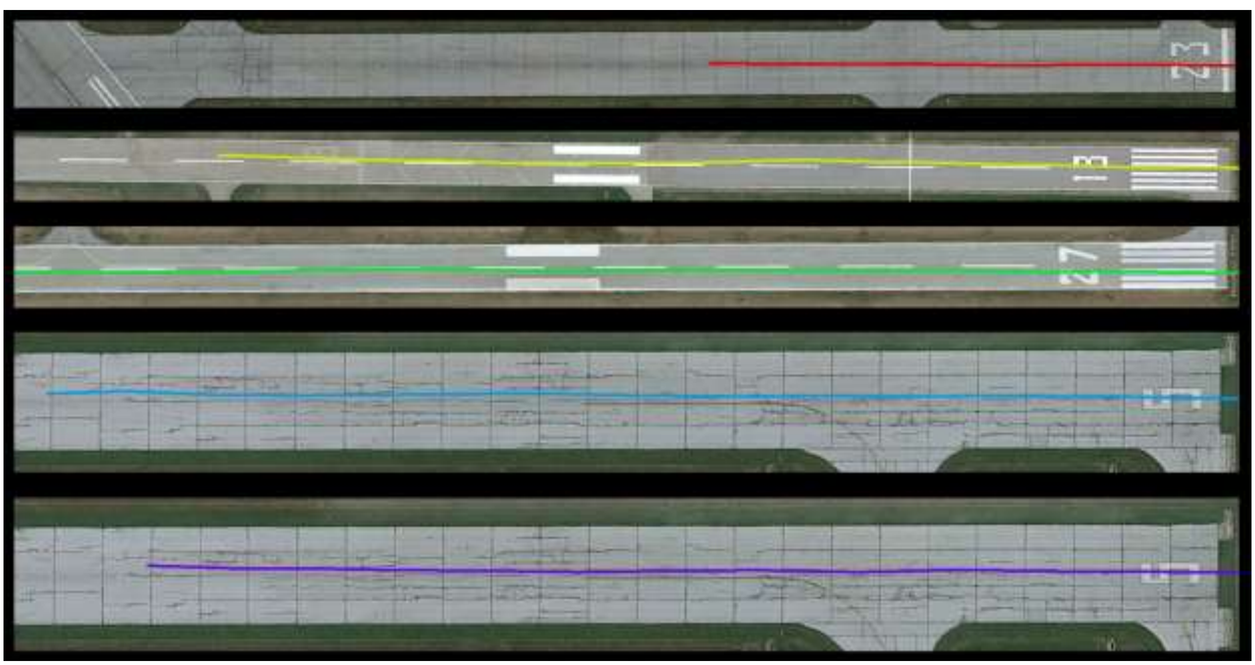


Figure 2 – Digital Landing Visualization – ‘Average’ Performance

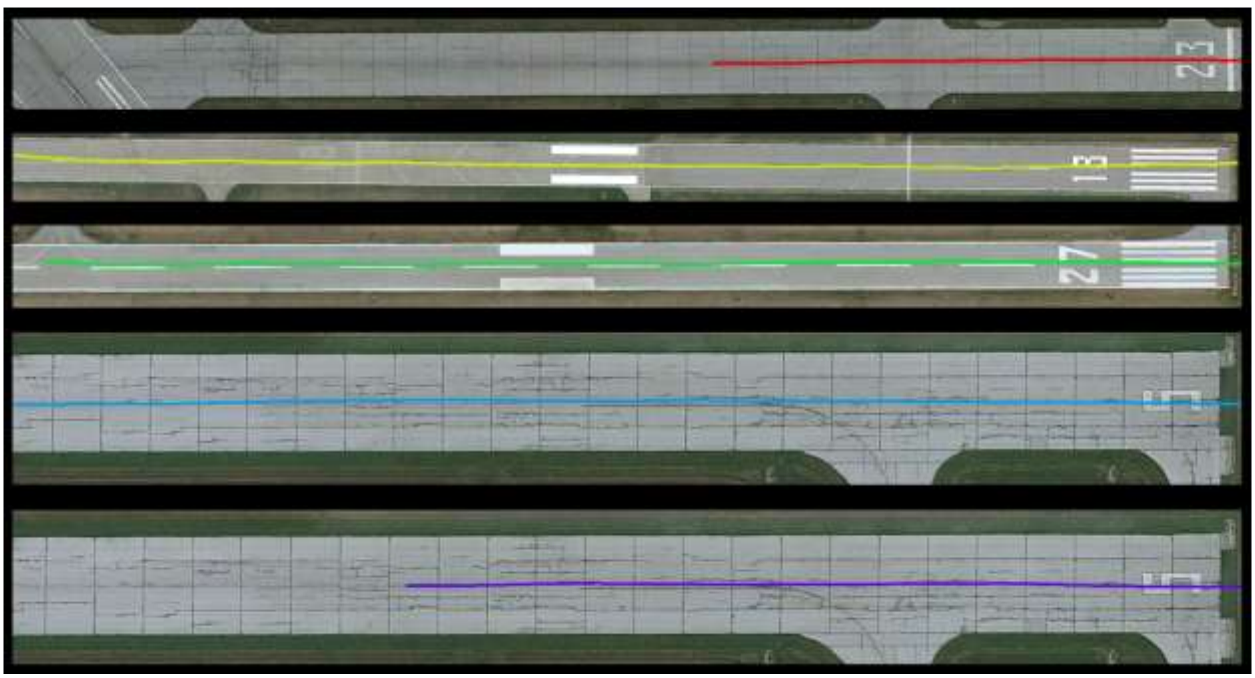


Figure 3 – Digital Landing Visualization – ‘Good’ Performance

Numerical Performance of Landings

Number of Landings	Minimum Centerline Deviation	Maximum Centerline Deviation	Average Centerline Deviation
44	5.33	26.78	11.9336

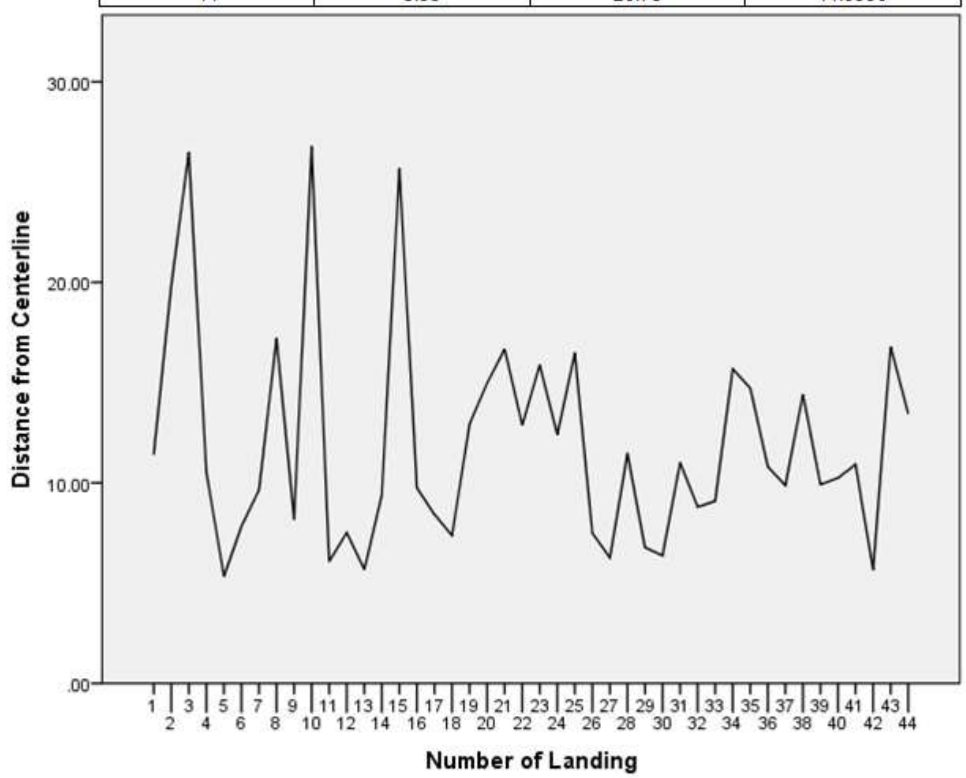


Figure 4 – Numerical Landing Visualization – ‘Poor’ Performance

Number of Landings	Minimum Centerline Deviation	Maximum Centerline Deviation	Average Centerline Deviation
44	3.45	11.47	7.1173

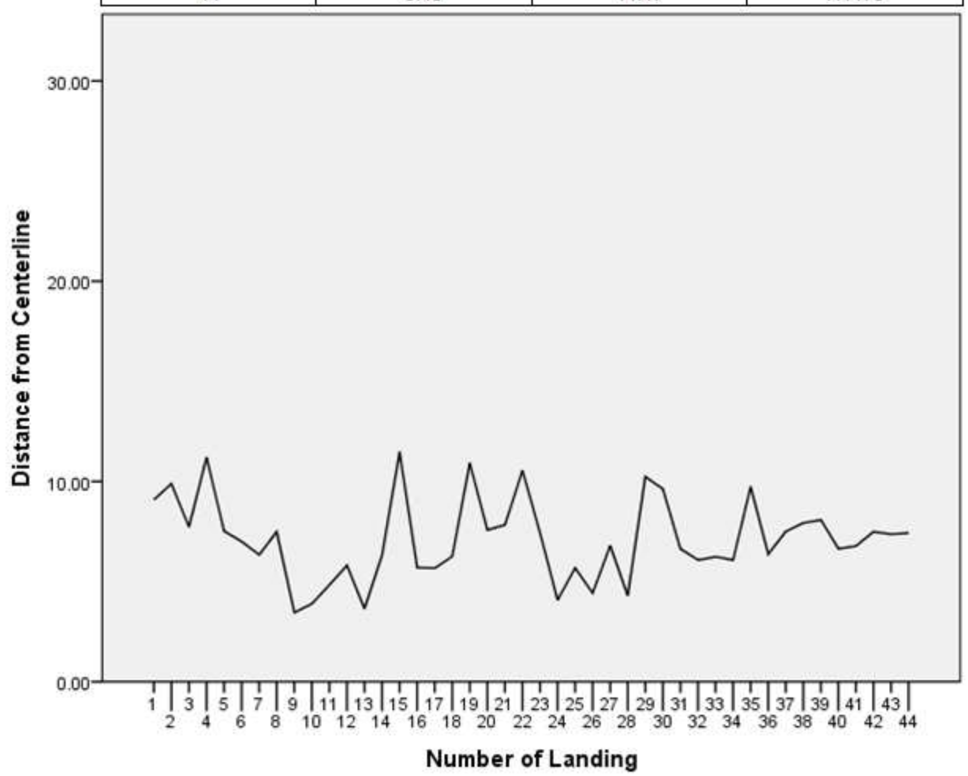


Figure 5 – Numerical Landing Visualization – ‘Average’ Student

Number of Landings	Minimum Centerline Deviation	Maximum Centerline Deviation	Average Centerline Deviation
44	1.43	9.65	4.5793

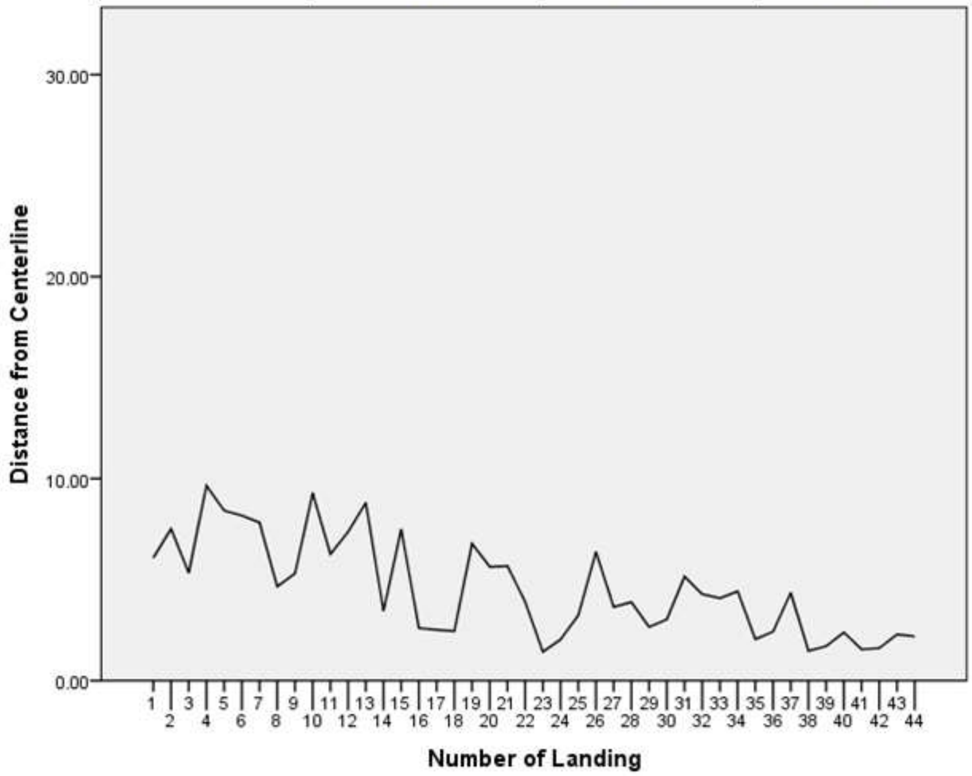


Figure 6 – Numerical Landing Visualization – 'Good' Performance

3.4.4 Student Profile Validation

Following the development of the levels of student performance, the information was presented to a flight instructor with more than five years of experience and 1000 hours of dual flight instruction given. This flight instructor made assessments of the accuracy of the representation of each level of student performance utilizing the digital information only for each level of performance developed. The profiles were presented individually and the flight instructor was asked to provide an assessment of performance for each profile. The flight instructor was also presented with the traditional logbook information only and was asked to provide a similar assessment for each student profile created. The flight instructor concurred with the accuracy of the representation of the student profiles for each level of student performance and for each type of information presented.

3.5 Variables

In order to measure how instructors evaluate students for solo flight, data was collected through a seven-point Likert-type scale on four dimensions: student landing competence, endorsement readiness, mentorability, likeability. Student landing competence had three items that measured the flight instructor's assessment of the student's ability to perform a landing. Participants were asked to determine if a hypothetical student was competent in landings, if they had the necessary skills for landing tasks, and if the student was able to handle the landing tasks. These three items for landing competence were averaged for a singular score for each participant as a measure to be used in statistical analysis. Endorsement readiness had three items that measured the flight instructor's readiness to provide a solo endorsement to the student. Participants were asked to determine if they would provide a solo endorsement to the student, to determine the likelihood that they would provide an endorsement for the final flight test, and to determine the likelihood that the student would successfully pass the flight test. These three items for endorsement readiness were averaged for a singular score for each participant as a measure to be used in statistical analysis. Mentorability had three items that measured the willingness of the flight instructor to provide support and guidance in the student's career aspirations as a pilot. Participants were asked to indicate if they would encourage the hypothetical student to continue their pursuit of a flight degree if they were considering changing majors, to provide educational advice concerning other aviation majors besides flight training, and if they would provide extra help towards a flying task in which the student was struggling with flight training. These three items for mentorability were averaged for a singular score for each participant as a measure to be used in statistical analysis. Likeability had three items that measured flight instructors' perceptions of the hypothetical student. Participants were asked how well they liked the flight student, if they wanted to get to know the student better, and whether or not the student would work well in a multi-crew cockpit environment. These three items for likeability were averaged for a singular score for each participant as a measure to be used in statistical analysis. The measures for landing competence, endorsement readiness, mentorability, and likeability were adapted from a study conducted by Moss-Racusin, Dovidio, Brescoll, Graham, and Handelsman (2012) to determine how faculty evaluated students' science abilities and their potential success as a graduate student after completion of their undergraduate degree.

3.5.1 Cognitive Load Items

Data on the cognitive load of the flight instructors for each of the above determinations was made using a nine-point Likert-type scale. This measurement has been used extensively by Sweller and Paas (2017) and has been labeled the “Paas Scale”. The Paas Scale consists of a single item asking participants to indicate how much effort they devoted to the task at hand with responses commonly made on a nine-point scale that rates effort from extremely low to extremely high (Sweller, & Paas, 2017).

3.6 Data Analysis

A 3 (‘poor’ vs ‘average’ vs ‘good’) x 3 (traditional vs digital vs traditional + digital) ANOVA factorial analysis was used for data analysis. Student performance level and type of information available to the flight instructor were the two factors analyzed. A Post Hoc Tukey analysis was utilized for comparison of factors after the ANOVA analysis.

CHAPTER 4 RESULTS

This study investigated the ability and cognitive load of instructors as they made assessments of student readiness for solo flight using either digital information provided by the aircraft (digital condition), a traditional logbook (traditional condition), or by using the combination of both the traditional logbook and digital information (traditional + digital condition). Demographic information and results from statistical analysis are discussed in this chapter.

4.1 Participants

After the recruitment process, a total of 298 participants gave consent to begin the study and initiated the Qualtrics survey. There were 49 participants that did not complete one or more questions within a section, failed to answer any questions besides the demographic information, or initiated the survey but did not answer any questions. These participants were removed from data analysis due to incomplete data. The final sample of the study included 249 participants, who completed all items in the survey. Survey participants had between one month and over 50 years of experience with an average of 19.88 years of experience as a flight instructor. Their total flight time ranged from 282 hours to over 5000 hours of experience with an average of 3129.9 hours. The amount of flight instruction given ranged from five hours to over 5000 hours with an average of 1696.04 hours of dual given. The number of students that the participants have endorsed for solo flight operations ranged from 0 to over 100 with the average being 24.6 students. Participants were asked to identify which region of the United States that they predominately provide flight instruction and the following areas were represented: Midwest – 103, South – 69, West – 37, Northeast – 31, and Pacific – 9. Table 5 shows the breakdown of participants in different experimental conditions.

Table 5 – Participants Who Completed the Survey

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="3">Student Performance</td><td colspan="2" rowspan="2">Totals</td></tr>
<tr><td>‘Poor’</td><td>‘Average’</td><td>‘Good’</td></tr>
<tr><td rowspan="3">Performance Data</td><td>Traditional</td><td>28</td><td>33</td><td>31</td><td>92</td><td rowspan="3">249</td></tr>
<tr><td>Traditional + Digital</td><td>30</td><td>23</td><td>30</td><td>83</td></tr>
<tr><td>Digital</td><td>25</td><td>22</td><td>27</td><td>74</td></tr>
<tr><td colspan="2" rowspan="2">Totals</td><td>83</td><td>78</td><td>88</td><td colspan="2" rowspan="2"></td></tr>
<tr><td colspan="3">249</td></tr>
</table>

4.1.1 Participant Expertise

The number of flight instruction hours given during their career was used to measure participant expertise. This is a standard measurement of experience used within the aviation industry. Table 6 shows the hours of flight instruction given for participants in the experimental groups.

Table 6 – Survey Participant Flight Instruction Hours Given

	N	Mean	Std. Deviation	Std. Error
‘Poor’ Landing Performance				
Traditional Condition	28	1649.68	1533.11	289.73
Traditional + Digital Condition	30	1830.37	1577.75	288.06
Digital Condition	25	1735.64	1577.25	315.45
‘Average’ Landing Performance				
Traditional Condition	33	1462.55	1432.94	249.44
Traditional + Digital Condition	23	1917.13	1550.76	323.36
Digital Condition	22	1536.73	1362.76	290.54
‘Good’ Landing Performance				
Traditional Condition	31	1796.39	1713.67	307.78
Traditional + Digital Condition	30	1681.10	1391.39	254.03
Digital Condition	27	1686.44	1441.08	277.34

An ANOVA analysis was conducted to compare the means of flight instruction hours given for each of the participant groups. There was no significant effect for the instructional hours given on the type of information available and the level of student performance $F(8, 240) = .242$, $p=.982$. This indicates that participants were similar in their flight instruction experience in each of the groups.

4.2 Descriptive Statistics

Table 7 shows the descriptive statistics for the ratings of landing competence that participants gave for the flight student in each of the condition combinations. The mean, median, and standard deviation is shown for each of the nine possible combinations of conditions that participants could have been asked to evaluate (traditional, traditional + digital, digital) by (‘poor’, ‘average’, and ‘good’). The average of three survey items were used to measure instructors’ ratings for the student’s landing competence. Landing competence is an assessment of the skillset for aircraft landings and an assessment of the proficiency level of the student for aircraft landing maneuver. Survey participants were asked to; determine the level of precision of the landings, the level of skill of the student performing the landings, and the level of qualifications of the student for landing performance.

Table 7 – Landing Competence Descriptive Statistics

	Mean	Median	Standard Deviation
‘Poor’ Landing Performance			
Traditional Condition	3.93	4.00	1.26
Traditional + Digital Condition	4.27	5.00	1.37
Digital Condition	5.31	5.33	0.80
‘Average’ Landing Performance			
Traditional Condition	5.25	5.67	1.02
Traditional + Digital Condition	5.10	5.33	1.07
Digital Condition	5.23	5.33	0.92
‘Good’ Landing Performance			
Traditional Condition	6.01	6.00	0.57
Traditional + Digital Condition	5.93	6.00	0.75
Digital Condition	5.64	6.00	0.96

Table 8 shows the descriptive statistics for the ratings of endorsement readiness that participants gave for the flight student in each of the condition combinations. The mean, median, and standard deviation is shown for each of the nine possible combinations of conditions that participants could have been asked to evaluate (traditional, traditional + digital, digital) by (‘poor’, ‘average’, and ‘good’). Endorsement readiness is an assessment as to whether or not the flight instructor would be willing to complete the paperwork necessary for the flight student to progress to the next phase of training. The average of three survey items were used to measure instructors’ ratings of the student’s endorsement readiness. Survey participants were asked to indicate whether or not they would; endorse the student for solo flight, endorse the student for the certification flight test, and to determine the likelihood that the student would pass the certification flight test at the end of their flight training.

Table 8 – Endorsement Readiness Descriptive Statistics

	Mean	Median	Standard Deviation
‘Poor’ Landing Performance			
Traditional Condition	4.21	4.67	1.05
Traditional + Digital Condition	4.41	4.67	1.34
Digital Condition	4.93	5.00	1.30
‘Average’ Landing Performance			
Traditional Condition	4.99	5.00	1.33
Traditional + Digital Condition	5.01	5.67	1.45
Digital Condition	5.18	5.50	1.14
‘Good’ Landing Performance			
Traditional Condition	5.58	6.00	1.13
Traditional + Digital Condition	5.78	5.67	1.13
Digital Condition	5.12	5.67	1.55

Table 9 shows the descriptive statistics for the ratings of mentorability that participants gave for the flight student in each of the condition combinations. The mean, median, and standard deviation is shown for each of the nine possible combinations of conditions that participants could have been asked to evaluate (traditional, traditional + digital, digital) by (‘poor’, ‘average’, and ‘good’). Mentorability is an assessment of the willingness of the survey participant to dedicate effort towards the flight student as they pursued their career objectives in the aviation industry. The average of three survey items were used to measure instructors’ ratings of the student’s mentorability. Participants were asked to indicate if they would encourage the hypothetical student to remain in a flight degree if they were considering changing majors, to provide educational advice concerning other aviation majors besides flight training, and if they would provide extra help towards a flying task in which the student was struggling with flight training.

Table 9 – Mentorability Descriptive Statistics

	Mean	Median	Standard Deviation
‘Poor’ Landing Performance			
Traditional Condition	6.12	6.33	0.73
Traditional + Digital Condition	6.21	6.50	0.94
Digital Condition	6.16	6.67	0.93
‘Average’ Landing Performance			
Traditional Condition	6.12	6.33	0.91
Traditional + Digital Condition	6.42	6.67	0.82
Digital Condition	6.00	6.33	1.00
‘Good’ Landing Performance			
Traditional Condition	6.00	6.00	0.83
Traditional + Digital Condition	6.42	6.67	0.66
Digital Condition	6.11	6.33	1.00

Table 10 shows the descriptive statistics for ratings of likeability that participants gave for the flight student in each of the condition combinations. The mean, median, and standard deviation is shown for each of the nine possible combinations of conditions that participants could have been asked to evaluate (traditional, traditional + digital, digital) by (‘poor’, ‘average’, and ‘good’). Likeability is an assessment of the personal perception that the survey participants have concerning the student condition presented in the survey. The average of three survey items were used to measure instructors’ ratings of the student’s likeability. Survey participants were asked to indicate whether or not they liked the student, whether they would like to get to know the student, and whether or not they would work well in a multi-crew cockpit environment.

Table 10 – Likeability Descriptive Statistics

	Mean	Median	Standard Deviation
‘Poor’ Landing Performance			
Traditional Condition	4.85	4.67	0.80
Traditional + Digital Condition	5.02	4.83	0.92
Digital Profile Only	5.00	5.33	0.87
‘Average’ Landing Performance			
Traditional Condition	4.97	5.00	0.94
Traditional + Digital Condition	5.29	5.33	0.84
Digital Profile Only	4.98	4.67	0.88
‘Good’ Landing Performance			
Traditional Condition	5.14	5.33	0.91
Traditional + Digital Condition	5.59	5.67	0.94
Digital Profile Only	4.90	4.67	0.95

Table 11 shows the descriptive statistics for cognitive load, which is a self-determined measure of the amount of effort required to determine the scores for the assessments of the flight student concerning landing competence, endorsement readiness, mentorability, and likeability. These four measures of cognitive load were averaged to determine the cognitive load for the student assessment. The mean, median, and standard deviation is shown for each of the nine possible combinations of conditions that participants could have been asked to evaluate (traditional, traditional + digital, digital) by (‘poor’, ‘average’, and ‘good’).

Table 11 – Cognitive Load Descriptive Statistics

	Mean	Median	Standard Deviation
‘Poor’ Landing Performance			
Traditional Condition	4.47	4.88	1.17
Traditional + Digital Condition	3.95	3.88	1.39
Digital Condition	3.79	3.75	1.02
‘Average’ Landing Performance			
Traditional Condition	3.83	3.50	1.41
Traditional + Digital Condition	3.89	4.00	0.95
Digital Condition	3.91	3.75	1.22
‘Good’ Landing Performance			
Traditional Condition	4.02	4.25	1.19
Traditional + Digital Condition	3.88	3.75	1.42
Digital Condition	5.44	5.50	1.55

4.4 ANOVA Analysis and Post-Hoc Tests

4.4.1 Landing Competence

Leven’s test on the data for landing competence violated the assumption of homogeneity of variance ($p < 0.00$). Due to this assumption violation, the Welch and Brown-Forsythe statistics were utilized in order to determine if the conditions had significant influence on the flight instructor’s assessment of landing competence. Results from the ANOVA suggested that the condition had a significant influence on flight instructor’s assessment of landing competence $F(8, 96.757) = 13.385$, $p = 0.000$.

Table 12 displays the ANOVA results for landing competence. A one-way ANOVA was performed so that the Welch and Brown-Forsythe statistics in Table 13 were considered for significance due to the violation of the assumption of homogeneity of variance.

Table 12 – Landing Competence ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	113.783	8	14.223	14.285	.000
Within Groups	238.964	240	.996		
Total	352.747	248			

Table 13 – Landing Competence Welch & Brown-Forsythe

	Statistic	df1	df2	Sig.
Welch	13.385	8	96.757	.000
Brown-Forsythe	14.311	8	196.817	.000

Games-Howell post-hoc analysis revealed that assessments of landing competence with the traditional condition on ‘poor’ performing students was significantly lower than all conditions except traditional + digital condition on ‘poor’ performing students. Post-hoc tests also revealed that traditional + digital condition on ‘poor’ performing students was evaluated significantly lower than digital condition with both ‘poor’ students and all conditions for ‘good’ performing students. Finally, post-hoc analysis revealed that evaluations for traditional condition on ‘good’ performing students was significantly higher than all conditions for ‘poor’ and ‘average’ performing students. The results of the Games-Howell post-hoc are shown in Tables 14 – 16.

Table 14 – Traditional Condition Landing Competence Games-Howell Post-Hoc Results

Group	Group	Mean Difference	Std. Error	Sig.
Traditional Condition on ‘Poor’ Performing Students	Digital Condition			
	‘Poor’ Performing Students	-1.38	0.29	0.001
	‘Average’ Performing Students	-1.30	0.31	0.003
	‘Good’ Performing Students	-1.71	0.30	0.000
	Traditional Condition			
	‘Average’ Performing Students	-1.32	0.30	0.001
	‘Good’ Performing Students	-2.08	0.26	0.000
	Traditional + Digital Condition			
	‘Poor’ Performing Students	-0.34	0.34	0.986
	‘Average’ Performing Students	-1.17	0.33	0.020
	‘Good’ Performing Students	-2.00	0.27	0.000
Traditional Condition on ‘Average’ Performing Students	Digital Condition			
	‘Poor’ Performing Students	-0.05	0.24	1.000
	‘Average’ Performing Students	0.03	0.27	1.000
	‘Good’ Performing Students	-0.39	0.26	0.843
	Traditional Condition			
	‘Good’ Performing Students	-0.76	0.21	0.015
	Traditional + Digital Condition			
	‘Poor’ Performing Students	0.99	0.31	0.053
	‘Average’ Performing Students	0.15	0.29	1.000
	‘Good’ Performing Students	-0.68	0.22	0.081
Traditional Condition on ‘Good’ Performing Students	Digital Condition			
	‘Poor’ Performing Students	0.70	0.19	0.016
	‘Average’ Performing Students	0.78	0.22	0.030
	‘Good’ Performing Students	0.37	0.21	0.715
	Traditional + Digital Condition			
	‘Poor’ Performing Students	1.74	0.27	0.000
	‘Average’ Performing Students	0.91	0.25	0.020
	‘Good’ Performing Students	0.08	0.17	1.000

Table 15 – Traditional + Digital Condition Landing Competence Games-Howell Post-Hoc Results

Group	Group	Mean Difference	Std. Error	Sig.
Traditional + Digital Condition on ‘Poor’ Performing Students	Digital Condition			
	‘Poor’ Performing Students	-1.04	0.30	0.025
	‘Average’ Performing Students	-0.96	0.32	0.086
	‘Good’ Performing Students	-1.38	0.31	0.002
	Traditional + Digital Condition			
	‘Average’ Performing Students	-0.83	0.34	0.260
	‘Good’ Performing Students	-1.67	0.28	0.000
Traditional + Digital Condition on ‘Average’ Performing Students	Digital Condition			
	‘Poor’ Performing Students	-0.21	0.27	0.998
	‘Average’ Performing Students	-0.13	0.30	1.000
	‘Good’ Performing Students	-0.54	0.29	0.640
	Traditional + Digital Condition			
	‘Good’ Performing Students	-0.83	0.26	0.065
Traditional + Digital Condition on ‘Good’ Performing Students	Digital Condition			
	‘Poor’ Performing Students	0.63	0.21	0.095
	‘Average’ Performing Students	0.71	0.24	0.108
	‘Good’ Performing Students	0.29	0.23	0.936

Table 16 – Digital Condition Landing Competence Games-Howell Post-Hoc Results

Group	Group	Mean Difference	Std. Error	Sig.
Digital Condition on ‘Poor’ Performing Students	Digital Condition			
	‘Average’ Performing Students	0.08	0.25	1.000
	‘Good’ Performing Students	-0.34	0.24	0.903
Digital Condition on ‘Average’ Performing Students	Digital Condition			
	‘Good’ Performing Students	-0.41	0.27	0.832

4.4.2 Endorsement Readiness

The results of the 3 (‘poor’ vs ‘average’ vs ‘good’) x 3 (traditional vs digital vs traditional + digital) ANOVA suggested that resource available (traditional vs digital vs traditional + digital) was not statistically significant $F(2, 240) = .287$, $p = 0.75$. Results also suggested that student performance was statistically significant, $F(2, 240) = 10.865$, $p < 0.001$ indicating a significant difference between ‘poor’ performing students, ‘average’ performing students, and ‘good’ performing students. Finally, the interaction between resource available and student performance was not statistically significant, $F(2, 240) = 1.679$, $p = 0.15$. Table 17 shows the ANOVA results for flight instructor assessment of endorsement readiness of the student.

Table 17 – ANOVA Results for Endorsement Readiness

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	49.407[a]	8	6.176	3.799	.000
Intercept	6126.750	1	6126.750	3768.939	.000
Resource Available	.934	2	.467	.287	.750
Student Performance	35.326	2	17.663	10.865	.000
Resource Available * Student Performance	10.920	4	2.730	1.679	.155
Error	390.142	240	1.626		
Total	6677.889	249			
Corrected Total	439.548	248			

A Tukey post-hoc analysis on student performance revealed that ‘poor’ performing students were scored significantly lower than ‘average’ performing students ($p = 0.018$) and ‘good’ performing students ($p < 0.00$). The results of the Tukey post-hoc are shown in Table 18.

Table 18 – Student Performance on Endorsement Readiness Tukey Post-Hoc

Student Performance		Mean Difference	Std. Error	Sig.
‘Average’ Performing Student	‘Good’ Performing Student	-0.39	0.20	0.125
	‘Poor’ Performing Student	0.55	0.20	0.018
‘Good’ Performing Student	‘Average’ Performing Student	0.39	0.20	0.125
	‘Poor’ Performing Student	0.94	0.20	0.000
‘Poor’ Performing Student	‘Average’ Performing Student	-0.55	0.20	0.018
	‘Good’ Performing Student	-0.94	0.20	0.000

4.4.3 Mentorability

The results of the 3 (‘poor’ vs ‘average’ vs ‘good’) x 3 (traditional vs digital vs traditional + digital) ANOVA suggested that resource available (traditional vs digital vs traditional + digital) was not statistically significant $F(2, 240) = 2.545$, $p = 0.81$. Results also suggested that student performance was not statistically significant, $F(2, 240) = .009$, $p = .991$. Finally, the interaction between resource available and student performance was not statistically significant, $F(4, 240) = .465$, $p = .761$. Table 19 shows the ANOVA results for flight instructor assessment of mentorability of the student.

Table 19 – ANOVA Results for Mentorability

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5.177[a]	8	.647	.847	.562
Intercept	9330.183	1	9330.183	12217.219	.000
Resource Available	3.887	2	1.944	2.545	.081
Student Performance	.014	2	.007	.009	.991
Resource Available * Student Performance	1.421	4	.355	.465	.761
Error	183.286	240	.764		
Total	9675.889	249			
Corrected Total	188.463	248			

4.4.4 Likeability

The results of the 3 (‘poor’ vs ‘average’ vs ‘good’) x 3 (traditional vs digital vs traditional + digital) ANOVA suggested that resource available (traditional vs digital vs traditional + digital) was statistically significant $F(2, 240) = 3.594$, $p = 0.29$ indicating a significant difference between traditional condition, digital condition, and traditional + digital condition. Results also suggested that student performance was not statistically significant, $F(2, 240) = 1.374$, $p = 0.185$. Finally, the interaction between resource available and student performance was not statistically significant, $F(4, 240) = .969$, $p = 0.425$. Table 20 shows the ANOVA results for flight instructor assessment of likeability of the student.

Table 20 – ANOVA Results for Likeability

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	12.155[a]	8	1.519	1.878	.064
Intercept	6322.860	1	6322.860	7814.637	.000
Resource Available	5.815	2	2.908	3.594	.029
Student Performance	2.748	2	1.374	1.698	.185
Resource Available * Student Performance	3.136	4	.784	.969	.425
Error	194.185	240	.809		
Total	6643.111	249			
Corrected Total	206.340	248			

A Tukey post-hoc analysis was performed on instructor resource revealed that traditional + digital condition scored significantly higher than digital condition information ($p = 0.048$). The results of the Tukey post-hoc are shown in Table 21.

Table 21 – Resource Available on Student Likeability Tukey Post-Hoc

Resource Available		Mean Difference	Std. Error	Sig.
Digital Condition	Traditional Condition	-0.03	0.14	0.976
	Traditional + Digital Condition	-0.34	0.14	0.048
Traditional Condition	Digital Condition	0.03	0.14	0.976
	Traditional + Digital Condition	-0.31	0.14	0.059
Traditional + Digital Condition	Digital Condition	0.34	0.14	0.048
	Traditional Condition	0.31	0.14	0.059

4.4.5 Cognitive Load

The results of the 3 (‘poor’ vs ‘average’ vs ‘good’) x 3 (traditional vs digital vs traditional + digital) ANOVA suggested that resource available (traditional vs digital vs traditional + digital) was not statistically significant $F(2, 240) = 2.637$, $p = 0.74$. Results also suggested that student performance was statistically significant, $F(2, 240) = 4.214$, $p = 0.016$ indicating a significant difference between ‘poor’ performing students, ‘average’ performing students, and ‘good’ performing students. Finally, the interaction between resource available and student performance was statistically significant, $F(4, 240) = 5.807$, $p < 0.001$ indicating a significant difference when the interaction between resource available and student performance is considered. Table 22 shows the ANOVA results for flight instructor assessment of the cognitive load necessary for survey participants to assess landing competence, endorsement readiness, mentorability, and likeability.

Table 22 – ANOVA Results for Cognitive Load

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	61.326[a]	8	7.666	4.653	.000
Intercept	4179.842	1	4179.842	2536.873	.000
Resource Available	8.689	2	4.344	2.637	.074
Student Performance	13.885	2	6.942	4.214	.016
Resource Available * Student Performance	38.269	4	9.567	5.807	.000
Error	395.432	240	1.648		
Total	4705.000	249			
Corrected Total	456.758	248			

A Tukey post-hoc analysis was performed on student performance and revealed that students with ‘good’ performing were rated higher (p = 0.021) in cognitive load effort than ‘average’ performing students. The results of the Tukey post-hoc are shown in Table 23.

Table 23 – Student Performance on Cognitive Load Tukey Post-Hoc

Student Performance		Mean Difference	Std. Error	Sig.
‘Average’ Performing Student	‘Good’ Performing Student	-0.54	0.20	0.021
	‘Poor’ Performing Student	-0.21	0.20	0.565
‘Good’ Performing Student	‘Average’ Performing Student	0.54	0.20	0.021
	‘Poor’ Performing Student	0.33	0.20	0.213
‘Poor’ Performing Student	‘Average’ Performing Student	0.21	0.20	0.565
	‘Good’ Performing Student	-0.33	0.20	0.213

A Tukey post-hoc analysis was performed on the type of flight instructor resource available and revealed that digital condition was rated higher (p = 0.032) in cognitive load effort than traditional + digital condition. The results of the Tukey post-hoc are shown in Table 24.

Table 24 – Resource Available on Cognitive Load Tukey Post-Hoc

Resource Available		Mean Difference	Std. Error	Sig.
Digital Condition	Traditional Condition	0.34	0.20	0.210
	Traditional + Digital Condition	0.52	0.21	0.032
Traditional Condition	Digital Condition	-0.34	0.20	0.210
	Traditional + Digital Condition	0.18	0.19	0.624
Traditional + Digital Condition	Digital Condition	-0.52	0.21	0.032
	Traditional Condition	-0.18	0.19	0.624

A profile plot in Figure 7 was created to represent the interaction effect between resource available and student performance on the cognitive load of flight instructors. The profile plot indicates that the amount of cognitive load varies based the combination of student performance and the resource condition. The profile plot appears to indicate that the type of resource condition is a determinant of the cognitive load necessary when assessing a 'good' performing student, but is not a determinant with an 'average' and 'poor' performing student.

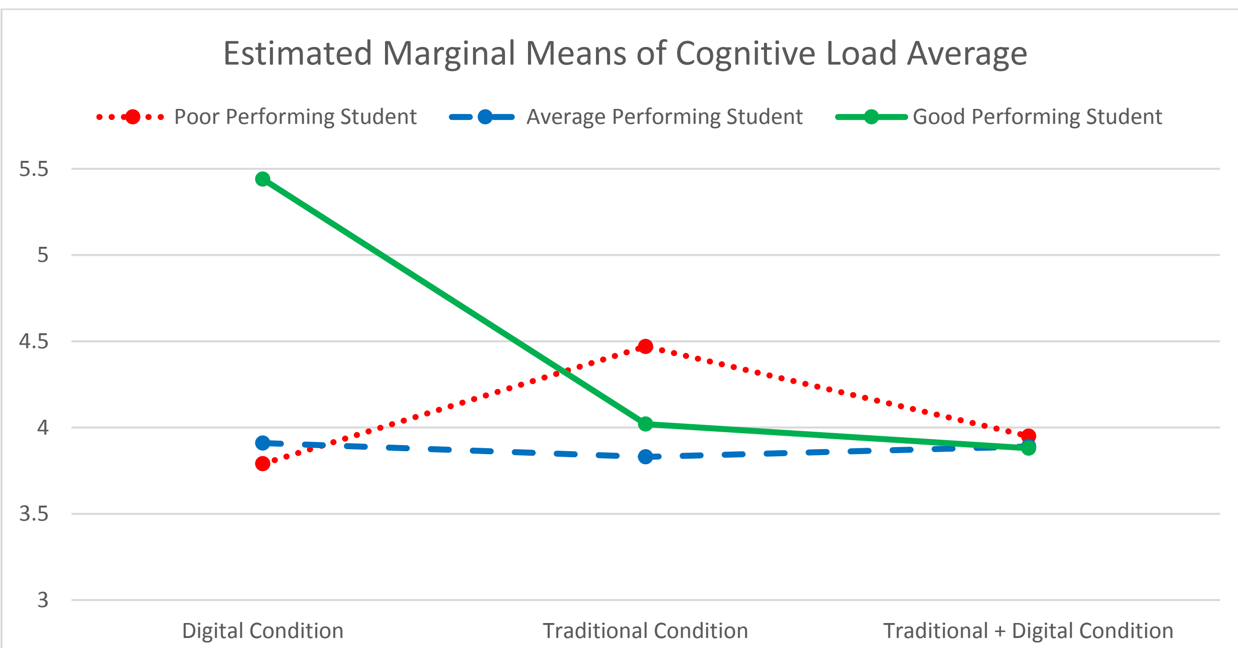


Figure 7 – Profile Plot of Cognitive Load

CHAPTER 5 DISCUSSION

Flight training is an expensive endeavor. The cost of an airplane and instructor can easily reach $200/hour and there is a minimum of 40 hours of flight time that must be accrued to become a private pilot. Additionally, the vast majority of instruction that is given in the United States is conducted by relatively inexperienced instructors as they are building flight time in order to qualify for a position with an airline. These two factors combine to create a situation where there is a risk of excessive expenses being paid for inefficient instruction. Additionally, the accuracy of student assessment is something that improves with experience but students early in a flight instructor’s career often receives feedback that could be dramatically improved by assessment tools. One of the ways that student assessment in flight instruction can be improved upon is by utilizing digital data that is available from the airplane. This would provide an opportunity to reduce the cost of flight training, improve the accuracy of student assessment, and make better informed decisions of student readiness for advancement in their flight training. This study investigated the ability and cognitive load of instructors as they made assessments of student readiness for solo flight while using a digital condition of information provided by the aircraft, a traditional condition, or a combination of both the digital and traditional condition. This chapter summarizes the findings and provides implications for using data and training pilots.

5.1 Discussion

Given three hypothetical student conditions (‘poor’, ‘average’, and ‘good’) across each of the conditions (traditional, digital, and traditional + digital), one would expect that CFIs would rate the student performance similar to the hypothetical student condition. However, participants did not rate students in a way that would match their profile in some of the conditions. The findings and possible hypothesis for why participants were not able to use the data provided (traditional, digital, or traditional + digital) to consistently rate the hypothetical student performance conditions are discussed.

5.1.1 Landing Competence

The first research question was to determine the influence of the use of digital flight data on instructor assessments of student competence for initial solo flight by a flight student. Student landing competence had three items that measured the flight instructor's assessment of the student's ability to perform a landing. Participants were asked to determine if a hypothetical student was competent in landings, if they had the necessary skills for landing tasks, and if the student was able to handle the landing tasks. These three items for landing competence were averaged for a singular score for each participant as a measure to be used in statistical analysis. Overall results suggested that there was a significant difference between the three conditions (traditional, digital, and traditional + digital). Post-hoc tests were conducted to examine difference between the three conditions for each of the three student performance levels ('poor', 'average', and 'good'). Comparisons of the combinations of information conditions and performance conditions are discussed further in the following sections.

Traditional Condition Comparisons

Results suggested that when evaluating students at any performance condition utilizing the traditional condition, participants rated them appropriately for all performance conditions. This is to say, 'poor' performing students were rated significantly lower than 'average' and 'good' performing conditions utilizing the traditional condition and 'good' performing students with the traditional condition were rated significantly higher than 'average' and 'poor' performing conditions utilizing the traditional condition. This result is to be expected when comparing student performance utilizing a traditional condition. It is also important to note that flight instructors are familiar with utilizing the traditional logbook when conducting student performance assessments. The results also suggested that when the digital condition was added to the traditional condition (traditional + digital condition) the results did not retain significance for the comparisons. The 'average' performing students utilizing traditional conditions were not rated significantly different than 'poor' and 'good' performing students utilizing traditional + digital condition. This would appear to indicate that when digital visualizations are added it creates less clear of a picture in the minds of the participants. This finding aligns with the study on medical patient handoffs by Young, et al. (2016), who found that when multiple sources of information were available there was an increase in extraneous cognitive load. It is possible that adding digital visualizations to the traditional logbook did not provide instructors with a unified visualization of student performance, rather it got utilized as a separate performance artifact that the participants then were required to evaluate. A unified visualization between digital and traditional measures, as well as, specific educational efforts will need to be taken prior to the realization of the full potential benefit of incorporating a digital visualization of student performance from aircraft data with a traditional logbook. Additionally, the use of multiple ways of viewing the data could provide better insight to the user in their assessment tasks (Buja et al., 1991; Green et al., 2009; Kosara et al., 2003; Tufte, 2001, 2006).

The results also indicated that 'poor' performing students were rated significantly lower for a traditional condition than in the digital condition. This appears to indicate that participants could not rate 'poor' performing students with similar accuracy when utilizing the traditional logbook versus digital visualizations. Previous research from Baker (2010) highlighted the identification of student abilities as a key purpose of visualizations and the findings in this study indicate that digital visualizations utilized in the digital condition did not provide flight instructors the information necessary to assess flight student abilities as well as the traditional condition. Lastly, 'average' performing students in the traditional condition were not significantly different from 'poor' and 'good' performing students utilizing the digital condition. This appears to indicate that participants could not utilize the digital visualizations to distinguish between an 'average' performing student in the traditional condition and 'poor' or 'good' performing student in the digital condition. Green et al. (2009) expressed the concern that the main purpose of the visualization of information is to facilitate the flow of reasoning and this flow is critical within the complex environment of flight instruction which is especially important to consider when developing visualizations (Ahlberg et al., 1992; Keim et al., 2008; Shneiderman, 2014; Van Wijk, 2005). The FAA has a list of requirements in Part 61 of Title 14 of the Code of Federal Regulations and the flight instructor must incorporate this flow of reasoning while determining the ability of students for safe aircraft operation that must be met prior to solo flight. Additionally, each flight instructor has their own perception of what performance requirements must be met by the student in order to demonstrate landing competency. Information that the flight instructor is utilizing is a variety of sources to think about, understand, and arrive at a logical conclusion in a sensible sequence and the visualizations provided to them must assist and not hinder this process. In this study, the instructors were better able to use the traditional logbook to evaluate and match the student performance yet the digital condition fell short of this objective. In concurrence with previous research, the digital condition needs to be revised so that it is better able to provide the information necessary for flight instructors to understand and make informed decisions regarding student performance.

Traditional + Digital Condition Comparisons

Results suggested that when evaluating ‘poor’ performing students in the traditional + digital condition, participants rated them significantly lower than ‘poor’ and ‘good’ performing students in the digital condition as well as ‘good’ performing student in traditional + digital condition. It is not surprising that ‘poor’ performing students were rated lower in landing competence than ‘good’ performing students regardless of the condition, but the significant difference between ‘poor’ performing students across different conditions was surprising. This appears to indicate that participants could not rate ‘poor’ performing students with the same accuracy when utilizing the traditional + digital or digital conditions. It could be possible that the digital condition did not provide scaffolds (such as, identification of landing improvement, problem areas to be addressed, and next steps in the instructional sequence) similar to traditional condition, which has been suggested as an important component in previous research (Larkin & Simon, 1987; Liu et al., 2008). Without proper scaffolds, an increase in cognitive load is possible during the assessment of student performance as the participants were utilizing the visualization in the digital condition (Dastani, 2002). It should be noted that the traditional + digital condition for ‘poor’ performing students was not statistically different for ‘average’ performing students in the digital or traditional + digital conditions. Furthermore, evaluations of ‘average’ performing students in the traditional + digital condition were not significantly different than ‘poor’ or ‘good’ performing students in the digital condition or ‘good’ performing students in the traditional + digital condition. This suggests that the participants had difficulty in combining the two resources from traditional logbook and digital visualizations when evaluating students in a way that could provide them an accurate picture of student’s performance. While previous research (Chandler & Sweller, 1991) has shown that the combination of diagrams and text can be a valuable tool to reduce extraneous load and visualizations for supporting information retrieval activities have significant utility (Morse et al., 2003), results from this study suggest that the combination of the traditional logbook with the digital visualizations did not help the instructors. Given that prior work has shown that written and drawn inputs can help increase understanding (Shelton, et al., 2016; Anderson et al., 2011), we need to engage certified flight instructors and get their input and guidance for future development of aviation visualizations.

Digital Condition Comparisons

Results suggested that when evaluating all student performance conditions with a digital condition, participants' ratings on landing competence were not significantly different with any performance condition. In other words, when a digital condition was utilized, participants were not able to accurately tell the difference between 'good', 'average', and 'poor' performing students. It is possible that the digital condition might have created extraneous cognitive load, which hindered the ability of the flight instructor to make an accurate assessment. Looking at the cognitive load results shows that for digital condition, participants had the highest cognitive load for 'good' performing students in the digital condition. This finding, in conjunction with previous research (Paas et al., 2003; Van Merrienboer & Sweller, 2005, Young et al., 2016), highlights the importance of ensuring that the additive effects of the forms of cognitive load do not make the task more difficult.

5.1.2 Endorsement Readiness

The second research question was to determine the influence of the use of digital flight data on instructor assessments of student readiness for an endorsement for initial solo flight. Providing an endorsement for solo flight is a significant step in the education sequence of a student working towards their private pilot certificate and is normally done with the utmost care and attention to detail. The weather, airspace, and traffic environment in which the student will ultimately fly solo are highly controlled and the flight instructor does all they can to ensure success and the avoidance of unforeseen issues. This aspect of the situation could affect a participant's responses regarding their willingness to provide an endorsement for a student with whom they've never flown. It is asking a lot for the participants to determine the readiness of the pilot for initial solo especially when they have never personally flown with the student in the past. Endorsement readiness had three items that measured the flight instructor's readiness to provide a solo endorsement to the student, the likelihood that the instructor would provide an endorsement for the final flight test, and the likelihood that the student would successfully pass the flight test. These three items for endorsement readiness were averaged for a singular score for each participant as a measure to be used in statistical analysis. That notwithstanding, results suggested that participants' evaluation of endorsement readiness was not influenced by the condition available to them (traditional, digital, or traditional + digital). There were, however, significant differences in how participants evaluated each of the student performance levels ('poor', 'average', and 'good') performing students. Not surprisingly, 'poor' performing students were rated significantly lower than 'average' and 'good' performing students; however, there were no significant differences between the 'average' and 'good' performing students. These results are not surprising given that the 'poor' performing student was not endorsement ready regardless of the condition that it was evaluated in by the participants. The lack of significant differences between 'good' and 'average' performing students suggests that participants were better able to determine whether or not a student was ready for a solo endorsement when their performance was on the extremes of either 'poor' or 'good' performance, but had difficulty distinguishing an 'average' student from a 'good' performing student.

Luan (2002) emphasized the importance of domain experts to be attuned to the meaning of the data visualizations in order for successful utilization of the information. It is possible that in the current study, either the digital condition did not accurately visualize the information concerning student flight performance or that CFIs were not fully ready to utilize the information to accurately evaluate a student to influence their decision to provide a solo endorsement. Galyardt and Goldin (2015) also pointed out that the visualization of digital data for educational assessments for the most recent observations are a better summary of the learner's mastery of a skill than the student's entire history of practice. The visualizations utilized in this study were for the entire footprint of traffic pattern practice and did not specifically highlight the degree of improvement of the student over time. It is possible that the use of limited visualizations to evaluate students in this study might not be sufficient for instructors' to accurately gauge their endorsement readiness for flight.

In addition, a CFI's reluctance to rely heavily on a digital visualization for student assessments could be understandable due to the fact that, in a very real sense, a CFI's certification is at stake when they provide this endorsement. There have been instances where an accident or incident has happened and the CFI's actions and decisions where brought into question and evaluated for potential administrative action. As pointed out by Thomas and Richards (2015), assessments of student flight performance are complex and there is significant variability between individual instructors in their decision-making process and there needs to be further development of how CFIs can be trained to use digital visualizations (Luan, 2002).

5.1.3 Mentorability

The third research question was to determine the influence of digital flight data on an instructor's willingness to mentor students who are preparing for initial solo flight. Mentorability had three items that measured the willingness of the flight instructor to provide support and guidance in the student's career aspirations as a pilot. Participants were asked to indicate if they would encourage the hypothetical student to remain in a flight degree if they were considering changing majors, to provide educational advice concerning other aviation majors besides flight training, and if they would provide extra help towards a flying task in which the student was struggling with flight training. These three items for mentorability were averaged for a singular score for each participant as a measure to be used in statistical analysis. The results for the assessment of mentorability exhibited no statistical difference between the conditions or student performance levels and there was also no significant interaction effect. Flight instructors are not commonly asked about their willingness to mentor a student and this may have caused confusion amongst the participants as to what that means. The FAA (2018) produced an advisory document called, "Best Practices for Mentoring in Flight Instruction" and within it they specify that a flight instructor mentor should have a personable mentality as well as being non-judgmental and to be a service to the student pilot. Furthermore, it is within the nature of pilots to be problem solvers and a student pilot that is struggling wouldn't necessarily affect a flight instructor's decision to continue working with them on their flight performance. One aspect that can have a negative impact for the flight instructor's continued effort is a lack of motivation or the unwillingness of the student to dedicate effort toward any homework or additional resources that have been assigned or recommended by the instructor. Additionally, flight instructors will often shift their efforts to other students if a student pilot does not show up on time or cancels multiple flight sessions. Thus, the student conditions used in this study might not have had enough information to influence an instructor's willingness to mentor a student.

However, with additional information like student reflections of their training could have influenced their assessment of student similar to the study by Stover et al. (2016). In that study, the authors had the ability to grasp a deeper understanding of the students' awareness of the topics rather than a simple assessment of the answers that they provided on an assignment. In much the same way, reflections of student effort and awareness of their strengths and weaknesses might provide an instructor with an assessment of a student's willingness to accomplish what is necessary for further development of their piloting abilities and therefore, enable an instructor to make an informed decision of their willingness to mentor such student. Dyckhoff (2012) stressed the importance of having indicators tied to specific questions and the results from this study support that aspect in regards to the assessment of a student and the willingness of a CFI to mentor them in the aviation industry.

5.1.4 Likeability

The fourth research question was to determine the influence of the use of digital flight data on instructor assessments of student likeability who are preparing for initial solo flight. Likeability had three items that measured how well the instructor might like the flight student, if the instructor wanted to get to know the student better, and whether or not the student would work well in a multi-crew cockpit environment. These three items for likeability were averaged for a singular score for each participant as a measure to be used in statistical analysis. Results suggested that participants' evaluation of likeability was influenced by the condition (traditional, digital, or traditional + digital). In addition, there was no significant difference in how participants evaluated each of the student performance levels ('poor', 'average', and 'good') performing students. Post-hoc results revealed that participants rated the students in traditional + digital condition significantly higher than the digital condition, which indicates that participants did not like the students with the digital condition as much as the student with the traditional + digital condition. Flight instructors can be influenced in regards to the assessment of student likeability when they have no prior experience with the student and they are presented with an overview of their performance. This result might be a response to the resource provided versus the student's likeability and it is unclear as to which aspect this likeability measure from the participants is focused. The implication of this is such that any new introductions of digital visualizations or a portal by which the instructor can view student flight performance will need to be introduced in a way that the flight instructor's personal viewpoint of the data does not negatively impact the instructor's assessment and personal opinion of the student. Intentional flight instructor involvement in the ongoing development of future digital visualizations to be utilized for student assessment will be of the utmost importance in order to ensure desired levels of success. Previous research by Anderson, et al. (2011) emphasized this critical step in regards to expert assessments and user studies. They found that verbal feedback and assessments of user performance was a typical method of judgment for visualizations and the iterative feedback loop from end users was especially helpful during the continued development for proper visualization development.

5.1.5 Cognitive Load

The final research question was to determine the cognitive load of a flight instructor when determining the readiness for solo flight when using varying types of information for assessment. The measure is an average of participant responses of the cognitive load required to determine the student's landing competence, endorsement readiness, mentorability, and likeability. Results suggested that students' performance levels ('poor', 'average', and 'good' performing) influenced participants' cognitive load; however, the resource condition (traditional, digital, or traditional + digital) did not significantly influence participants' cognitive load. The results also revealed a significant interaction effect between the student performance level and condition. The post hoc results indicated that assessments of 'average' and 'good' performing students took significantly different amounts of cognitive load with the assessment of 'good' performing students taking more cognitive load than 'average' performing students. These findings, in conjunction with previous research (Keim et al., 2008) could attribute these results to the fact that instructors were able to easily dismiss 'poor' performing students and there may not have been enough information available to accurately assess the 'average' performing students. When assessments of the 'good' performing students were performed it took more effort on the part of the instructors to provide an assessment of their landing competence than the other measures. Results also suggested that participants reported experiencing significantly higher cognitive load for digital condition as compared to traditional + digital condition. The interaction effect of the cognitive load required appears to indicate that the digital condition resulted in mixed amounts of cognitive effort necessary as compared to the traditional and the traditional + digital conditions. For 'good' performing students, the participants had to expend a significantly higher amount of cognitive load when just the digital condition was utilized. As was cautioned by Paas et al. (2003) and Young et al. (2016), the findings from this study would appear to indicate that the digital condition, when presented alone, creates extraneous cognitive load during the assessments of 'good' performing students. The use of flight data that is retrievable from the airplane is relatively new in the general aviation training sector. Previous research (Anderson et al., 2011; Kosara et al., 2003; Van Wijk, 2005; Yau, 2013) identified the difficulty in determining the best choice of visualization for a particular data set and it is possible that the visualization utilized by this study was not the best choice for a data set representing landing performance. The participants of this study were asked to utilize a visualization of landing performance from digital aircraft data and there is no guidance from the Federal Aviation Administration as to how best to incorporate such visualizations with traditional logbooks. More importantly, there is no guidance as to how the data should be presented in a visualization that the flight instructors can utilize in order to assess student performance. Future digital visualizations will require careful development in order to limit the extraneous load and reduce the intrinsic load for student flight assessment.

5.3 Future Research

The accuracy of flight student assessments and the ability to utilize visualizations for student assessment are areas that provide significant opportunities for future research. The digital condition from this study could be modified so that it provides a usable visualization for flight instructors during their assessment of student performance. Once an appropriate visualization is developed, flight instructors will need guidance on how to interpret the visualization in conjunction with traditional logbook information. Future research should develop hypothetical student profiles with feedback from subject matter experts and include data visualizations that provide a complete and objectively accurate picture of students' flight performance rather than use a selection of traffic pattern performance to evaluate student abilities. Additionally, this study assessed student performance levels as 'good', 'average', and 'poor', but did not provide a definitive measure of student performance that flight instructors would most likely determine that they are prepared for solo flight. It is possible that even the hypothetical performance level representing 'good' student performance used for this study would not be ready for solo flight and the accuracy of the measurement of the participants' evaluations were not realistic for the information provided in the survey.

Additional research should use think-alouds with participants to better understand how they make use of the information and resources available for assessment when rating students on their flight readiness. Specifically, asking participants about the type of information being used to evaluate the student would provide insight into how role digital data played in their evaluation as well as inform future education of instructors on how to use digital data.

Research in cognitive load on actual instructors in their assessment of student performance needs continued assessment in order to determine the cognitive effort necessary under actual flight instruction circumstances. The task put before the survey participants in this study, while not uncommon, is one that was forced upon the participants. Given the high pressure task that flight instructors face, future research should focus on cognitive load in flight training. Assessing the cognitive load of flight instructors while they are performing assessments on actual students will provide a better picture of the reality of that effort. In this context the introduction of a digital visualization can be assessed against a traditional logbook and a clearer picture of the differences between the two resources available could be determined.

Another technology that has widespread implications for flight training and provides an area for research is the use of virtual reality (VR), augmented reality (AR), and mixed reality (MR) and is an area with tremendous research opportunities. Simulators have been used in flight training for many years, but the sole focus in simulation is to replicate the airplane with the level of fidelity necessary for the training task that is to be accomplished. The most advanced simulators cost in excess of $10 million, but even the lower fidelity devices approach costs in the $50,000 - $100,000 range. VR, AR, and MR focus on replicating the flying experience and environment and can be acquired for a small investment with commercial off the shelf (COTS) technology and some of the AR technologies can be accessed through apps on standard iOS devices. While the FAA has historically been reluctant to reduce the number of hours needed to be "ready" for a pilot certificate, these technologies, if used properly, have the potential of limiting the flight training footprint in excess of what the FAA requires as a minimum. The ideal mix of technologies and experiences is unknown at the present time but the development of small group trials is an ideal technique to determine what is helpful for teaching flight training concepts and skills. Blumenstyk (2016) argued that universities should be the main drivers of advancements of educational technologies rather than industry. As an example, students enrolled in the eight-week flight training toward a Private Pilot at Purdue University can use VR, AR, or MR in addition to initial solo flight, cross country solo flight, and final practical exam for certification. However, there is little research on how digital data from these tools can be utilized to assess student performance and final certification. I see potential of using data from tools to evaluate student performance objectively and even compare which of these technologies can be effective in supporting flight instructors to train and assess student pilots.

5.4 Implications for Certified Flight Instructors

Flight instruction is moving into an era of change and opportunity. The way that pilots have been trained and are currently training involves incorporating new technologies and pedagogies. At the start of this research project, there were companies that were taking data from an aircraft or using the technology of accelerometers and GPS position within personal cellphones to replicate a flight for review by flight instructors, students, or anyone that was willing to utilize the system that was developed. The issue with that technology was the length of time that it took to review the flight and to provide feedback which was almost the same amount of time it took to complete the flight. A two-hour flight would take almost two hours to review electronically and is a major deterrent for scalability, which is a limitation of deployment of visualizations in educational context (Baker, 2014). However, new technologies, like CloudAhoy (CloudAhoy Digital Flight Assessment Tool), that use artificial intelligence to segment a flight based upon the maneuvers in the Airmen Certification Standards (ACS) are readily accessible and can be used to discuss the flight immediately following a flight. Tools like CloudAhoy have the potential to provide flight instructors with the ability to assess their student's learning and to determine areas where they could focus on educational improvement. The benefits of visualizations in flight instruction could allow instructor to not only evaluate students, but also provide them with helpful feedback and influence their training similar to how Mazza and Dimitrova (2004) used CourseVis visualizations to influence online discussions. However, in order for visualization technologies to be valuable to flight instructors, they need to provide instructors information in a way that can easily be digested and used for the benefit of the student. For example, next iterations might be to start providing a score of a maneuver based upon the ACS standards, which provides an objective measure of student performance that can be used as a basis of evaluation and a determinant of further instruction. In addition, findings from this study suggest that the incorporation of visualizations will need to be completed in a manner that allows flight instructors to readily incorporate it into their assessment process or it could increase the cognitive load and influence accuracy of student assessment. This would involve educating flight instructors on how to make use of the digital visualizations. There are two opportunities that are readily available to accomplish this task: first, new modules could be embedded in CFI refresher courses to educate instructors; second, in the application of the standards that are required of a CFI during the practical exam.

Flight instructors are required to pass a practical exam, much like a private pilot, but the comprehensiveness of their knowledge is much more significant. The current standard for the practical exam is the Practical Test Standard (PTS) which was most recently published in June 2012. The FAA has been transitioning from the PTS to the Airmen Certification Standards (ACS) since 2014 when they published the first Private Pilot ACS. Since then, FAA has developed ACS guidance for the Instrument Rating, Commercial Pilot Certificate, and have recently released the ACS for an Airline Transport Pilot. The ACS is the ultimate standard for the expectation of pilot performance and all guidance documents and student preparation is based upon the ACS/PTS. The flight instructor ACS will be developed in the near future and this is an opportunity to incorporate a testing component for digital aircraft data assessment of student performance. There is a common phrase in the aviation industry that goes, “If you test it, they will teach it.” In order to incorporate best practices and encourage widespread use of new technologies, there needs to be a requirement to demonstrate the knowledge on a practical exam.

After their initial certification, flight instructors are also required to renew their instructor certificate every 24 calendar months. There are many ways to renew a flight instructor certificate, but the most common is to complete a Flight Instructor Refresher Clinic (FIRC), which can be done online or in person through an approved provider. FIRCs are standardized and approved by the FAA, which also provides an opportunity to incorporate modules on the use of digital data in the assessment process of student pilots. Demonstrations of assessment practices, demonstrations of improved accuracy of assessments, and demonstrations of reduction in flight time could show the opportunity and benefit of using digital data in flight training. FIRCs provide an ideal place to educate flight instructors to the opportunities of using digital data in student assessment.

5.5 Limitations

Flight training is highly regulated and is reliant, to some extent, upon the FAA's mandates and guidance for advancement. While an individual flight instructor may have the ability to affect their individual students, widespread change is difficult to achieve unless it is regulated from the government. That notwithstanding, advancements in aviation can be achieved through an advocate that is able to articulate the rationale and benefit for change. Within this study there were several limitations that could impact the future advancement of the use of digital data. A primary limitation was that given the lack of use of digital data currently in the field, the instructors could have disregarded the digital information completely and relied solely on the traditional logbook information. Another limitation is the forced nature of the survey questions. Flight instruction is an effort that results in an outcome of student assessment in an organic nature. As Thomas and Richards (2015) highlighted, the assessment of a flight student's performance and their readiness for the next phase of flight training is multi-dimensional and this study forced an instructor to assess student performance in a situation in which they may not have been able to make an accurate assessment. This dynamic could have impacted how flight instructor's responded and evaluated the student profiles. As discussed previously, the use of think-alouds might be a way to learn more about how flight instructors make use of digital data when evaluating student pilots.

REFERENCES

Ahlberg, C., Williamson, C., & Shneiderman, B. (1992, June). Dynamic queries for information exploration: An implementation and evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 619-626). ACM.

Anderson, E., Potter, K., Matzen, L., Shepherd, J., Preston, G., & Silva, C. (2011, June). A user study of visualization effectiveness using EEG and cognitive load. In *Computer Graphics Forum* (Vol. 30, No. 3, pp. 791-800). Blackwell Publishing Ltd.

Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of cognitive neuroscience*, *4*(3), 281-288.

Baker, R. S. J. D. (2010). Data Mining for Education. *International Encyclopedia of Education*, 7(3), 112-118.

Baker, R. S. (2014). Educational data mining: An advance for intelligent systems in education. *IEEE Intelligent systems*, *29*(3), 78-82.

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *US Department of Education, Office of Educational Technology*, 1-57.

Blumenstyk, G. (2016). As big-data companies enter the classroom, a pioneer issues a warning. *The Chronicle of Higher Education*, *62*(25), A28

Buja, A., McDonald, J. A., Michalak, J., & Stuetzle, W. (1991, October). Interactive data visualization using focusing and linking. In *Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on* (pp. 156-163). IEEE.

Carlson, R., Chandler, P., & Sweller, J. (2003). Learning and Understanding Science Instructional Material. *Journal of Educational Psychology, 95*(3), 629-40.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and instruction*, *8*(4), 293-332.

Crapo, A., Waisel, L., Wallace, W., Willemain, T., Stolfo, S., & Parsa, I. (2000). Visualization and the process of modeling: A cognitive-theoretic view. *Proceeding of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 218-226.

Dastani, M. (2002). The role of visual perception in data visualization. *Journal of Visual Languages & Computing*, *13*(6), 601-622.

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Journal of Educational Technology & Society*, 15(3), 58.

Federal Aviation Administration. (2018). Best Practices for Mentoring in Flight Instruction. Retrieved March, 2019, from https://www.faa.gov/training_testing/training/media/mentoring_best_practices.pdf

Federal Aviation Administration. (2017). Title 14, Chapter I, Subchapter D, Part 61. (14 CFR 61.87). Washington, DC: Electronic Code of Federal Regulations.

Federal Aviation Administration. (1997). Title 14, Part 121, et al., Revisions to Digital Flight Data Recorder Rules, Final Rule, 62(137). Washington, DC: Electronic Code of Federal Regulations.

Galyardt, A., & Goldin, I. (2015). Move your lamp post: Recent data reflects learner knowledge better than older data. *JEDM-Journal of Educational Data Mining*, *7*(2), 83-108.

Green, T., Ribarsky, W., & Fisher, B. (2009). Building and applying a human cognition model for visual analytics. *Information visualization*, *8*(1), 1-13.

Keim, D. D., Andrienko, G., Fekete, J., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual analytics: Definition, process, and challenges. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4950*, 154-175

Kosara, R., Hauser, H., & Gresh, D. (2003). An Interaction View on Information Visualization.

Larkin, J., & Simon, H. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, *11*(1), 65-100.

Liu, Z., Nersessian, N., & Stasko, J. (2008). Distributed cognition as a theoretical framework for information visualization. *IEEE transactions on visualization and computer graphics*, *14*(6).

Luan, J. (2002). Data mining and its applications in higher education. *New directions for institutional research*, 2002(113), 17-36.

Mazza, R., & Dimitrova, V. (2004, May). Visualising student tracking data to support instructors in web-based distance education. In *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters* (pp. 154-161). ACM.

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Miller, G., & Kintsch, W. (1994). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review, 101*(2), 343-352.

Morse, E., Lewis, M., & Olsen, K. A. (2000). Evaluating visualizations: using a taxonomic guide. *International Journal of Human-Computer Studies*, *53*(5), 637-662.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474-16479.

National Transportation Safety Board. (2017). Title 49, Subtitle B, Chapter VII, Part 830, Subpart C, 830.10. (49 CFR 830.10). Washington, DC: Electronic Code of Federal Regulations.

National Transportation Safety Board. (2017). Title 49, Subtitle B, Chapter VIII, Part 831, 831.12. (49 CFR 831.12). Washington, DC: Electronic Code of Federal Regulations.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, *38*(1), 1-4.

Rönkä, M., Tolvanen, H., Lehikoinen, E., von Numers, M., & Rautkari, M. (2008). Breeding habitat preferences of 15 bird species on south-western Finnish archipelago coast: applicability of digital spatial data archives to habitat assessment. *Biological Conservation*, *141*(2), 402-416.

Shelton, A., Smith, A., Wiebe, E., Behrle, C., Sirkin, R., & Lester, J. (2016). Drawing and writing in digital science notebooks: Sources of formative assessment data. *Journal of Science Education and Technology*, *25*(3), 474-488.

Shneiderman, B. (2014). The big picture for big data: Visualization. *Science (New York, N.Y.), 343*(6172), 730.

Stover, K., Yearta, L., & Harris, C. (2016). Formative Assessment in the Digital Age. *The Reading Teacher, 69*(4), 377-381.

Sweller, J., & Chandler, P. (1994). Why Some Material Is Difficult to Learn. *Cognition and Instruction, 12*(3), 185-233.

Sweller, J. & Paas, F. (2017). Should self-regulated learning be integrated with cognitive load theory? A commentary. *Learning and Instruction,* Learning and Instruction.

Thomas, M., & Richards, C. (2015). Determining Readiness for Solo Flight Training. *Aviation Psychology and Applied Human Factors*.

Tufte, E. R. (2006). Beautiful evidence. *New York*.

Tufte. E. R. (2001). The visual display of quantitative information.

University Aviation Association. (2019). Retrieved August, 2018, from http://www.uaa.aero/

Van Merrienboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational psychology review, 17*(2), 147-177.

Van Wijk, J. (2005). The value of visualization. *Visualization, 2005. VIS 05. IEEE,* 79-86.

Yau, N. (2013). *Data points: Visualization that means something*. John Wiley & Sons.

Young, J., Ten Cate, O., O'Sullivan, P., & Irby, D. (2016). Unpacking the Complexity of Patient Handoffs Through the Lens of Cognitive Load Theory. *Teaching and Learning in Medicine, 28*(1), 88-96.

APPENDIX

SURVEYS

SURVEY INTRODUCTION

Welcome to the research study!

This survey is part of a PhD Dissertation evaluating how Flight Instructors make decisions of student performance for initial solo. You are under no obligation to participate in this survey. If you do participate you may stop at any time and for any reason. Your answers on the survey are anonymous and cannot be used in any way for identification.

This survey should take approximately 15 minutes to complete, and consists of 21 questions. If you have any questions regarding this survey, please contact;
Brian G. Dillman at (765) 409-4501 or dillman@purdue.edu
Ala Samarapungavan at 765-494-7321 or ala@purdue.edu.

This research study has been reviewed by the Institutional Review Board at Purdue University. IRB Reference Number 1712019989.

Answers to the first six questions of the survey will provide us with demographic information and information concerning your instructional experience. The last sixteen questions ask for evaluation of a student's readiness for initial solo.

By clicking on the continue button you agree to the following statements:
I am at least 18 years old
I have a current Flight Instructor certificate from the Federal Aviation Administration
I give my consent to voluntarily participate in this study
I have read and understood the above information

Thank You for your participation!

- I consent, begin the study

- ○ I do not consent, I do not wish to participate

How long have you possessed a Flight Instructor Certificate?

(Select maximum amount if more than 50 years)

	0	10	20	30	40	50
Years						

Approximately how many total hours do you currently have?

(Select maximum amount if more than 5000 hours)

	0	1000	2000	3000	4000	5000
Total Hours						

Approximately how many total instructional hours do you currently have?

(Select maximum amount if more than 5000 hours)

	0	1000	2000	3000	4000	5000
Instructional Hours						

Approximately how many students have you endorsed for initial solo?

(Select maximum amount if more than 100 students)

	0	20	40	60	80	100
Signoffs for Initial Solo						

I find working with computer technology and the associated information displays to be very easy.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Per the US Census Bureau Regions and Divisions Graphic, select the Region where you predominately conduct flight training.

- West
- Midwest
- Northeast
- South
- Pacific

CONTROL GROUP SURVEY QUESTIONS

This first series of questions is the Control and represents the traditional type of information that a Flight Instructor has available when evaluating a flight student with whom they have never flown with before. In the survey there were three levels of student performance with which a participant in the study was presented (‘poor’, ‘average’, and ‘good’). The level in this series of questions represents ‘poor’ performance. Each participant in the survey study only viewed one level of performance and one type of information.

Jacob Smith has been working on his first solo. Jacob started his flight training at the beginning of the summer and has been flying 2 or 3 times per week for the last several weeks. His instructor, Steve, acquired the hours necessary to get a job with the Regional Airlines and left for his starting class date at the beginning of last week. The Chief Flight Instructor at Jacob’s school contacted you and asked that you continue working with him toward his goal of becoming a Private Pilot. The Chief Pilot did mention that Jacob has been paying for his flight lessons out of his savings that he had acquired working a part time job. The Chief CFI asked that you not duplicate any flights that weren’t necessary for fear that Jacob would run out of money prior to him completing his flight training. Look over the records that Jacob’s previous instructor left and answer the questions concerning the student's readiness for initial solo.

Logbook Record

Flight	Dual Given	Number of Landings	Comments
Landing Lesson 1	1.3	8	Good attempt at first landings. Keep aileron correction in to the wind and increase deflection as you get slower as required especially upon touchdown.
Landing Lesson 2	1.5	8	Landing sequence is good, but still need work on airspeed control on final approach and flaring. Generally flaring too late and flat.
Landing Lesson 3	1.4	6	Improved but still landing firm occasionally. As aircraft slows over the runway, need to add increasing back pressure to prevent it from sinking into the runway with the nose dropping.
Landing Lesson 4	1.4	7	Flare was high. Make sure your eyes are at the end of the runway to see if flattening out. Runway width may cause visual illusion.
Landing Lesson 5	1.6	9	Flare a little aggressive, landing a little flat/sideways, remember xwind correction
Landing Lesson 6	1.1	6	Landing flare improving, but speed on final wasn't stable. Small corrections for approach & crosswind. Don't stop flare when adding crosswind correction. After touchdown continue adding aileron while stopping to correct for wind and keep wings level.

CONTROL GROUP ASSESSMENT OF LANDING COMPETENCE

Did the applicant strike you as competent for this flying task?

- Extremely competent
- Moderately competent
- Slightly competent
- Neither competent nor incompetent
- Slightly incompetent
- Moderately incompetent
- Extremely incompetent

How likely is it that the applicant has the necessary skills for this flying task?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How qualified do you think the applicant is to handle this flying task?

- Extremely qualified
- Moderately qualified
- Slightly qualified
- Neither qualified or unqualified
- Slightly unqualified
- Moderately unqualified
- Extremely unqualified

How much effort was devoted to determining your responses concerning student competence?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

CONTROL GROUP ASSESSMENT OF ENDORSMENT READINESS

How likely would you be to endorse the student for initial solo flight operations?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How likely would you be to endorse the student (once they've completed their training) to conduct a private pilot certification flight test with a designated examiner?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How likely do you think it is that the flight student will successfully pass the private pilot certification flight test?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How much effort was devoted to determining your responses concerning student readiness for an endorsement?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

CONTROL GROUP ASSESSMENT OF MENTORABILITY

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Encourage the student to stay in the field if he was considering changing majors?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Encourage the student to continue to focus on flight training if he was considering switching focus to aviation administration, Unmanned aerial systems, or maintenance technician training?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Give the student extra help if he was having trouble mastering a difficult flying task and their instructor is having difficulty providing guidance for improvement?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How much effort was devoted to determining your responses concerning any possible mentoring interaction?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

CONTROL GROUP ASSESSMENT OF LIKEABILITY

Based upon the information presented, how much did you like the flight student?

- Like a great deal
- Like a moderate amount
- Like a little
- Neither like nor dislike
- Dislike a little
- Dislike a moderate amount
- Dislike a great deal

Would you characterize the flight student as someone you want to get to know better and possibly work with during future flight training?

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Disagree
- Strongly disagree

Would the flight student fit well within a crew environment in a multi-crew cockpit?

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Disagree
- Strongly disagree

How much effort was devoted to determining your responses concerning the likability of the student?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

EXPERIMENTAL GROUP 1 SURVEY QUESTIONS

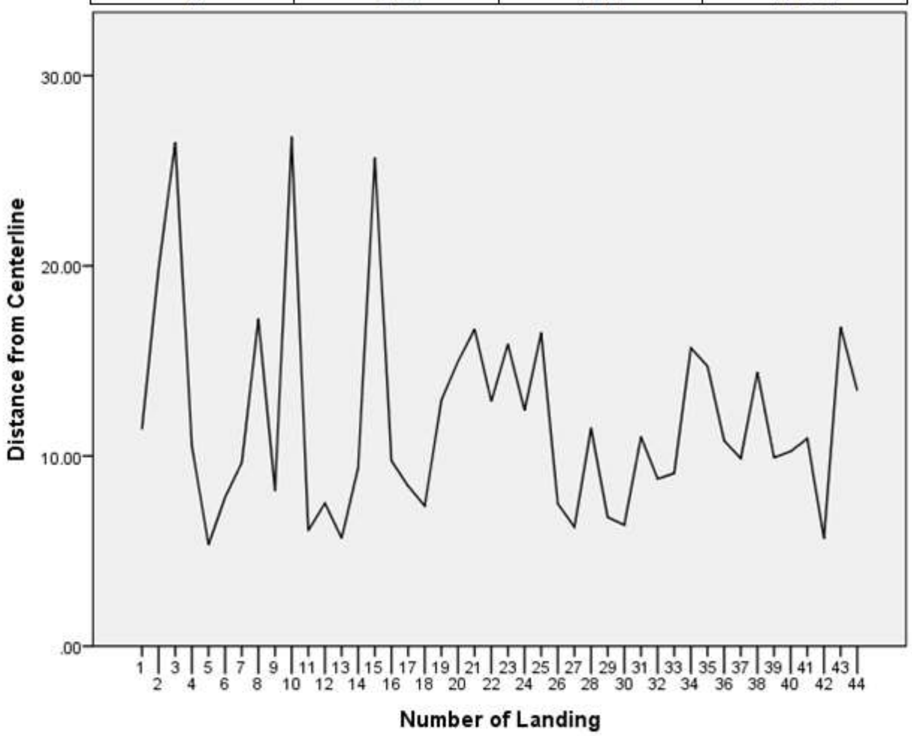
This second series of questions is the Experimental group 1. Like the Control group, there were three levels of student performance with which a participant in the study was presented (‘poor’, ‘average’, and ‘good’). The level in this series of questions represents ‘poor’ performance. Each participant in the survey study will only view one level of performance.

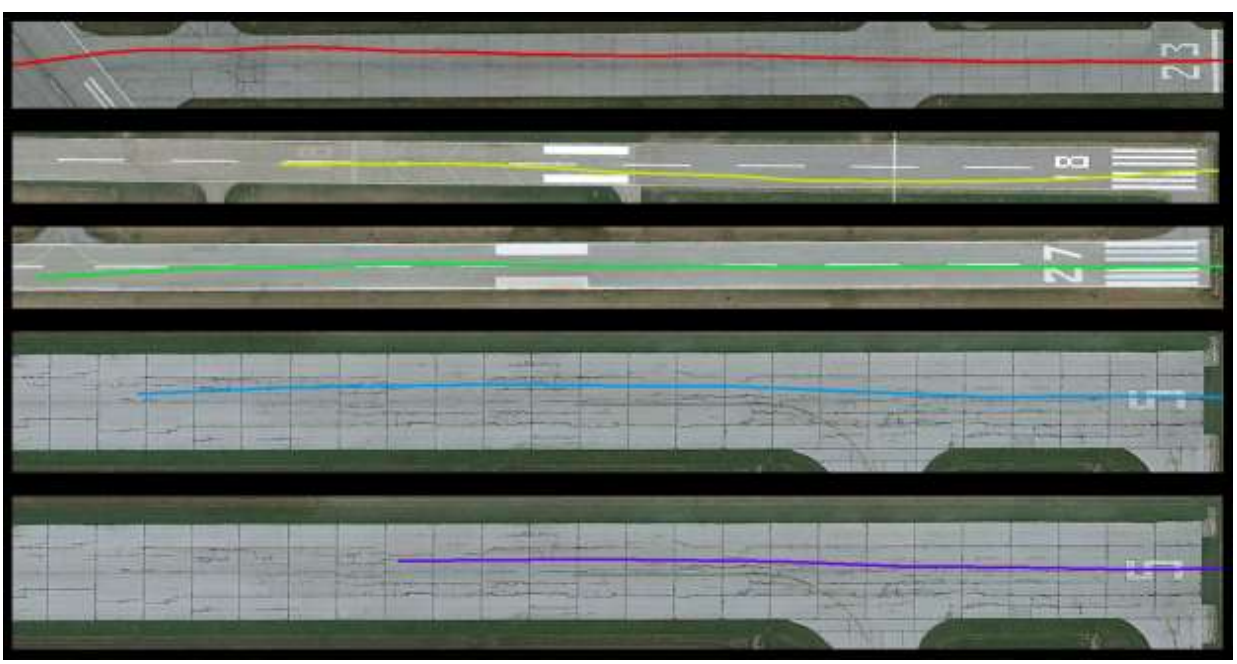
Jacob Smith has been working on his first solo. Jacob started his flight training at the beginning of the summer and has been flying 2 or 3 times per week for the last several weeks. His instructor, Steve, acquired the hours necessary to get a job with the Regional Airlines and left for his starting class date at the beginning of last week. The Chief Flight Instructor at Jacob’s school contacted you and asked that you continue working with him toward his goal of becoming a Private Pilot. The Chief Pilot did mention that Jacob has been paying for his flight lessons out of his savings that he had acquired working a part time job. The Chief CFI asked that you not duplicate any flights that weren’t necessary for fear that Jacob would run out of money prior to him completing his flight training. Look over the records that Jacob’s previous instructor left and answer the questions concerning the student's readiness for initial solo.

Logbook Record

Flight	Dual Given	Number of Landings	Comments
Landing Lesson 1	1.3	8	Good attempt at first landings. Keep aileron correction in to the wind and increase deflection as you get slower as required especially upon touchdown.
Landing Lesson 2	1.5	8	Landing sequence is good, but still need work on airspeed control on final approach and flaring. Generally flaring too late and flat.
Landing Lesson 3	1.4	6	Improved but still landing firm occasionally. As aircraft slows over the runway, need to add increasing back pressure to prevent it from sinking into the runway with the nose dropping.
Landing Lesson 4	1.4	7	Flare was high. Make sure your eyes are at the end of the runway to see if flattening out. Runway width may cause visual illusion.
Landing Lesson 5	1.6	9	Flare a little aggressive, landing a little flat/sideways, remember xwind correction
Landing Lesson 6	1.1	6	Landing flare improving, but speed on final wasn't stable. Small corrections for approach & crosswind. Don't stop flare when adding crosswind correction. After touchdown continue adding aileron while stopping to correct for wind and keep wings level.

Number of Landings	Minimum Centerline Deviation	Maximum Centerline Deviation	Average Centerline Deviation
44	5.33	26.78	11.9336





EXPERIMENTAL GROUP 1 ASSESSMENT OF LANDING COMPETENCE

Did the applicant strike you as competent for this flying task?

- Extremely competent
- Moderately competent
- Slightly competent
- Neither competent nor incompetent
- Slightly incompetent
- Moderately incompetent
- Extremely incompetent

How likely is it that the applicant has the necessary skills for this flying task?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How qualified do you think the applicant is to handle this flying task?

- Extremely qualified
- Moderately qualified
- Slightly qualified
- Neither qualified or unqualified
- Slightly unqualified
- Moderately unqualified
- Extremely unqualified

How much effort was devoted to determining your responses concerning student competence?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

EXPERIMENTAL GROUP 1 ASSESSMENT OF ENDORSEMENT READINESS

How likely would you be to endorse the student for initial solo flight operations?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How likely would you be to endorse the student (once they've completed their training) to conduct a private pilot certification flight test with a designated examiner?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How likely do you think it is that the flight student will successfully pass the private pilot certification flight test?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How much effort was devoted to determining your responses concerning student readiness for an endorsement?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

EXPERIMENTAL GROUP 1 ASSESSMENT OF MENTORABILITY

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Encourage the student to stay in the field if he was considering changing majors?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Encourage the student to continue to focus on flight training if he was considering switching focus to aviation administration, Unmanned aerial systems, or maintenance technician training?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Give the student extra help if he was having trouble mastering a difficult flying task and their instructor is having difficulty providing guidance for improvement?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How much effort was devoted to determining your responses concerning any possible mentoring interaction?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

EXPERIMENTAL GROUP 1 ASSESSMENT OF LIKEABILITY

Based upon the information presented, how much did you like the flight student?

- Like a great deal
- Like a moderate amount
- Like a little
- Neither like nor dislike
- Dislike a little
- Dislike a moderate amount
- Dislike a great deal

Would you characterize the flight student as someone you want to get to know better and possibly work with during future flight training?

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Disagree
- Strongly disagree

Would the flight student fit well within a crew environment in a multi-crew cockpit?

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Disagree
- Strongly disagree

How much effort was devoted to determining your responses concerning the likability of the student?

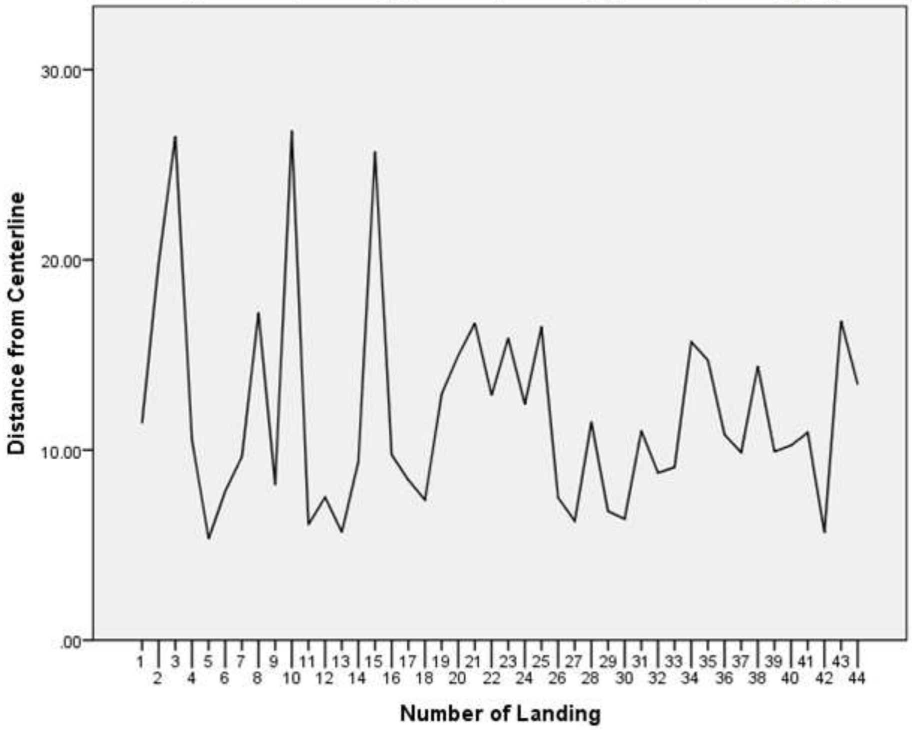
- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

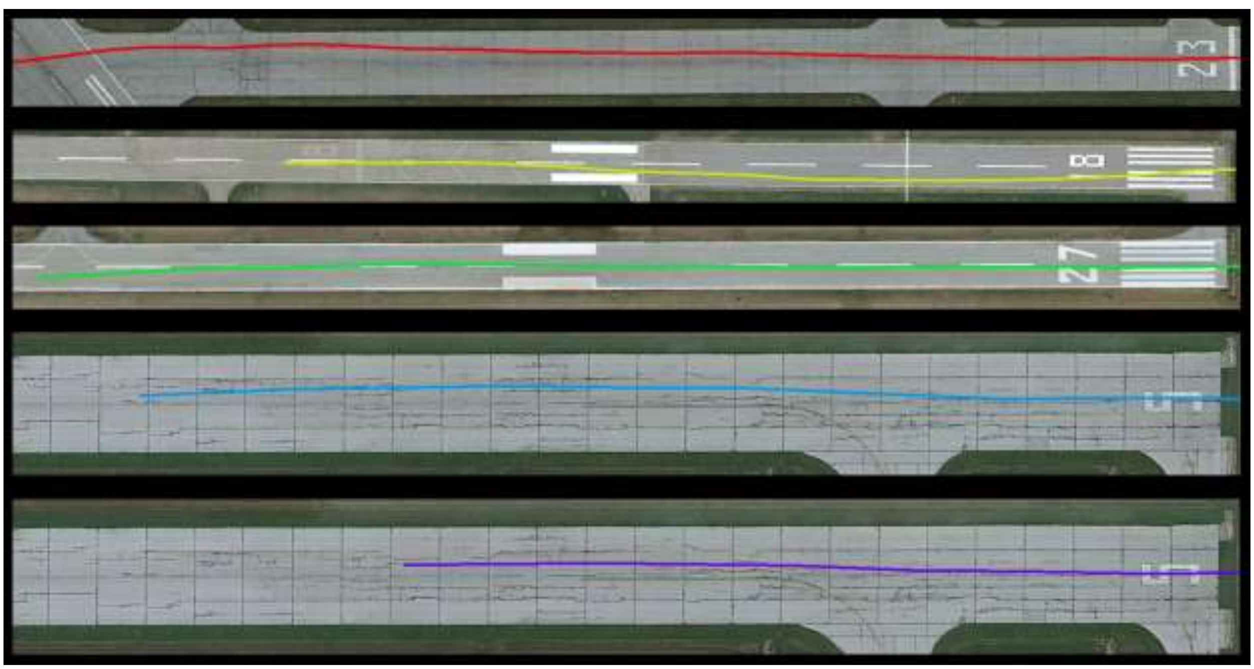
EXPERIMENTAL GROUP 2 SURVEY QUESTIONS

This third series of questions is the Experimental Group 2. Like the Control and Experimental Group 1, there were three levels of student performance with which a participant in the study was presented (‘poor’, ‘average’, and ‘good’). The level in this series of questions represents ‘poor’ performance. Each participant in the survey study will only view one level of performance.

Jacob Smith has been working on his first solo. Jacob started his flight training at the beginning of the summer and has been flying 2 or 3 times per week for the last several weeks. His instructor, Steve, acquired the hours necessary to get a job with the Regional Airlines and left for his starting class date at the beginning of last week. The Chief Flight Instructor at Jacob’s school contacted you and asked that you continue working with him toward his goal of becoming a Private Pilot. The Chief Pilot did mention that Jacob has been paying for his flight lessons out of his savings that he had acquired working a part time job. The Chief CFI asked that you not duplicate any flights that weren’t necessary for fear that Jacob would run out of money prior to him completing his flight training. Look over the records that Jacob’s previous instructor left and answer the questions concerning the student's readiness for initial solo.

Number of Landings	Minimum Centerline Deviation	Maximum Centerline Deviation	Average Centerline Deviation
44	5.33	26.78	11.9336





EXPERIMENTAL GROUP 2 ASSESSMENT OF LANDING COMPETENCE

Did the applicant strike you as competent for this flying task?

- Extremely competent
- Moderately competent
- Slightly competent
- Neither competent nor incompetent
- Slightly incompetent
- Moderately incompetent
- Extremely incompetent

How likely is it that the applicant has the necessary skills for this flying task?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How qualified do you think the applicant is to handle this flying task?

- Extremely qualified
- Moderately qualified
- Slightly qualified
- Neither qualified or unqualified
- Slightly unqualified
- Moderately unqualified
- Extremely unqualified

How much effort was devoted to determining your responses concerning student competence?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

EXPERIMENTAL GROUP 2 ASSESSMENT OF ENDORSEMENT READINESS

How likely would you be to endorse the student for initial solo flight operations?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How likely would you be to endorse the student (once they've completed their training) to conduct a private pilot certification flight test with a designated examiner?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How likely do you think it is that the flight student will successfully pass the private pilot certification flight test?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How much effort was devoted to determining your responses concerning student readiness for an endorsement?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

EXPERIMENTAL GROUP 2 ASSESSMENT OF MENTORABILITY

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Encourage the student to stay in the field if he was considering changing majors?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Encourage the student to continue to focus on flight training if he was considering switching focus to aviation administration, Unmanned aerial systems, or maintenance technician training?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

<u>If you encountered this student on the flight line, how likely would you be to. . .</u>

Give the student extra help if he was having trouble mastering a difficult flying task and their instructor is having difficulty providing guidance for improvement?

- Extremely likely
- Moderately likely
- Slightly likely
- Neither likely nor unlikely
- Slightly unlikely
- Moderately unlikely
- Extremely unlikely

How much effort was devoted to determining your responses concerning any possible mentoring interaction?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

EXPERIMENTAL GROUP 2 ASSESSMENT OF LIKEABILITY

Based upon the information presented, how much did you like the flight student?

- Like a great deal
- Like a moderate amount
- Like a little
- Neither like nor dislike
- Dislike a little
- Dislike a moderate amount
- Dislike a great deal

Would you characterize the flight student as someone you want to get to know better and possibly work with during future flight training?

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Disagree
- Strongly disagree

Would the flight student fit well within a crew environment in a multi-crew cockpit?

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Disagree
- Strongly disagree

How much effort was devoted to determining your responses concerning the likability of the student?

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither low nor high mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

VITA

Brian G. Dillman

dillman@purdue.edu

Employment History

Purdue University, West Lafayette, IN July 2007 - Present

Associate Professor

- Research in flight certification, data analysis, and safety
- Flight education at all levels of pilot competencies
- Establishing safety and positive flight control expectations

Federal Aviation Administration, Indianapolis, IN July 1998 - Present

FAA Safety Team Representative

- Promoting and fostering aviation safety
- Counseling pilots following potentially unsafe acts
- Assisting FAA in necessary capacities

Federal Aviation Administration, Indianapolis, IN December 2006 - Present

Designated Pilot Examiner

- Private Pilot
- Commercial Pilot
- Instrument Rating
- Multi-Engine Rating
- Flight Instructor - Airplane
- Instrument Instructor
- Multi Engine Instructor
- Sport Pilot
- Remote Pilot

Education

Purdue University, School of Aviation Technology, West Lafayette, IN

- PhD of Educational Psychology Expected Spring 2019
- Master's of Education in Educational Technology August 2000
- Bachelor of Science with distinction in Professional Pilot Technology May 1995

Flight Time

Total Time..............5025	Multi-engine..............2120	Dual Given………...3137
Pilot In Command.. 4415	Instruments..................355	Aerobatics...................450

Certificates and Ratings

- Airline Transport Pilot
- Certified Flight Instructor – Single and Multi-engine Instrument Airplane
- Advanced Ground Instructor

Professional Memberships

- University Aviation Association
  - Safety Committee Member
  - Flight Education Committee Member
- Aircraft Owner & Pilots Association
- Experimental Aircraft Association

Refereed journal articles

Keller*, J., Borsa, R., Duran, L., & Dillman, B. (In Process). Effects of runway midpoint markings and general aviation pilot decision–making

Bromfield*, M., & Dillman, B. G. (2015). The effects of using an angle of attack system on pilot performance and workload during selected phases of flight. *Procedia Manufacturing, 3*, 3222–3229.

Leib*, S.M., Dillman, B.G., Petrin, D.A., Young, J.P. (2012). A Comparison of the Effect of Variations to U.S. Airport Terminal Signage on the Successful Wayfinding of Chinese and American Cultural Groups. The Journal of Aviation Technology and Engineering.

Ropp*, T.D. & Dillman, B.G., (2010). Standardized Measures of Safety: Finding Global Common Ground for Safety Metrics. Technology Interface Journal. Spring 2010 issue. V.10 No. 3. ISSN#1523-9926.

Dillman*, B.G., Hendricks, R.E., Petrelli, M.A., Elliott, S.J. (2010). Biometric Access to Training Devices as a Security Protocol in Flight Training. Journal of Air Transport Studies. Hellenic Aviation Society, Athens, Greece. 1, (1), 82 – 97.

<u>Government, university, industrial reports and standards</u>

Dillman, B. G. Pruchnicki, S. A., & Wilt, D., (In Review). Angle of Attack for General Aviation Operations (FAA Technical Report). Washington, DC: US Department of Transportation.

Fanjoy, R. O., Winter, S. R., Dillman, B. G., Young, S. B., Johnson, M. E., Rosser, T. G., Pruchnicki, S. A., Rao, A., Podhipak, N., & Maharaj, I., (In Review). Acquisition Distance of LED–based Runway Closure Lighting (FAA Technical Report). Washington, DC: US Department of Transportation.

Hawkins, G. H., & Dillman, B. G., (In Process). Evaluation of Midpoint Runway Marking (FAA Technical Report). Washington, DC: US Department of Transportation.

Dillman, B.G. & Ropp, T. D., (2008). Purdue Safety Gap Analysis (University Report). West Lafayette, IN. Purdue University, Department of Aviation Technology.

<u>Refereed conference proceedings (with presentation)</u>

Dillman, B. G., Wilt, D., Pruchnicki, S., <u>Rudari, L</u>., <u>Ball, M</u>., & <u>Pomeroy, M</u>. (2015). *Flight Operational Quality Assurance (FOQA) – Do Exceedances Tell the Whole Story?* Dayton, OH: 18th International Symposium of Aviation Psychology.

Ferguson, D.M., Dubikovsky, S. I., Dillman, B.G. (2014). Mindset Changes in an Engineering Technology Design Course. Joint International Conference on Engineering & International Conference on Information Technology. Riga, Latvia

Dillman, B.G., Mott, J.H. (2013). Proceedings of the 12th Annual Oklahoma Aerospace Education Research Symposium. Development and Implementation of a Capstone Experience in Professional Flight Technology. Stillwater, OK.

Borsa, R., Dillman, B.G., Ropp, T.D. (2013). .Proceedings of 17th International Symposium on Aviation Psychology. Exploring the Impact of ADS-B NextGen Technology Requirements for General Aviation. Dayton, OH.

Conference Presentations

Dillman, B.G., Davis, M.D., Esser, D.A., Higgins, J., McCarroll, L.D., September 2011. Recording and Monitoring Flight Data in Primary Training Aircraft. Presented to the University Aviation Association, Education Session, Indianapolis, IN.

Ruiz, J.R., Robertson, M.F., Voges, J.K., Dillman, B.G., Lu, C. & Mott, J. October 2010, Safety Management Systems at Collegiate Aviation Programs: Possible Collaborative Synergy and Projects. Presented to the University Aviation Association, Education Session II, St. Paul, MN.