An Alternative Method of Identification of a Failed Engine in Twin-Engine Propeller Aircraft

Andrey Babin

Follow this and additional works at: https://commons.erau.edu/edt

Part of the Aerospace Engineering Commons

Scholarly Commons Citation

https://commons.erau.edu/edt/398
AN ALTERNATIVE METHOD OF IDENTIFICATION OF A FAILED ENGINE IN TWIN-ENGINE PROPELLER AIRCRAFT

by

Andrey Babin

A Thesis Submitted to the College of Aviation, School of Graduate Studies, in Partial Fulfillment of the Requirements for the Degree of Master of Science in Aeronautics

Embry-Riddle Aeronautical University
Daytona Beach, Florida
May 2018
AN ALTERNATIVE METHOD OF IDENTIFICATION OF A FAILED ENGINE IN TWIN-ENGINE PROPELLER AIRCRAFT

by

Andrey Babin

This Thesis was prepared under the direction of the candidate’s Thesis Committee Chair, Dr. Andrew R. Dattel, Assistant Professor, Daytona Beach Campus, and Thesis Committee Members Dr. Margaret F. Klemm, Associate Professor, Daytona Beach Campus, and Dr. Cass D. Howell, Adjunct Faculty, Daytona Beach Campus, and has been approved by the Thesis Committee. It was submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science in Aeronautics

Thesis Committee:

[Signatures]

Andrew R. Dattel, Ph.D.
Committee Chair

Margaret F. Klemm, Ph.D.
Committee Member

Cass D. Howell, Ed.D.
Committee Member

Donald S. Metscher, D.B.A.
Program Coordinator
Master of Science in Aeronautics

Antonio Cortés, Ph.D.
Associate Dean
School of Graduate Studies,
College of Aviation

Christopher Grant, Ph.D.
Associate Vice President of Academics

[Date]
Acknowledgements

This thesis was a great life and career achievement to me, and it wouldn’t be possible without those who were around me. I am grateful to my family for their help and ideas. They encouraged my academic advancements and always helped me get on the right track and pursue education in areas that are of interest to me. A big thank you goes to my girlfriend, Jiyeon Song, for her great love and care. She was always by my side when I needed help, and the support she provided is an enormous contribution to this work.

I would like to thank my committee chair, Dr. Andrew R. Dattel, for his understanding and guidance on preparation and conduct of this research. I am thankful to everyone who took part in this project, and, in particular, to Dr. Margaret F. Klemm and Dr. Cass D. Howell for their advices and recommendations, and to Dr. Margaret F. Klemm for her help with the video script. I am also thankful to Joey Maxwell for creating the voiceover of the video script, to Ziyi Dong for her help with video editing, to Steve Wilson (whom I have contacted through X-Plane.Org forum) for his help with programming, to Stephanie Fussel for her help with piloting the study, and to the personnel of the safety department of one U.S. airline (whose names I will not include for confidentiality purposes) for their help with the survey data collection.

I would like to express special gratitude to the College of Aviation of Embry-Riddle Aeronautical University, and Dr. Alan Stolzer and Dr. Antonio Cortes in particular, for their help with the funding that was provided to me to conduct the research.

Lastly, I would like to thank everyone else who contributed to this thesis and everyone who, by their actions or otherwise, encouraged and motivated me to further my career in aviation.
Abstract

Researcher: Andrey Babin

Title: AN ALTERNATIVE METHOD OF IDENTIFICATION OF A FAILED ENGINE IN TWIN-ENGINE PROPELLER AIRCRAFT

Institution: Embry-Riddle Aeronautical University

Degree: Master of Science in Aeronautics

Year: 2018

Previous research has revealed that wrong identification of a failed engine during flight is not an uncommon event in an aircraft cockpit. A number of fatal accidents in the past, including the recent TransAsia Flight 235 accident, resulted from engine misidentification. Most accidents of this type happened on takeoff when pilot workload was at its highest level. This thesis consists of two research studies. For Research Study 1, a survey was created to gather opinions of airline pilots who operate twin-engine turboprop aircraft. Twenty-nine percent of respondents to the survey agreed with the statement that there could be a better method of identification of a failed engine. Thirty-four percent of respondents who provided suggestions for improvement of a current method recommended adding a visual indicator of some kind.

Research Study 2 was designed on the assumption that the current method of identification of a failed engine, called “dead leg – dead engine,” was not efficient enough, and an alternative method was introduced and tested. The alternative method was based on a visual sensory channel and it involved the use of a failed engine indicator called Engine Status Panel.
**Method:** To test the proposed training method, 50 student pilots from Embry-Riddle Aeronautical University who had not obtained multi-engine rating (MEL) were sampled and assigned to two groups – either the Traditional or the Alternative method. Participants performed three takeoffs in a flight training device, and an engine failure was simulated during each takeoff. Participant accuracy of identification and response time to an engine failure were measured and compared across takeoffs and between groups. **Results:** Participants in the Alternative Group were able to identify a failed engine significantly faster than the participants in the Traditional Group. Additionally, Participants in the Alternative Group reported being generally less confused in regard to which engine was failing and more confident that their identification was correct. For further development of the matter, it is recommended to measure the effectiveness of the alternative method among pilots who are rated for multi-engine aircraft operations. Implementation of a visual indicator similar or identical to the Engine Status Panel in twin-engine general aviation aircraft may improve pilot response to identification of a failed engine.
Table of Contents

Thesis Review Committee .......................................................... ii
Acknowledgements ..................................................................... iii
Abstract .................................................................................. iv
List of Tables ............................................................................. ix
List of Figures .............................................................................. x

Chapter

I  Introduction................................................................................. 1
    Significance of the Study ....................................................... 2
    Statement of the Problem ...................................................... 3
    Purpose Statement ............................................................... 4
    Hypotheses ........................................................................... 5
    Delimitations ........................................................................ 5
    Limitations and Assumptions ................................................. 5
    Definition of Terms .............................................................. 6
    List of Acronyms ................................................................. 6

II  Literature Review................................................................. 8
    Engine Misidentification ..................................................... 8
    Error Types .......................................................................... 10
    Human Capabilities and Limitations ..................................... 11
       Working Memory .............................................................. 11
       Stress .............................................................................. 12
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>34</td>
</tr>
<tr>
<td>Accuracy of Identification of a Failed Engine</td>
<td>36</td>
</tr>
<tr>
<td>Response Time to an Engine Failure</td>
<td>37</td>
</tr>
<tr>
<td>Hypothesis Testing</td>
<td>38</td>
</tr>
<tr>
<td>Qualitative Data</td>
<td>39</td>
</tr>
<tr>
<td>Reliability Testing</td>
<td>42</td>
</tr>
<tr>
<td>V Discussion</td>
<td>44</td>
</tr>
<tr>
<td>Limitations</td>
<td>48</td>
</tr>
<tr>
<td>Recommendations</td>
<td>49</td>
</tr>
<tr>
<td>Further Research</td>
<td>50</td>
</tr>
<tr>
<td>Conclusion</td>
<td>53</td>
</tr>
<tr>
<td>References</td>
<td>54</td>
</tr>
<tr>
<td>Appendices</td>
<td></td>
</tr>
<tr>
<td>A Survey for Twin-Engine Turboprop Pilots</td>
<td>59</td>
</tr>
<tr>
<td>B Participant Demographics Questionnaire</td>
<td>61</td>
</tr>
<tr>
<td>C Participant Post-Flight Questionnaire</td>
<td>63</td>
</tr>
<tr>
<td>D Training Video Script</td>
<td>65</td>
</tr>
<tr>
<td>E Training Video Quiz</td>
<td>70</td>
</tr>
<tr>
<td>F Participant Instructions</td>
<td>72</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pros and Cons of the Current Method of Identification of a Failed Engine in Twin-Engine Turboprop Aircraft</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>Response Times for the Traditional Group</td>
<td>37</td>
</tr>
<tr>
<td>3</td>
<td>Response Times for the Alternative Group</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>Group x Flight Interaction with Mean Response Times to an Engine Failure</td>
<td>39</td>
</tr>
<tr>
<td>5</td>
<td>Participant Confidence in Accuracy of Failed Engine Identification</td>
<td>40</td>
</tr>
<tr>
<td>6</td>
<td>Participant Confidence in Adequacy of Time Spent on Identification of a Failed Engine</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>Method Used as Reported by Participants</td>
<td>41</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Classification of Errors</td>
<td>11</td>
</tr>
<tr>
<td>2  Flight Training Device Setup</td>
<td>23</td>
</tr>
<tr>
<td>3  Engine Status Panel</td>
<td>25</td>
</tr>
<tr>
<td>4  Flight Experience in Turboprop Aircraft</td>
<td>31</td>
</tr>
<tr>
<td>5  Flight Experience in all Types of Multi-Engine Aircraft</td>
<td>32</td>
</tr>
<tr>
<td>6  Categorized Suggestions for Improvements to the Current Method of</td>
<td>34</td>
</tr>
<tr>
<td>Identification of a Failed Engine</td>
<td></td>
</tr>
<tr>
<td>7  Participant Age (All Groups)</td>
<td>35</td>
</tr>
<tr>
<td>8  Participant Flight Experience (All Groups)</td>
<td>36</td>
</tr>
<tr>
<td>9  Suggestions Proposed by the Participants After the Three Test</td>
<td>42</td>
</tr>
<tr>
<td>Flights Were Completed</td>
<td></td>
</tr>
</tbody>
</table>
Chapter I

Introduction

In 2015, TransAsia ATR 72-600, a twin-engine turboprop passenger aircraft, crashed shortly after takeoff, killing 43 out of 53 people on board. The investigation team determined that the crew had received an engine flameout warning, but the captain did not properly identify the affected engine and instead shut down the wrong, perfectly working, engine. Having no power at only 1,500 feet, the crew could not restart either of the two engines, stalled, and crashed into a river (Aviation Safety Council, 2016).

The TransAsia crash is not the first of its kind – similar aircraft accidents have happened in the past. For example, in 1992, a DHC-6 Twin Otter crashed in California, killing 16 out of 22 occupants, and in 2009, an engine misidentification resulted in a crash of SA Airlink Jetstream 41 near Durban Airport, South Africa (National Transportation Safety Board, n.d.; South African Civil Aviation Authority, n.d.). These two accidents are notably similar to the TransAsia crash. As such, engine failure in all three cases happened on takeoff and each accident involved a twin-engine turboprop aircraft.

An engine failure in a propeller aircraft is more dangerous to a pilot than a failure in an aircraft powered by a jet engine. A jet engine generates thrust by air expansion, but a propeller engine generates thrust by propeller blades which generate lift and move the airplane forward. In case of an engine failure, the propeller starts creating drag that will reduce airspeed. In twin-engine propeller airplanes, this drag is also followed by imbalance in thrust and a significant yaw. An engine failure in such aircraft results in a loss of up to 80% of climb performance (Federal Aviation Administration, 2016).
Significance of the Study

A performance loss of such magnitude can be most detrimental in situations that involve an engine failure immediately after takeoff when the aircraft is at low altitude and has a lower airspeed, thus a pilot has less time to recover and stabilize the aircraft. A procedure has been developed that teaches pilots to reduce the drag to a minimum and climb to a safe altitude in the event of an engine failure on takeoff.

Because a failing engine creates yaw, a pilot needs to compensate for the yaw and center the aircraft by pushing one of the rudder pedals. To determine which engine is failing, pilots are taught to refer to the “dead leg–dead engine” principle, which states that the leg not pushing the rudder pedal is on the side of the failed engine. Identification is verified by pulling back the throttle of the presumably dead engine; if no change in engine sound or aircraft direction of flight occurs, the identification was correct. After that, the engine is feathered. This procedure is called identify-verify-feather (I-V-F) (Gardner, Schiff, & Bringloe, 2011). Although no exact date for when the method had been introduced was found, the oldest publication available to the researcher that mentions the “dead leg–dead engine” principle dates back to 1973 (Bramson & Birch, 1973). Thus, this method has been recommended and practiced for at least 45 years.

It is important to ensure that the procedure used for identification of a failed engine is efficient enough for engine failures at any point in the flight and particularly at the critical phase of flight. However, multiple accidents that occurred in the past had involved an engine misidentification by a pilot which led to a complete loss of power and a subsequent crash. An alternative method of identification of a failed engine in propeller aircraft was proposed and tested in this study in an effort to improve the procedure for identifying a failed engine. Pilots who use the alternative method were expected to identify an engine failure quicker and be more
accurate in the identification. Accuracy and response time are two key factors in the performance of identification of a failed engine.

**Statement of the Problem**

A notable part of training literature available to multi-engine pilots is devoted to engine failures on takeoff. These training materials define recommended actions for engine failure during the takeoff roll, as well as after the aircraft lifts off the ground. The first hundreds of feet after takeoff are the most crucial and require a pilot to be prepared for an abnormality, including an engine failure. Despite the abundance of available literature and training, there is evidence that pilots have issues coping with engine failures (Aviation Safety Council, 2016; National Transportation Safety Board, n.d.; Plant & Stanton, 2012; Sallee & Gibbons, 1999; South African Civil Aviation Authority, n.d.). It is highly probable that the poor efficiency of the I-V-F method is a contributing factor for these troubles with handling engine failures, and it has a potential to lead to more errors instead of problem rectification.

The reliance on the haptic sensory channel for information input in the “dead leg – dead engine” method may be a disadvantage. First, identification of a leg exerting force, as well as identification of any type of force exerted, requires mental resources, which are limited at a critical phase of flight. As a result, it leads to an unnecessary increase of workload. Second, the haptic channel is used for decisions, but only simple and more reactive decisions. In an emergency situation, which is frequently associated with stress, the information processed through the haptic channel might not be given priority or credibility despite previous training. This method may only lead to greater confusion.

Furthermore, the recovery procedure requires deflecting the rudder to compensate for the yaw, but rudder deflection alone does not solve the problem. During normal flight, pilots
sometimes deflect rudder in the direction of a turn to increase its rate. For example, pushing on the left rudder pedal while banking to the left will make the airplane change its heading faster. This creates a mental model that associates direction of turn with the rudder pedal pushed. Such a mental model might also be unintentionally used by pilots in emergency situations; in the engine failure situation, this model will be completely wrong, and its application can lead to a disaster. Rudder deflection in the wrong direction (e.g. a pilot pushing on the right leg when the airplane is already yawing to the right) can easily go unnoticed – deflection of ailerons in the opposite direction will center the airplane. Peter (1981) reported observing such behavior in the past. This does not include the effort required to measure and compare the forces exerted by both legs on rudder pedals to identify which leg applies more power.

**Purpose Statement**

The proposed alternative method of identification of an engine failure, which would eliminate the use of the haptic sensory channel, may reduce the use of pilot mental resources. The method is based on a visual sensory channel (as opposed to haptic), and it requires an addition of system (panel) which would illuminate a red light (on either right or left side of the panel) and indicate to a pilot which engine has failed. The assumption is that looking at an indicator instead of trying to feel which leg exerts the force is quicker, more intuitive, and less resource-demanding. In real emergencies (and not simulated), when a person is subjected to life-threatening dangers, their performance is impaired (Baddeley, 2000). Therefore, avoiding any complex procedures for problem-solving is crucial, and a method for failed engine identification needs to be intuitive. The purpose of this study was to test the effectiveness of using this alternative method of identification of a failed engine when compared to the traditional method which is recommended for use nowadays. Participants were asked to perform three flights, and
each flight involved an engine failure. The effectiveness of the method was measured as response time to a failure and accuracy of identification.

**Hypotheses**

The following hypotheses were stated:

$H_{01}$: There is no difference in accuracy of engine identification between participants using the traditional and the alternative method.

$H_{02}$: There is no difference in average response time for all three flights between participants using the traditional and the alternative method.

$H_{03}$: There is no difference in response time across the three flights between participants using the traditional and the alternative method.

**Delimitations**

The scope of this research was limited to student pilots from Embry-Riddle Aeronautical University (ERAU). Only student pilots with no experience in multi-engine aircraft were eligible for participation. The delimitation was in place to avoid possible confounds related to previous experiences with flying multi-engine aircraft or learning about the “dead leg – dead engine” concept.

**Limitations and Assumptions**

Due to restrictions to participants imposed in this study, the results of this study could only be generalized across student pilots at ERAU who have not started their multi-engine training. The effect of the proposed method on pilots who have experience in the industry as well as pilots who are certified to operate multi-engine aircraft is unknown.

Performing flights in a real airplane would significantly increase mundane realism of the study. However, participants in this study were not rated to fly a multi-engine aircraft, and
simulating an engine failure on takeoff in a real-world setting would not only be complicated, but also unsafe. In such cases, the risks of performing the study would outweigh the benefits. Therefore, the experiment was performed with the use of a flight training device station that resembled a cockpit of a twin-engine general aviation aircraft.

**Definitions of Terms**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical phase of flight</td>
<td>Typically taxi, takeoff, and final approach for landing.</td>
</tr>
<tr>
<td>Engine misidentification</td>
<td>An occurrence in which a pilot does not correctly identify the engine (e.g., if right engine fails, and the pilot thinks that it is the left engine).</td>
</tr>
<tr>
<td>Engine throttle</td>
<td>A mechanism that controls the amount of fuel supplied to the engine, thus increasing or reducing output power.</td>
</tr>
<tr>
<td>Engine feathering</td>
<td>A procedure that involves moving propeller blades parallel to the airflow to reduce drag after an engine failure.</td>
</tr>
</tbody>
</table>

**List of Acronyms**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/C</td>
<td>Air Conditioning</td>
</tr>
<tr>
<td>ATC</td>
<td>Air Traffic Control/Controller</td>
</tr>
<tr>
<td>ESP</td>
<td>Engine Status Panel</td>
</tr>
<tr>
<td>GA</td>
<td>General Aviation</td>
</tr>
<tr>
<td>I-V-F</td>
<td>Identify, Verify, Feather</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>MEL</td>
<td>Multi-Engine Land</td>
</tr>
<tr>
<td>NTSB</td>
<td>National Transportation Safety Board</td>
</tr>
<tr>
<td>SA</td>
<td>Situation Awareness</td>
</tr>
<tr>
<td>WM</td>
<td>Working Memory</td>
</tr>
</tbody>
</table>
Chapter II

Literature Review

Engine misidentification

According to Sallee and Gibbons (1999), for the period from 1985 to 1997, of 13 documented in-flight engine shutdowns in turboprop aircraft, six involved the wrong engine, i.e. in almost 50% of the cases, the pilots shut down the wrong engine. However, Sallee and Gibbons (1999) advise that this data be interpreted with caution because the number of the flights was too small. The latter is due to the fact that not many turboprop aircraft, especially at the time of the research, were equipped with flight data recorders or cockpit voice recorders. Despite the limitation, the results of the study conducted by Sallee and Gibbons underline the importance of pilot error in failed engine identification. Similar results were acquired by Wildzunas, Levine, Garner, and Braman (1999; as cited in Aviation Safety Council, 2016) after they interviewed U.S. Army UH-60 helicopter pilots. Forty percent of interviewees stated that they had confused engine throttle levers during an emergency in an actual or simulated flight.

Research performed by Wildzunas et al. (1999; as cited in Aviation Safety Council, 2016) and Sallee and Gibbons (1999) provides compelling evidence to the issue. To acquire further information related to the problem and gather opinions from people who have experience flying twin-engine propeller aircraft, and who do that on a daily basis, a survey was distributed among pilots of a U.S. regional airline that operates twin-engine turboprop aircraft. This survey served as a follow-up to the research studies performed by Sallee and Gibbons and Wildzunas et al.

For turbofan engine powered airplanes, the trend is not very different – out of 79 analyzed events that happened over a period of 39 years (from 1958 to 1997), 23 (or
approximately 29 percent) involved the shutdown of a good engine, and four involved throttling a good engine (Sallee, & Gibbons, 1999). British Midland Flight 92 that crashed short of East Midlands airport in the United Kingdom in 1989 (also known as the Kegworth Air Disaster), is an example of consequences of engine misidentification in turbofan engine powered aircraft (Department of Transport, 1990). Furthermore, the research performed by Sallee and Gibbons did not include general aviation (GA) accidents and incidents because GA aircraft rarely have flight data recording devices installed. Nonetheless, the National Transportation Safety Board (NTSB) database includes reports of multiple fatal accidents involving GA aircraft that resulted from an engine failure on takeoff (NTSB, 2018). Although the definite cause of those accidents will never be determined, it is probable that improper engine identification was a contributory cause to many of them.

These findings indicate that wrong engine shutdown problem is widespread, and it is not restricted to one aircraft or engine type only. The factors and human limitations that can possibly lead to wrong identification of a failed engine in twin-engine aircraft will be covered throughout this chapter.

In a discussion on wrong engine shutdown accidents, Plant and Stanton (2012) addressed the Schema Theory, which claims that previous experiences shape a human’s view of the world and situation that he/she is in. The decision is made based on this view; thus, past experiences play a key role in decisions made during the flight. Neisser’s Perceptual Cycle model (Plant & Stanton, 2012) expands this view and describes human behavior in the environment as cyclical. An operator makes a decision about which action to perform at a specific point of time based on her or his mental representation of the world, or a schema. The schema is created from past experiences with the environment; these experiences will differ between individuals. A triggered
action performed by an operator changes the environment, shaping its representation in the operator’s mind. The change in the environment will trigger another schema, which will call for the next action. The Perceptual Cycle Model describes these processes as continually changing from bottom-up (environment triggering the schemata) to top-down (operator’s actions changing the environment). This cyclical nature of the model is also referred to as reciprocity (Plant & Stanton, 2012).

**Error types**

Human error accounts for at least 70% of accidents that happen in aviation (Sarter & Alexander, 2000). There are classifications that categorize errors based on the initial intention for the action and the decision-making process that triggers the action. The simplest classification commonly met in research is the division of errors in omissions and commissions. Omissions occur when a pilot does not perform a required action. In case of commissions, however, an action is performed, but it is performed either at an inappropriate time or in an inappropriate manner (Strayer & Drews, 2007). Furthermore, another theory divides errors into mistakes, slips, and lapses by their nature. Mistakes can be described as bad decision making as to what action to take. Slips and lapses, on the other hand, involve good decision making but bad execution of an action. More precisely, slips are actions taken in an inappropriate manner, while lapses are actions taken at an inappropriate time (Sarter & Alexander, 2000). Therefore, mistakes can be considered equal to omission errors, and slips and lapses are related to commission errors. Figure 1 represents a schematic classification of errors.
Figure 1. Classification of errors. Commissions are classified as slips and mistakes, and lapses are considered another form of omissions (adapted from Sarter & Alexander, 2000; Strayer & Drews, 2007).

Statistically, slips and lapses are detected more often than mistakes (Kontogiannis & Malakis, 2009). However, the general trend is surprising – humans tend to unconsciously ignore mistakes that they make. An analysis of reports submitted to the Aviation Safety Reporting System (ASRS) revealed a low error detection rate for pilots who commit errors. In 76% of cases, other crew members or air traffic controllers (ATC) noticed errors committed by pilots (Sarter & Alexander, 2000).

**Human Capabilities and Limitations**

During a flight, a pilot can be subjected to stress, workload, fatigue, and time pressure. Any one of these factors may play a significant role in their performance (Staal, 2004; Tiwari, 2011). Therefore, it is important that these factors are taken into consideration in training design.

**Working memory.** Working memory (WM) is a type of memory that holds only a limited amount of information for a shorter time. A characteristic that distinguishes WM from short-term and long-term memory is that it is used while an operator is performing another task.
or a cognitive process. A phone number indicated in the address book will be stored in WM while a caller dials the number on their phone (Wickens, 1992).

The limitations of WM are capacity and time. On average, the human brain can store 7±2 chunks of information in memory (Wickens, 1992). A chunk is a single unit of information. In addition to that, information is stored in working memory for only a certain amount of time. Information stored in short-term memory will decay over time, and for that reason, it is frequently reactivated. A well-known example of information reactivation is its constant repetition (saying the information out loud). Researchers have determined that three chunks of information stored in working memory decay by half in approximately 7 seconds and one chunk of information decays in 70 seconds (Wickens, 1992).

**Stress.** Stress is a set of psychological and behavioral responses to a situation that depends on how much control an operator feels he or she has over a situation (Olver, Pinney, Maruff, & Norman, 2015). People face stressful situations frequently – the extent and degree of stress, however, differ from one person to another. In most situations, being late for work has less of an impact on individual's behavior than being late for a flight. Two types of stressors known to affect human performance are environmental (i.e., acting from outside) or psychological (i.e., acting from within). Examples of environmental stressors are temperature, lighting, and air cleanliness. Most frequently experienced psychological stressors are anxiety and fatigue. In addition to performance, stress is known to impact mental aptitude (Wickens, 1992).

Stress induced by events that occur in the aircraft cockpit is more likely to be associated with psychological stressors. Cockpit environment can add to the overall stress condition. Psychological stressors are more dangerous because they can affect everyone differently, leading to different actions. The threat of stress lies in the way it shapes pilot’s assessment of the
situation in the cockpit. In situations when evidence in the environment (e.g. instrument indications) conflicts with individual’s perceived expectation of the environment’s state, that individual might still rationalize the observed evidence just to support their well-established models. In cockpit environment, time pressure and the urge to get to the destination faster might force pilots to ignore conflicting evidence (Kontogiannis & Malakis, 2009).

**Workload.** From a safety standpoint, takeoff and landing are typically considered the most important stages of flight. Pilots need to perform many tasks almost simultaneously, such as monitor altitude and speed, follow the departure/arrival procedures, communicate with ATC, as well as plan an emergency procedure in the event of an engine (or any other critical aircraft system) fails. These tasks may increase workload, demanding more cognitive resources.

Effects of high workload can be detrimental. Casto and Casali (2013) found that pilot performance decreases with an increase in workload when it is coupled with other factors that, in one way or another, relate to perception of pertinent information. In the research performed by Casto and Casali, this factor was low quality of communication signal. On the other side, planning was found to positively influence performance. As such, pilots who are aware of an upcoming increase of workload and schedule high-workload tasks in advance are less affected by increasing task demand (Andre & Heers, 1995). Thus, a key point to solid performance in a high-workload environment is careful preparation.

Morris and Leung (2007) divided the construct of workload into three stages based on the level of tasks that they include – flying, rule-based, and high cognitive demand. Flying, the first stage, implies simply controlling the aircraft. These are the primary actions, which are usually automated. Therefore, pilots do not need focus on completion of this type of tasks. Rule-based tasks are non-automated tasks that require an action defined by a rule. This is represented by the
logic of "if this event happens, perform that action." Controlling aircraft systems or complying with ATC instructions are good examples of rule-based tasks. High cognitive demand tasks, on the other hand, do not fall within flying or rule-based, i.e. high-cognitive are non-standard tasks that do not have a pre-defined solution. Because there is no solution, these tasks require pilots to think thoroughly about further actions, leading to an increase in workload. The results of the study performed by Morris and Leung (2007) indicated that increasing workload reduced pilots' ability to perform the primary task of flying the aircraft. Additionally, the researchers found that with the increase of workload, pilots started having troubles with prioritizing tasks, even though they were instructed that their top priority was to control the aircraft (Morris & Leung, 2007).

Morris and Leung (2007) also shed more light on factors that could possibly contribute to mistakes in identifying which engine has failed. The researchers predicted that pilots lose concentration while trying to fly the aircraft and resolve engine problems at the same time and cannot decide which task is more important. If a pilot cannot keep track of her/his actions, it might lead to more errors (Kontogiannis & Malakis, 2009). Additionally, as Hart and Bortolussi (1984) found, pilot error may be a source of workload itself. Indeed, shutting down a wrong engine when another one is failing leads to a loss of all engine power available to the flight crew. This creates additional tasks for pilots, such as speed control in order to avoid stalling, engine restart, search for possible landing spots, and preparation for an emergency landing. Combined with shorter time for decision-making, these tasks may require allocation of additional cognitive resources.

The TransAsia accident (Aviation Safety Council, 2016) is a good example of how pilot error can induce workload in the cockpit. A flameout is unusual, yet not a very rare event that happens in turboprop aircraft. Pilots are trained to handle flameouts in flight simulators, and so
was the captain of the accident flight. Referring to task classification mentioned before, dealing with flameouts can be considered a rule-based task. However, the captain did not correctly identify the problem and ordered the first officer to shut down a perfectly working engine (Aviation Safety Council, 2016). As a result, what could have been handled as a rule-based task suddenly became a high-cognitive task, dragging the captain’s attention and requiring more cognitive resources that were already limited.

**Attention.** Pilots’ attention in the cockpit is essential. Even with modern airplanes that perform most tasks autonomously, it is important to monitor aircraft instruments to ensure that there are no abnormalities. However, like other human capabilities, attention is limited, and can be affected by the environment. Three types of attention commonly discussed are selective attention, focused attention, and divided attention. In the case of selective or focused attention, one task is given more priority over others. Good selective attention involves the ability to concentrate on the tasks that are most important in the current situation, ignoring the rest, even the most salient stimuli. However, under such factors as stress and fatigue, selective attention can lead to cognitive tunneling. Cognitive tunneling implies concentrating on one task to such extent that important information is ignored (Strayer & Drews, 2007). Conversely, attention can be not selective enough, resulting in failure to complete the task correctly. For example, hearing the echo of an individual’s own voice while speaking will most likely result in disruptions of their speech (Howell & Powell, 1987). Lastly, a crucial aspect that has a potential of affecting cockpit operations is expectancy. When performing tasks, especially automated tasks, people tend to foresee outcomes and expect the environment to change based on their predictions. It could be a change in engine parameters, flight paths, or ATC instructions. In reality, events may occur in a way that is contrary to the expected. Expectancy poses a threat to selective attention. In
environments with multiple events, expected information may have priority over other information (Strayer & Drews, 2007). If the expected outcome is unreliable or not consistent with the actual situation, it might lead to pilot error.

Divided attention differs from selective or focused attention. The term "divided" means that attention is separated between ongoing tasks. In different situations, a person might select the incorrect task to attend to or might not be able to focus on it because of other distractions in the environment. This is one of the limitations of divided attention (Wickens, 1992). The main limitation is the number of tasks that a human can focus on before performance starts to decrease. Several researchers (Fisher, 1984; Julesz, 1981; James, 1980; as cited in Strayer & Drews, 2007) have suggested the human brain can pay attention to up to four tasks concurrently. According to other research (Pylyshyn and Storm, 1988; as cited in Strayer & Drews, 2007; Pylyshin, 2004; as cited in Strayer & Drews, 2007) the accuracy of attending to four tasks simultaneously can be as high as 87%. However, the addition of more tasks results in the drastic decrease of speed and accuracy. Wickens (2002) proposed a Multiple Resource Model, which suggests that resources to which the human brain attends can be grouped based on their characteristics. Tasks in different dimensions (e.g. visual and auditory) generally will not affect performance because different cognitive resources are involved. However, the tasks will interfere if the use of working memory is required to process received information.

Apart from its limitations, attention can be influenced by outside factors. For example, anxiety, which is frequently associated with an in-flight emergency, can make pilot's gaze behavior more chaotic, and has a potential to reduce his or her performance (Allsop & Gray, 2014). Such gaze behavior leads to an individual’s inability to read aircraft instrument parameters correctly, wrong understanding of the situation, and, consequently, errant decisions.
Mental Models

During flight training and real life, pilots build mental models, or an understanding of how elements of their surroundings operate (Rowe & Cooke, 1995). For example, a mental model is a pilot's expectation of an increase in aircraft speed with forward displacement of the engine throttle lever. Pilots need mental models to know what aircraft systems and parameters to monitor at various stages of flight. If the aircraft behavior does not go along with the pilot’s mental model, a cognitive mismatch occurs (Baxter, Besnard, & Riley, 2007). Cognitive mismatches can be either real, when the discrepancy between pilot's mental model and reality exists, and perceived, when a pilot thinks there is a discrepancy, but in fact there is not (Baxter et al., 2007).

A threat that mental models impose on operator’s action is that they are considered valid by an operator as long as the information input is consistent with the operator’s expectations. If the information coming from the environment is not consistent with the operator’s mental model, the operator is more likely to reject that information than update the established mental model (Besnard, Greathead, & Baxter, 2004). A use of an incorrect mental model might work most of the time, but it can lead to bad consequences when a situation changes. Dattel, Durso, and Bedard (2009) provided an example in which one student pilot associated increase of engine thrust with the increase in altitude. Such a model was incorrect, because an increase of thrust results in an increase in airspeed, but not necessarily altitude. The mental model worked flawlessly during normal flights, but nearly caused a fatal accident when a spin occurred in flight. In another example, the author of this paper experienced engine overheating in his personal vehicle while driving and using the air conditioning (A/C) system. The overheating was caused by a leak in the engine cooling system. However, this event led to a wrong mental model
that working A/C caused the engine to overheat, whereas in reality the A/C and the engine cooling system in that vehicle were not connected to each other.

Therefore, it is crucial that pilots develop appropriate mental models. Along with accuracy, another critical characteristic of a mental model is time sensitivity. Whenever applied, mental models should relate to the most current state of the system. For example, when the flight crew of British Midland Flight 92 felt engine vibrations and smelled smoke in the cockpit, the captain, knowing the air conditioning system supplied air to the cockpit from the right engine, identified it to be troubling, and shut it down. However, the captain of the aircraft usually operated Boeing 737-300 aircraft, but on that flight, he was piloting the -400 variant. The operation of the air conditioning system in -400 series differed from the -300 series, and, consecutively, from the captain’s mental model (Department of Transport, 1990).

**Situation Awareness in Pilots**

An important construct that is vital for safe aircraft operation is situation awareness (SA). SA can be divided into three levels. The first level is perception of current situation. The second level is the understanding of it. The third level is projection. At the projection level, an operator uses all information that he or she received and processed (Level 1 and Level 2) to predict the events that might occur in a given situation (Endsley, 1995). In other words, a pilot with good SA will be aware of what is currently happening, know what to expect during each stage of the flight, and understand what consequences her or his actions will bring. Understanding and projection are considered the most important levels of SA in the aircraft cockpit. Conversely, most accidents happen because of a Level 1 SA failure (Jones & Endsley, 1996). Hartel, Smith, and Prince (1991; as cited in Salas, Prince, Baker, & Shrestha, 1995) have determined that lack of good SA has contributed to many accidents in the past. SA is also closely related to the
construct of mental model – an analysis of aviation incident and accident reports revealed that nearly 7% of the occurrences resulted from either poor mental models or lack thereof (Jones & Endsley, 1996).

**Haptic vs visual channels**

More than 80% of information perceived by the human brain comes through the visual channel (Geruschat & Smith, 2010). Hence, the use of a different sensory channel for identification of a failed engine may increase its accuracy and reduce the risk of a mistake. This assumption is supported by the data showing that people are more likely to notice visual cues over auditory or haptic (Hecht & Reiner, 2008). Moreover, if both visual and haptic channels are involved, people tend to rely on the former, even if they are aware that information perceived through the visual channel is less reliable than the information perceived through the haptic channel (Xu, O’Keefe, Suzuki, & Franconeri, 2012). When applied to the aircraft cockpit environment, this finding suggests that pilots are more inclined to look at what leg is not pushing the rudder pedal rather than sense it.

There is anecdotal evidence of the benefits of using visual channel in an aircraft cockpit, which is mirrored in aircraft documentation and practical flying. Aircraft checklists for emergency situations emphasize visual scanning of instruments to determine aircraft condition after the situation is stabilized. Similarly, pilots frequently report scanning aircraft instruments for any inconsistencies during abnormal situations in flight. In the Kegworth Air Disaster example (Department of Transport, 1990), Plant and Stanton (2012) claim that had the pilots looked at the engine parameters, they could have avoided the engine misidentification which led to the accident.
Summary

Research shows that misidentification of a failed engine is not an unlikely event, and almost 50% of inflight engine shutdown cases involving turboprop aircraft that occurred in a 10-year timespan involved a shutdown of a wrong engine. The misidentification problem is widespread and is not pertinent to turboprop aircraft only.

A pilot error is frequently the result of other factors that affect performance. Human capabilities can be influenced by external factors (e.g., environment) and internal factors (e.g., stress, fatigue, or increased workload). Psychological stressors, which are induced by an operator, can impair the operator’s judgment. It is not unlikely for a stressed person to ignore unexpected or unwanted information in favor of well-established models of the environment.

Errors related to engine misidentification often occur in high-workload situations, particularly on takeoff. Increased workload can cause difficulties with task prioritization which, in turn, leads to inability to focus and loss of situation awareness in the cockpit. Lack of SA makes a pilot more vulnerable to an error. Sometimes, on the contrary, errors can become the source of additional workload. Therefore, it is crucial for a pilot to always maintain good SA and have developed well-established mental models. A pilot who is aware of her or his surroundings and has an understanding of the environment is more likely to be prepared for abnormal situations that can result from inappropriate actions.

This research focused on testing a method of identification of a failed engine that is based on a visual sensory channel, as opposed to the haptic sensory channel. Several studies in the past have uncovered the benefits of visual channel over both haptic and audio channels. Moreover, in several situations, visual channel received higher credibility than haptic channel even though participants were aware that the information coming through the visual channel was less reliable.
Chapter III

Method

Research Study 1 (Airline Survey)

Participants. Forty-nine pilots from one regional U.S. airline participated in the survey. All participants were employed as pilots and had experience operating twin-engine turboprop aircraft. Participation was voluntary; each pilot received a link with an invitation to participate.

Materials and Apparatus.

The survey was created through https:// surveymonkey.com website (SurveyMonkey). Collected data was categorized and analyzed with the use of Microsoft Excel and IBM Statistical Package for Social Sciences (SPSS) software.

The survey consisted of 10 questions. There were five open-ended questions, four questions with categorized (Yes/No) answers, and one scale item. The questions asked pilots about their experience (in years and flight hours) in flying twin-engine turboprop aircraft and twin-engine aircraft in general, experience with handling an engine failure (in simulator and in real life), and their opinions regarding the current method of identification of a failed engine. Three questions included categorized answers that asked the respondents to manually type in additional comments. Appendix A is a full copy of the survey.

Procedure. The link to the survey was distributed to the pilot group of the airline by the airline’s safety department via email. The email contained a short description of the survey and its purpose, researcher’s name and contact information, and the link to the survey. Once the link was opened, pilots were prompted to read the consent form and choose to either agree or disagree to participate in the survey. Pilots who agreed to participate were redirected to the
survey. Pilots who did not agree would be redirected to the last page of the survey and prompted to close the browser window.

The results of the survey were scored automatically by SurveyMonkey. The scored data was exported into an Excel spreadsheet and used by the researcher for the analysis afterward.

**Research Study 2 (Simulated engine failure)**

**Participants.** Fifty participants were sampled from ERAU flight students on the Daytona Beach, FL, campus. The recruiting process included advertisements posted on the bulletin boards around the campus and an announcement email that was sent out via the Flight Training Department to all student pilots who were eligible for participation. The sample was limited to participants who had obtained their private pilot licenses (PPL) but who did not have a multi-engine rating (MEL) and had not started the corresponding training at the time of the participation. This restriction was put in place to limit any potential confounds due to possible previous experience with emergency procedures in multi-engine training that a participant might have acquired.

Participants were randomly assigned to two groups: 25 participants in the Traditional Method Group (Traditional Group) and 25 participants in the Alternative Method Group (Alternative Group). Each participant was paid $20 for about 45 minutes of participation.

**Materials and Apparatus**

**Flight training device.** Laminar Research X-Plane 11 flight training device was used to simulate the flight environment. X-Plane 11 is capable of simulating a flight in any part of the world, and it provided the researcher an opportunity to set weather conditions (visibility, cloud layer types and ceilings, temperature, etc.) and choose from a number of pre-installed aircraft, including twin-engine propeller aircraft. To increase mundane and experimental
realism, participants used a computer with a three-monitor setting with a 120-degree view for better surrounding view, and a control panel with a yoke, rudder pedals, six levers (left/right engine throttle levers, left/right engine propeller RPM levers, left/right engine mixture levers), and switches similar to switches installed in a real GA twin-engine aircraft.

The computer setup was installed on a stationary platform (see Figure 2). A limitation of the setup was lack of motion simulation; as such, participants were not able to feel the aircraft yaw from an engine failure or receive any other haptic feedback from the aircraft. Thus, this was not considered a high-fidelity training device. A use or implementation of a full-motion platform would exceed time and budget constraints imposed on this research study. Despite the limitation, participants were able to detect an engine failure by visually noticing a change in aircraft yaw and hearing a change in engine sound.

Figure 2. Flight training device setup.
X-Plane 11 internal features were used for recording aircraft altitude, gear position (up/down), commanded throttle lever position, commanded propeller lever position, and commanded mixture lever position for each engine. The rate of recording was 10 samples/sec. The recorded parameters were automatically saved in a Data.txt file; the data were manually transported to a Microsoft Excel spreadsheet. These parameters allowed the researcher to identify the exact moment of the engine failure with higher precision during the data scoring stage.

Audacity® recording software was used to record participant oral responses and call-outs. Participant responses were recorded with the use of a headset. IBM Statistical Package for Social Sciences was used to run statistical analyses of collected data. Adobe Premiere Pro CC was used to create training videos for participant groups.

To simulate cockpit lights that would indicate which engine had failed, an add-on was created and programmed with the help of X-Plane.Org community website forum user Steve.Wilson. The add-on was named “Engine Status Panel,” and it consisted of a white rectangle with two circles (one for each engine), which would simulate an engine status indicator. The panel was programmed in such a way that each circle would be shown in green if fuel flow for the corresponding engine was above one gallon/hour, and in red any time the fuel flow value reduced below one gallon/hour.
To control engine operation during the experiment, an application called Control Pad, developed by Laminar Research, was used on an Apple iPad device. The application was connected to X-Plane 11 training device via a Wi-Fi network. Among many options, Control Pad includes a possibility to manage aircraft failures during the flight in X-Plane. In this experiment, fuel pump failure was simulated with the use of a Control Pad.

**Questionnaires.** Before the experiment began, each participant was asked to fill out a demographic questionnaire. The questionnaire included questions about participant age, academic standing (freshman, sophomore, junior, etc.), obtained pilot licenses, and flight experience in flight hours.

After the experiment, each participant was asked to fill out a post-flight questionnaire. The questionnaire asked participants’ confidence in accuracy and response time in identification of a failed engine for each flight, overall confusion in regard to which engine was failing, what method was used for identification, and it included room for suggestions regarding...
improvements to the method of identification of a failed engine used by the participant. Appendices B and C are copies of the questionnaires.

**Training video.** All participants watched an 8-minute long training video that covered the main aspects of operation of a twin-engine propeller aircraft using a Beechcraft Baron 58 as an example. The video explained thrust imbalance created by the engines and the concept of the critical engine and included basic precautions for flying a twin-engine aircraft. The second part of the video focused on a step-by-step set of actions for handling an engine failure in a twin-engine propeller aircraft, including the Identify-Verify-Feather procedure. Two versions of the video were created – one for the Traditional and one for the Alternative Groups. The video for the Traditional Group instructed participants to identify a failing engine by using the "dead leg – dead engine" principle. The video for the Alternative Group included instruction for the alternative method of failed engine identification. Instead of using the “dead leg – dead engine” method, participants were instructed to look at the Engine Status Panel to identify an engine that was failing. This was the only difference between the two videos. Appendix D is a copy of the script which was recorded and played over the video. The script was created by the researcher and checked for accuracy by an experienced twin-engine turboprop aircraft pilot.

A short quiz with four multiple-choice questions was created to be shown after the video. The questions referred to the main principles discussed in the training video. The quiz was used to assess participants’ overall understanding of the video, as well as reiterate the main aspects of operation of multi-engine aircraft. Most participants scored four out of four on the quiz; therefore, it was not used in the analysis of collected data. A copy of the quiz is in Appendix E.

**Flights.** There was a total of four flights to be performed during the experiment – one practice flight and three test flights. The practice flight was configured to a X-Plane 11 default
Beechcraft Baron 58 aircraft which was located at Daytona Beach International Airport (KDAB) runway 25R. The test flights were also configured to Baron 58, and the routes were designed as follows:

- Flight 1: KDAB – KDED;
- Flight 2: KSGJ – KDAB;
- Flight 3: KSFB – KDAB.

Flight conditions were set to daylight with clear skies, no winds, and no precipitation. No other traffic aircraft was simulated to avoid participant distraction. The aircraft navigation system was set for direct route to the destination, and the track was shown on the aircraft’s navigational display. Participants were provided with a paper sheet with basic instructions for each flight and weather information in the form of a routine weather report (METAR) for each departure and destination airport. Participants were allowed to refer to this information at any time during each flight. Appendix F is a copy of the information sheet.

Each flight was designed to include an engine failure on takeoff. Flight 1 included a right engine failure 30 seconds after aircraft rotation. Flight 2 included a left engine failure 20 seconds after the aircraft rotation. Flight 3 included a left engine failure 45 seconds after the aircraft rotation. The rotation was defined as the moment immediately after the aircraft started the climb. The order of flights was the same for each participant. To simulate an engine failure, the researcher manually engaged fuel pump failure for the corresponding engine using the Control Pad application. This simulated failure immediately reduced the fuel mixture parameter to zero, leading to a complete loss of engine power.

**Procedure.** Before the experiment began, each participant was given a consent form that described the purpose of the study, risks and benefits, and contact information of the research
team. After signing the consent form, participants were asked to fill out a demographics questionnaire. After the questionnaire, each participant watched a training video which corresponded to the group to which a participant was assigned (Traditional/Alternative). At the end of the video, participants completed the quiz. Participant score was assessed after the completion of the quiz, and if a participant chose a wrong answer for any of the questions, that concept was discussed during the practice flight to ensure that the participant had a good understanding of the topic.

After the quiz, participants proceeded to the flight training device platform. Each participant was initially provided about 5 minutes to perform a short practice flight, get accustomed to the training device setup and practice the procedure for identifying a failed engine. Participants were instructed to takeoff and climb to 2,000 feet. Once at the assigned altitude, the experimenter demonstrated an engine failure to the participant. The experimenter announced to the participant which engine was failing, then pulled back the mixture lever of the corresponding engine to zero for 5 seconds, then returned the lever to the full mixture position. The demonstration was performed once for each engine. Each participant was asked to verbally confirm that she or he had the understanding of what a failing engine looked like in the flight training device setting. After the demonstration, each participant was given an opportunity to practice the procedure of handling an engine failure. Participants were instructed to expect an engine failure, but they did not know which engine it was. The experimenter intentionally failed the right engine midflight and asked participants to follow the steps from the training video to identify and feather the failed engine.

After the practice, participants were asked to fly three test flights. Participants were instructed to takeoff, fly runway heading, and climb to 2,000 feet, 3,000 feet, and 4,000 feet in
Flights 1, 2, and 3 correspondingly. Participants were not required to maintain ATC communications and were instructed to ignore any airspace restrictions for the experiment.

The researcher simulated an engine failure for each flight in the order discussed before. To analyze the decision-making pattern and compare response times, participants were asked to comment aloud their actions after an engine failure, as well as say aloud which engine had failed as soon as they identified it. Participant callouts, videos of each flight, and some flight parameters were recorded. Each flight was concluded after the failed engine was feathered, and participants were not required to land.

After the end of the experiment, each participant was asked to fill out a post-flight questionnaire to identify techniques or measures utilized to detect a failing engine and measure participant’s confidence in the correct identification.

Data collection and analysis. This study was a 2 × 3 mixed design. Between-subjects variable was Training method (Traditional/Alternative), and within-subjects variable was Flight (Flight 1/Flight 2/Flight 3). After the data was collected, each flight video was replayed and checked for accuracy of engine identification (Identified correctly: Yes/No). For flights with a correctly identified and feathered engine (i.e. positive result), response times to a failure were measured.

The response time was measured between the moment the engine failed and until the failure was identified by a participant. Participants were asked to say aloud which engine had failed as soon as they identified it. However, in several cases, participants either were not able to comment their actions aloud due to concentrating on coping with the engine failure, or they reacted to the failure before announcing the identified engine. If a participant did not comment aloud his or her actions, the researcher measured the time difference between an engine failure
and a moment at which the participant started to reduce the throttle of the failed engine to zero. These data were recorded by X-Plane 11 automatically. The difference was defined as response time to an engine failure.

Because the researcher was able to manually engage fuel pump failure for each flight, the recorded fuel flow parameter was used for each flight to identify the exact moment when an engine had failed. For the data analysis portion, the moment of the engine failure was stated as the moment when the engine fuel flow value dropped to zero.

Participant audio recordings were replayed to determine when the participant identified and verbally announced the failing engine. The difference between the engine failure and identification of that engine as announced by the participants was measured and scored as the response time. The response times for each of the three flights were compared later.

To test the hypothesis, four variables were measured: a) accuracy in engine identification for each participant (across three flights), b) accuracy of engine identification between groups (For Flights 1, 2, and 3), c) response times across three flights (Flight 1, Flight 2, and Flight 3), and d) response times between groups.
Chapter IV

Results

Research Study 1

A total of 73 pilots agreed to participate in the airline survey and were redirected to the survey. However, only 49 respondents completed the survey.

The average experience flying twin-engine turboprops was $M = 8.97$ years ($SD = 11.21$ years) and $M = 6,230$ flight hours ($SD = 8,695.11$ flight hours). The highest experience was 40 years and (more than) 30,000 flight hours. The average experience flying all types of multi-engine aircraft was $M = 13.91$ years ($SD = 12.53$ years) and $M = 7,229$ flight hours ($SD = 8,924.87$ flight hours). The highest experience in flying all types of multi-engine aircraft was 46 years and (more than) 30,000 flight hours. Figures 3 and 4 show participant flight experience.

Figure 4. Flight experience in turboprop aircraft.
Figure 5. Flight experience in all types of multi-engine aircraft.

Approximately one fifth (18.75%) of respondents had to utilize the Engine-Out procedure at least once during their experience with the twin-engine turboprop aircraft type that the airline owns. Of those nine cases where the procedure was utilized, four were caused by high engine oil pressure, three were defined by the respondents as “engine failure,” and for the other two, the respondents did not provide any details. When asked about experience with handling an engine failure in the simulator training, 22.86% of respondents admitted having problems with identifying a failed engine at least once, 5.71% of respondents had problems with feathering an engine, and the rest did not have any problems. Fifty-three percent of respondents have experienced some engine problems (low power, engine failure, etc.) during their real-life experience. To measure pilot comfort level with the Identify-Verify-Feather method, a scale from 1 to 5 was introduced, with 1 being “Not comfortable at all,” and 5 being “Very comfortable.” Most respondents (71.43%) indicated that they were very comfortable with the
current method, 24.49% were somewhat comfortable, 2% felt neutral, and 2% felt somewhat uncomfortable with the method.

Participants were asked to list the pros and cons of the current method. The replies to these questions were open-ended and were categorized by the researcher. If the response could not be attributed to any category, it was labeled as “Other.” Example answers that were labeled as “Other” included “universal applicability” for pros and “drag that the failed engine creates if not feathered” for cons. For the pros, most respondents indicated that it is accurate, redundant, simple, and systematic. For the cons, participants stated that the method has a likelihood of error or misunderstanding, is time-consuming, and adds to the workload (see Table 1 below).

Table 1.

*Pros and cons of the current method of identification of a failed engine in twin-engine turboprop aircraft*

<table>
<thead>
<tr>
<th>Pros</th>
<th>Count</th>
<th>Cons</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redundant</td>
<td>10</td>
<td>Likelihood of error/misidentification</td>
<td>8</td>
</tr>
<tr>
<td>Accurate</td>
<td>10</td>
<td>Time-consuming</td>
<td>7</td>
</tr>
<tr>
<td>Simple</td>
<td>8</td>
<td>Other</td>
<td>7</td>
</tr>
<tr>
<td>Systematic</td>
<td>6</td>
<td>High workload</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>Distracting</td>
<td>1</td>
</tr>
</tbody>
</table>

*Note.* Pros and Cons of the current method of identification of a failed engine in twin-engine turboprops. Items that were not relevant were categorized as “Other.”

Lastly, when asked if there could be a better method of identification of a failed engine, 29% of respondents agreed with the statement. The question included a field for suggestions, which were later also categorized. Of those pilots who agreed that there is a need for improvement, 34% suggested adding a visual indication, 22% suggested other improvements of the indications, 22% suggested adding an audio indication, and the other 22% of respondents
recommended implementing automated systems for handling an engine failure. Figure 5 provides more detail.

![Pie chart showing categories of suggestions for improvements to identify a failed engine.](image)

*Figure 6. Categorized suggestions for improvements to the current method of identification of a failed engine.*

It is worth mentioning that seven respondents to the survey (14.2%) mentioned improvement or an addition of a failed engine indication (either visual or oral). Three participants (6.1%) specifically mentioned that a cockpit light indicating which engine has failed would be the best way to identify a failing engine.

**Research Study 2**

**Demographics.** Fifty participants, 42 male and 8 female, were sampled for the study. Traditional Group had 22 males and 3 females, and Alternative Group had 20 males and 5 females.

Descriptive statistics analysis was performed to measure participant demographics. Participant age ranged from 18 years to 29 years, with mean age at $M = 20.22$ years.
The mode age for all participants was 18 years. Flight experience ranged from 50.3 flight hours to 495 flight hours. The mean flight experience for all participants was $M = 145.90$ hours ($SD = 75.45$ hours), and mode flight experience was 130 hours.

Figure 7. Participant age (all groups). Note: N = 49 because one participant did not indicate their age.
Figure 8. Participant flight experience (all groups).

**Accuracy of identification of a failed engine.** To measure the accuracy of identification of a failed engine during each of the three test flights, participant flight data and voice recordings were reviewed to reveal if participants feathered the correct engine. Among 50 participants, there was no case in which a wrong engine was feathered during any of the three flights. One participant from the Traditional Group feathered the wrong engine during the practice portion of the experiment. However, this data was not included in the analysis because it occurred during the practice flight and not during the test flight. Therefore, overall accuracy of identification was 100% for all participants.

A few participants did initially identify a wrong engine during a test flight. In each of these cases, participants moved the throttle of a presumably failed engine as per the procedure to verify identification and moved the throttle back to full position afterward. This resulted in
longer response time to an engine failure for the participant. Eight participants moved the wrong throttle initially. Of those eight, six participants were in the Traditional group, and two participants were in the Alternative group. Only one participant verbally announced the incorrect engine and corrected himself after moving the throttle of the incorrect engine. The other seven participants identified the correct engine verbally.

**Response time to an engine failure.** Response times to an engine failure were measured for each flight separately as well as averaged for three flights. The lowest response time for all participants in both groups was 0.21 seconds, which was also the lowest response time for the participants in the Alternative Group. The highest response time for the Alternative Group was 10.27 seconds. The lowest response time for the participants in the Traditional Group among the three flights was 0.61 seconds, and the highest response time was 16.35 seconds. The mean response time for all 50 participants was $M = 4.09$ seconds, $SD = 2.36$ seconds. See tables 2 and 3 for further detail.

Table 2.

*Response times for the Traditional group*

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flight 1</td>
<td>25</td>
<td>1.16</td>
<td>16.35</td>
<td>5.2308</td>
<td>3.59859</td>
</tr>
<tr>
<td>Flight 2</td>
<td>25</td>
<td>1.19</td>
<td>9.32</td>
<td>4.8304</td>
<td>2.42732</td>
</tr>
<tr>
<td>Flight 3</td>
<td>25</td>
<td>.61</td>
<td>15.22</td>
<td>5.2240</td>
<td>3.60429</td>
</tr>
<tr>
<td>All flights (average)</td>
<td>25</td>
<td>1.26</td>
<td>10.58</td>
<td>5.0951</td>
<td>2.43043</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hypothesis testing. Three hypotheses were stated in this research study. The goal of the research study was to compare accuracy of failed engine identification and response time to an engine failure between Traditional and Alternative Groups. Because all participants in both groups feathered the correct engine in all three flights, the overall accuracy value was 100%. Therefore, there was no difference in accuracy between participants in Traditional and Alternative Groups, and the first hypothesis (H\(_01\)) was retained.

To measure the difference in response time for participants within the three flights and between the two groups, a two-way mixed (2x3) Analysis of Variance (ANOVA) test was run with Group (Traditional or Alternative) as the between-subjects variable and Flight (Flight 1/Flight 2/Flight 3) as the within-subject variable.

The Mauchli’s test for sphericity was significant at \(\chi^2(2) = 13.566, p < 0.05\), indicating that the assumption of sphericity was violated. A test of within-subjects variable with Greenhouse-Geisser correction applied was not significant. Therefore, H\(_02\) was retained.

Table 4 shows mean response times for Flights 1, 2, and 3 between Traditional Group and Alternative Group. No significance was found for the Group x Flight interaction. However, as can be seen in the table, mean response time for participants in the Alternative Group was

---

Table 3.

**Response times for the Alternative group**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flight 1</td>
<td>25</td>
<td>.21</td>
<td>10.27</td>
<td>3.9836</td>
<td>2.84272</td>
</tr>
<tr>
<td>Flight 2</td>
<td>25</td>
<td>.59</td>
<td>8.25</td>
<td>2.6932</td>
<td>1.96077</td>
</tr>
<tr>
<td>Flight 3</td>
<td>25</td>
<td>.59</td>
<td>7.91</td>
<td>2.5864</td>
<td>1.89154</td>
</tr>
<tr>
<td>All flights (average)</td>
<td>25</td>
<td>.63</td>
<td>7.05</td>
<td>3.0877</td>
<td>1.84227</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
reducing steadily from Flight 1 to Flight 3, while there was no such reduction for the participants in the Traditional Group.

Table 4.

*Group x Flight interaction with mean response times to an engine failure*

<table>
<thead>
<tr>
<th>Group</th>
<th>Flight</th>
<th>Mean</th>
<th>Std. Error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>Traditional</td>
<td>1</td>
<td>5.231</td>
<td>.649</td>
<td>3.927</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.830</td>
<td>.441</td>
<td>3.943</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.224</td>
<td>.576</td>
<td>4.067</td>
</tr>
<tr>
<td>Alternative</td>
<td>1</td>
<td>3.984</td>
<td>.649</td>
<td>2.680</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.693</td>
<td>.441</td>
<td>1.806</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.586</td>
<td>.576</td>
<td>1.429</td>
</tr>
</tbody>
</table>

The test of the between-subjects variable, Group, was significant at $F(1,48) = 10.83$, $p = 0.002$. This result shows that the average response time for all three flights was significantly lower for participants in the Alternative Group ($M = 3.09$ seconds, $SD = 1.84$ seconds) than the response time for participants in the Traditional Group ($M = 5.09$ seconds, $SD = 2.43$ seconds). Therefore, $H_{03}$ was rejected.

**Qualitative data**

Each participant was asked to fill out a short survey after the test flights were competed (post-flight survey). The purpose of the post-flight survey was to measure participant confidence in identification of the failed engine and provide an opportunity to suggest improvements to the method used.

The majority of participants felt that they identified a failing engine correctly for each of the three flights. No participants from the Alternative Group responded negatively to this.
question, and only two participants in the Traditional Group responded negatively to this question. More participants were unsure whether they identified a failing engine in an adequate amount of time. Four participants from the Traditional Group and two participants from the Alternative Group responded negatively to this question. Tables 5 and 6 provide more details.

Table 5.

*Participant confidence in accuracy of failed engine identification*

<table>
<thead>
<tr>
<th>Do you feel that you identified a failed engine correctly for each of the three flights?</th>
<th>Flight 1</th>
<th>Flight 2</th>
<th>Flight 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Group</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Alternative Group</td>
<td>25</td>
<td>0</td>
<td>25</td>
</tr>
</tbody>
</table>

*Note.* Data represented as a count of participants who gave a corresponding answer (Yes/No) to the question.

Table 6.

*Participant confidence in adequacy of time spent on identification of a failed engine*

<table>
<thead>
<tr>
<th>Do you feel that you identified a failed engine in adequate amount of time during each of the three flights?</th>
<th>Flight 1</th>
<th>Flight 2</th>
<th>Flight 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Group</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>Alternative Group</td>
<td>24</td>
<td>1</td>
<td>24</td>
</tr>
</tbody>
</table>

*Note.* Data represented as a count of participants who gave a corresponding answer (Yes/No) to the question.

Participants were asked how confused they were about which engine was failing and were presented with a Likert scale ranging from 1 to 10, with 1 labeled as “Not confused at all”
and 10 labeled as “Very confused.” An independent samples t-test was run to compare the mean confusion value between the two groups. The result of the t-test was not significant at $t(48) = 1.343, p > .05$. The mean confusion value for the Traditional Group was $M = 2.28$, $SD = 1.27$. The mean confusion value for the Alternative Group was $M = 1.84$, $SD = 1.03$.

A majority of the participants reported using the procedure for identification of a failed engine which they were instructed to use while watching the training video (i.e., Traditional Group used the “dead foot – dead engine” principle, and Alternative Group used the Engine Status Panel). Eleven Traditional Group participants and five Alternative Group participants (16 participants in total) did not indicate that they used the instructed method. Most of these 16 participants reported that they relied on other engine instruments or the direction in which aircraft was yawing to identify which engine was failing. This data must be treated carefully because participants might have still used the instructed method but did not report it. See Table 7 for more details.

Table 7.

Method used as reported by participants

<table>
<thead>
<tr>
<th></th>
<th>Traditional Group</th>
<th>Alternative Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Dead foot”</td>
<td>13</td>
<td>ESP</td>
</tr>
<tr>
<td>Yaw</td>
<td>6</td>
<td>Yaw</td>
</tr>
<tr>
<td>Instruments</td>
<td>3</td>
<td>Instruments</td>
</tr>
<tr>
<td>RPM</td>
<td>2</td>
<td>Dead foot</td>
</tr>
</tbody>
</table>

Note. ESP – Engine Status Panel, Yaw – participants used the direction of yaw to identify which engine had failed, Instruments – participants used other engine or airplane instruments, RPM – participants used engine Revolutions Per Minute gauge to identify which engine had failed.

The last question asked participants to provide suggestions for improvements to the method of identification of a failed engine, if any. Ten Traditional Group participants and seven
Alternative Group participants recommended improvements to the method. The rest of the participants did not have any suggestions. This was an open-ended item, so the suggestions were later categorized by the researcher. Most participants proposed adding a warning sound that would announce an engine failure, adding other instruments, or training pilots to look at other aircraft instruments for identification of a failed engine. Figure 6 provides more details.

![Pie Chart]

_Figure 9._ Suggestions proposed by the participants after the three test flights were completed. Audio – adding a warning sound in the cockpit. Instruments – adding another instrument in an aircraft or teaching pilots to reference other instruments. Other (ESP) – suggested improvements to the Engine Status Panel. Other – suggestions which could not be categorized due to low count.

**Reliability testing.** Only participants with no previous experience in multi-engine aircraft were sampled for the experiment. However, more experienced pilots in one group than another would possibly affect the overall accuracy and response time for that group.

To test for reliability of the results and check for confounds, an independent-samples t-test was performed between the Traditional and Alternative Groups. The test revealed no significant difference in mean flight hours between the two groups, meaning that neither group
included participants who were significantly more experienced. The average experience for participants in Traditional Group was $M = 158.68$ hours, $SD = 93.75$ hours, and the average experience for participants in Alternative Group was $M = 133.13$ hours, $SD = 49.93$ hours.
Chapter V

Discussion

The purpose of the study was to test the effectiveness of the alternative method of identification of a failed engine. Two research studies were completed in the study. A survey that was distributed among airline pilots who operate twin-engine turboprop aircraft (Research Study 1) provided better insight on the effectiveness of the current method in the industry and pilots’ attitudes toward the procedure.

The analysis of the survey demographics data shows that pilots who responded to the survey questions were well-experienced in flying twin-engine turboprop aircraft. The results indicate that 1/5 of participated pilots have had difficulties with identification of a failed engine during their simulator training in the past. This, along with the fact that almost 1/3 of respondents agreed that there could be a better method of identifying a failed engine, shows that the method of failed engine identification in twin-engine turboprop aircraft (“dead foot – dead engine”) might not be efficient enough, and many real-world pilots would benefit from a potentially better method. Several survey respondents have mentioned that time is of the essence in situations with an engine failure during a critical phase of flight, such as takeoff, and the current method can be time-consuming.

Many pilots considered the current method of identification of a failed engine systematic and simple and did not feel that there was a need for an improvement. This does correlate with survey respondents reporting that they were very comfortable with the method. These ratings could be related to the fact that the training procedure for handling an engine failure is well-taught and practiced but is rarely used. Among all survey respondents, only 19% reported having utilized the procedure in their career. If used one or two times in real life, the method might feel
convenient for that particular pilot in a particular situation. The pilots might also feel more comfortable if the aircraft operated by the airline are equipped with automated systems to aid with engine identification. For example, the ATR-72 in the TransAsia accident was equipped with a system that feathered one of the engines automatically after the flameout occurred (Aviation Safety Council, 2016). In other cases, pilots seemed reluctant to change in the current training methods and procedures. Moreover, one pilot proposed adding a few extra steps to the current method to make it more systematic.

The feedback from the airline pilots does slightly align with the results of the post-flight survey filled out by the Research Study 2 participants who were student pilots. Although the majority of Traditional Group participants in Research Study 2 did not have any suggestions regarding the method of the failed engine identification that they used, those who left comments indicated, for the most part, that a visual or oral indication would be a beneficial improvement to the method.

The idea for this survey was developed from the study completed by Wildzunas et al. (1999; as cited in Aviation Safety Council, 2016). However, it was not intended to directly replicate the study mainly because, despite several attempts, the author of this paper was unsuccessful in retrieving the original publication. The latter could be due to the fact that the research was performed by the U.S. Army Aeromedical Research Laboratory. Additionally, the original study was targeting twin-engine helicopter pilots. Despite these differences, the Research Study 1 survey findings are, to some degree, parallel to the findings published by Wildzunas et al. The information presented above indicates that engine misidentification is not local to one aircraft type or category, but, rather, is consistent across the aviation industry.
The main assumption for this research was that identifying a failed engine using a visual sensory channel is more beneficial (i.e. faster) than using a haptic sensory channel. To test the assumption, an alternative method was proposed and compared to the traditional method that is recommended for use nowadays. Participants in the Alternative Group were instructed to use an Engine Status Panel (a visual indicator) to identify a failed engine. The results revealed that Alternative Group participants were significantly faster at the identification of a failed engine in three simulated flights, and they were able to feather the engine quicker than participants in the Traditional Group. The analysis of the participant data showed that the mean response time (time measured from engine failure to a moment at which a participant announced which engine had failed) for all three flights and for each flight separately was lower for the Alternative Group participants. On average, participants in the Alternative Group identified a failed engine 2 seconds faster than participants in the Traditional Group. Moreover, after all three flights were completed, some Traditional Group participants who were instructed to follow the “dead leg – dead engine” method reported referring to the visual cues, such as observing propeller rotation, for identification of a failed engine. These findings are in line with findings by Xu et al. (2012) who claimed that humans tend to prioritize the use of visual sensory channel over haptic sensory channel. Participants in Alternative Group also reported being less confused in regard to which engine was failing (based on a Likert scale that was provided to the participants), and they were more confident in their response to a failure. More specifically, more Alternative Group participants reported that they identified a correct engine and that they identified the correct engine in an adequate amount of time.

Although no significant effect was found, the average response time reduced from Flight 1 to Flight 3 for the Alternative Group participants, while no such reduction was noted for the
Traditional Group participants. Rather, the average response times for Flight 1 and Flight 3 for Traditional Group were equal. This may suggest that participants who were taught to use a visual indicator for identification of an engine failure became acquainted with the procedure faster than the participants who were taught to use the “dead leg – dead engine” principle, and they needed less practice to improve their performance. With more power, the difference could have been significant.

Overall accuracy of engine identification was at 100%, i.e. all participants in both groups feathered the correct engine. However, the analysis of commanded throttle position revealed that eight participants initially moved back the throttle of a working engine. Although this might count as misidentification, most participants identified the correct engine verbally. These results must be treated with caution for two reasons. First, it is possible that these participants hastily moved the wrong throttle lever before being able to detect and process the information required to identify a failed engine. Second, for many of the participants, this research study was their first experience in flying a twin-engine aircraft training device. Not being used to the new setting, they could have identified the correct engine, but moved the opposite throttle due to lack of experience with twin-engine aircraft.

In addition to the test flight findings, some pilots who answered the survey questions in Research Study 1 stated that they would benefit from a different method because it is very easy to get confused “in the heat of the moment.” As the TransAsia accident revealed, such mistake is not impossible, and it can very easily lead to tragic consequences. Furthermore, 30% of respondents to the last survey question (“If you think there could be a better way to identify a failed engine, what is it?”) suggested using a light for identification of a failed engine, which provides additional evidence toward the support of the theory proposed in this study.
Limitations

Engine identification accuracy for this study was based on whichever engine was feathered. In other words, the accuracy for any flight would count as 0 if a participant feathered the wrong engine, and as 1 if a participant feathered the correct engine. It was mentioned before that some participants feathered the correct engine but moved the wrong throttle initially. Presuming that participants did identify a wrong engine and moved the wrong throttle in these eight cases, a correlation test was performed but was not significant. There is a possibility that, had participant actions been observed and noted during the experiment, the researcher would be able to detect instances in which participants grabbed the wrong throttle initially but did not move it. However, as stated above, only one participant announced the wrong engine verbally, and the rest did announce the correct engine.

Participant behavior was an unexpected challenge that was discovered during the data collection phase for Research Study 2. Although the training video for participants in both groups instructed them to use a certain procedure when dealing with an engine failure during the experiment (the procedure was adapted from Gardner et al., 2011), not all participants followed the prescribed steps. As such, the video instructed to ensure that full power was set and landing gear were retracted by moving the corresponding levers and handles before trying to identify which engine was failing. However, a few pilots skipped that part and began the identification of the failed engine directly after they noticed the failure. Moreover, some participants did not retract landing gear. Such behavior was comprehensible because for most of the student pilots who participated, this was the first experience in flying a twin-engine aircraft with retractable gear. Thus, they were not acclimated to retracting landing gear after takeoff.
It was initially planned to measure the time needed for a participant to feather the engine after it fails as a response time. The difference in participant behavior, however, resulted in differences in the response time depending on whether a participant followed the prescribed procedure or not. The time spent on retracting landing gear and moving engine power levers forward could hardly be considered related to the cognitive process of engine identification. Thus, using this approach to measure the response time would have low reliability. The time between an engine failure and a verbal identification by the participant was measured instead, giving a more precise value for the response time.

**Recommendations**

The implementation of a visual indicator that identifies a failed engine and informs a pilot about the failure will be of particular benefit to the GA industry. Modern commercial aircraft are equipped with annunciators and automated systems which largely aid in operation of these aircraft, as well as in handling emergency procedures. GA aircraft, on the other hand, are lacking such sophisticated systems due to the costliness of their installation and low profitability of operations, which implies no possibility of covering high costs of installation. The integration of a system similar to the Engine Status Panel will certainly incur additional expenses to the aircraft owners who elect to use the system. However, it will greatly reduce the risk of human error in emergency situations, which are known to impair performance (Baddeley, 2000) and attentional control (Allsop & Gray, 2014). Moreover, the implementation of such a system in newly-built aircraft can be performed at a smaller cost due to the widespread integration of glass cockpit displays in GA aircraft. According to the NTSB (2010), glass cockpit displays were installed in more than 90 percent of piston-powered light aircraft produced in 2006. An indicator similar to
the Engine Status Panel used in this study can possibly be integrated into the aircraft’s firmware to be shown on a glass cockpit display.

An imaginable concern against the implementation of such an indicator is additional clutter. Clutter has been found to impact search performance in individuals (Adamo, Cain, & Mitroff, 2015; Ho, Scialfa, Caird, & Graw, 2001). It is highly unlikely, however, that this issue arises in an aircraft cockpit equipped with an Engine Status Panel or its equivalent. Clutter becomes problematic when targets are grouped. An indicator for a failed engine is only required for reference when an emergency occurs, and a pilot does not need to have it constantly in sight. Thus, it may be installed aside from other instruments so that it will not be cluttered, and it will not affect pilot cognitive performance if a pilot knows its location in the cockpit. Additionally, in aircraft with glass cockpits, it may be designed as a pop-up panel that only appears if an engine fails. The Engine Status Panel in the Research Study 2 was located approximately 40 degrees to the right from a participant. It is noteworthy that some participants, on the contrary, suggested moving the Engine Status Panel closer to the center of the screen. Thus, effects of clutter in this case shall be considered minimal.

**Further research**

The Research Study described in this paper was targeting single-engine pilots with no experience in multi-engine aircraft. This allowed the researcher to sample participants with nearly equal experiences and knowledge of twin-engine aircraft operations. However, if implemented, the training method will affect not only pilots who are new to multi-engine operations, but also those who do already have MEL and are operating twin-engine propeller aircraft.
The results of the survey from Research Study 1 indicate that at least part of current MEL-rated pilots who operate twin-engine aircraft do not see the need for improvements, and, on the contrary, believe that the current method is efficient enough. To some extent, this seemed as a reluctance to new methodology or training approaches. The Schema Theory (Plant & Stanton, 2012) suggests that experience has a significant effect on human decisions, which does in part explain possible reluctance. Respondents to the survey in Research Study 1 have been trained the traditional method and have acquired thousands of hours in twin-engine turboprop aircraft. It is likely that these pilots have developed solid mental models in relation to operation of twin-engine aircraft and actions for handling an engine failure. Moreover, procedures for handling an engine failure differ between GA and transport category aircraft. A change of these mental models cannot occur instantly and may require compelling evidence for justification. Participants in Research Study 2, on the other hand, did not have experience in twin-engine aircraft, and thus were more open to learning the methodology of handling an engine failure that was provided to them.

Therefore, performing research with multi-engine-rated pilots would help determine the extent to which past training and experience affect the ability to get accustomed to a novel methodology which is based on a different sensory channel and which involves a different set of actions. It is recommended to conduct an experiment in a similar fashion, but sampling participants who are MEL-rated. Pilot training shall be developed and modified so that it would account for past experiences and underline the importance of the change in the methodology.

The implementation of a system for identification of an engine failure in aircraft will involve technical considerations. The Engine Status Panel used in this study was designed to refer to the fuel flow value for each engine. The engine status light would turn from green to red
as soon as the fuel flow parameter decreased below a pre-set number. The fuel flow parameter was chosen for feasibility purposes – the researcher was able to simulate a complete loss of an engine by simulating a fuel pump failure for the corresponding engine. Undoubtedly, a complex system, which an aircraft engine is, never fails in the same way, and thus fuel flow alone would never be a good reference. Further research is required to identify engine parameters which are the most indicative of an engine failure in twin-engine propeller aircraft. For example, Gardner et al. (2011) recommend using engine Exhaust Gas Temperature (EGT) as a reference when identifying a failing engine in flight. Using several parameters should increase reliability of the system and account for various conditions at which an engine may fail.

When asked to suggest improvements to the method of identification of an engine used in the study, many participants in Research Study 2 mentioned using audio systems as a means of alerting a pilot of an engine failure. An aural alert can be most beneficial in cruise and descent phases of flight. Engine power is reduced during the cruise flight, and during the descent phase of flight, engine throttle levers are frequently moved to idle to reduce aircraft speed. If an engine fails, the yaw will not be so salient, and an engine failure can be not instantly noticed. Because it is omnidirectional, an aural warning can attract attention faster than a visual indicator, and thus it will immediately alert the pilot that a critical issue has arisen (which is one step before identifying a failed engine) and reduce the risk of not detecting an engine failure in cruise and descent stages of flight. However, an additional audio warning can also create more confusion. It is important to understand that participants in Research Study 2 were flying in ideal conditions. In a real flight, external sounds, such as engine and wind, create additional noise inside the cockpit. An extra sound that rings or buzzes during an emergency situation can be ignored or regarded as a nuisance. Edworthy, Meredith, Hellier, and Rose (2013) found that some specific
alarms were hard to identify even after considerable amount of practice. More research is needed to identify the effect of aural warnings on pilot’s ability to timely detect a failure and identify a correct engine in flight.

**Conclusion**

This research study looked at an alternative method of identification of a failed engine in twin-engine propeller aircraft based on the assumption that the current method is lacking efficiency. Three takeoffs with an engine failure were simulated. The analysis of the collected data revealed that participants who used the alternative method could identify which engine had failed significantly faster than the participants who used the traditional method which is recommended nowadays; additionally, participants in the Alternative Group were more confident in correct identification of a failed engine and less confused in regard to which engine was failing.

These findings indicate that the use of method based on a visual sensory channel for identification of a failed engine is more efficient compared to the method based on a haptic sensory channel (the traditional method) and has a potential to reduce pilot error in a situation involving an engine failure. It is suggested to develop this research and measure the effectiveness of the proposed method in pilots who are MEL-rated. The implementation of a failed engine indicator is assumed to be most beneficial to the general aviation industry due to lack of sophisticated automated systems installed in GA aircraft as compared to airliners.
References


Appendix A

Survey for Twin-Engine Turboprop Aircraft Pilots
1) How many years of experience do you have in a twin-engine **turboprop** aircraft?  
______________ Years

2) How many hours do you have in a twin-engine **turboprop** aircraft? ________ Hours

3) How many years of experience do you have flying **all types** of multi-engine aircraft?  
__ Years

4) How many hours do you have flying **all types** of multi-engine aircraft? __________ Hours

5) In your experience flying the [Aircraft type has been de-identified for confidentiality reasons], have you ever had to utilize the Engine-Out Procedure at any time? (Not including simulator training)  
a) Yes  
b) No

6) During your simulator training, have you ever had any difficulties with handling an engine failure on takeoff? What were they?  
a) Difficulties identifying a failed engine   Yes / No  
b) Difficulties feathering a failed engine   Yes / No  
c) Other (please indicate)_____________________

7) During your real-life experience, have you ever experienced any difficulties (low power, engine failure, etc.) with an engine in a multi-engine aircraft? Please list the issues. If no, please write N/A.  
a) __________________________

8) How comfortable do you feel with the Identify-Verify-Feather procedure?  
i) (Scaled item 1-10, where 1 is Not comfortable at all and 5 is Very comfortable)

9) In your opinion, what are the pros and cons of the Identify-Verify-Feather method?  
a) Pros:____________________________
   b) Cons:___________________________

10) Do you think there could be a better way to identify a failed engine in turboprop aircraft?  
a) Yes  
b) No  
c) If Yes, what could it be?  
i) ________________________________
Appendix B

Participant Demographics Questionnaire
1. What is your age? ______ years
2. You are:
   a. Freshman
   b. Sophomore
   c. Junior
   d. Senior
   e. Graduate student
   f. Not a student
3. List all licenses and ratings you have obtained
   ______________________________________________________________
4. How many total flight hours have you accumulated? _________
5. How many of those were in multi-engine aircraft? _______
Appendix C

Participant Post-Flight Questionnaire
1. Do you feel that you identified a failed engine correctly for each of the three flights?
   a. Flight 1: Yes / No
   b. Flight 2: Yes / No
   c. Flight 3: Yes / No

2. Do you feel that you identified a failed engine in adequate amount of time during each of the three flights?
   a. Flight 1: Yes / No
   b. Flight 2: Yes / No
   c. Flight 3: Yes / No

3. On a scale from 1 to 10, to what degree were you confused about which engine was failing (overall for three flights)? 1 is Not confused, and 10 is Very confused.

   1  2  3  4  5  6  7  8  9  10
   Not confused at all   Very confused

4. What instrument/technique helped you identify a failing engine better?

5. Do you feel that there needs to be an improvement in the method of identification of a failing engine? What could be improved?
Appendix D

Training Video Script
Welcome to this short introduction to flying a twin-engine propeller aircraft. In this video, we will discuss what challenges a second engine brings to planning and performing a flight when compared to a single-engine airplane. During the experiment, you will fly a Beechcraft Baron 58, so it is important that you pay close attention to the following rules and understand the principles of operation of a twin-engine propeller aircraft.

It is clear that a twin is more powerful due to increased horsepower. Although not used at full power all the time, twin-engine aircraft can accelerate and cruise faster than single-engine airplanes. This is a great advantage when both engines operate properly. However, problems may arise when either of the two engines fails. In a single-engine aircraft, a failure of that only engine does not have a big impact on flight controls, except for the lack of thrust. In twins, on the other side, power loss of one engine is followed by differences in how an aircraft handles. Let’s imagine that we are cruising at 3000 feet. In this situation, both engines produce the same thrust, thus the airplane is well-balanced in relation to the vertical axis. So, what happens if the right engine fails? First, right wing is affected by the loss of thrust, which means that it is not pushed forward anymore, and it will move slower than the left wing. This will result in a yaw toward the right side. Second, a pilot now experiences reduced horsepower, and thus needs to increase pitch angle to maintain the same altitude. The increase in pitch produces an increase in a P-factor, the force that is generated by a descending propeller blade and that causes turning motion. Last, our right propeller is now windmilling, producing drag, and, consequently, increasing yawing moment toward the side of the failed engine. Conversely, if left engine fails, we will observe an opposite reaction of the airplane, which is a significant yaw to the left.

The concept of P-factor has to be well explained and understood. In a propeller engine, descending blades of the propeller do most of the work. Although some twins have counter-
rotating propellers, in the Baron 58, both propellers rotate clockwise. The descending blades of the right propeller are further from the aircraft centerline than the descending blades of the left propeller. Thus, more momentum is created by the right propeller. If the right engine fails, the P-factor from the left engine will cause a yaw to the right, but manageable to overcome the yaw by incorporating proper flight controls. If left engine fails, however, the P-factor from the right engine will create a greater yaw toward the failed engine. In addition to that, the slipstream created by the right engine is further from rudder centerline, meaning that the rudder will have less authority in this situation. For this reason, the left engine is called a critical engine – its failure will cause more significant control problems to the pilot. To ensure that the airplane is operative in case of an engine failure, a pilot must stay above \( V_{MC} \) – Minimum Control speed. This is the minimum speed at which an airplane will remain controllable in case if a critical engine fails. For Baron 58, \( V_{MC} \) is 81 knots.

There are several rules that apply to flying twin-engine props.

1. **Respect the aircraft limitations.** It may seem that losing one engine will result in 50% loss of vertical performance. But think about this – if your engine quits, not only does it stop producing thrust, but it also creates drag due to windmilling. Add to that the dead weight of the engine, and your vertical performance loss can go up to 90%, leaving you with only 10% of nominal performance. The effect of windmilling can be reduced by feathering the failed engine. A feathered engine has the propeller blades turned parallel to the airflow, so it minimizes drag and lets the air flow past the propeller, generating some lift for the wing. To feather the propeller, you should move the throttle, then the propeller RPM lever all the way back on the failed engine.
2. **Plan like a pro.** When doing a flight in a twin, you are flying a faster, bigger, and heavier airplane. Therefore, it’s in your best interest to prepare for possible failures or abnormalities in advance by thinking ahead of possible problems.

3. **Expect the failure.** It’s always good to have a so-called constructive paranoia, or a tendency toward suspiciousness. A safe approach is to prepare, and even expect that one of your engines will quit at some point during takeoff.

Imagine this: You just initiated you climb after takeoff, and all of a sudden you feel the yaw and can hear the difference in engine sound on the left and on the right. You have an engine failure on takeoff!

What are the best practices if you are airborne? First of all, compensate the yaw by pushing the rudder pedals in the direction opposite to the yaw. Stabilize the aircraft, and make sure that you are not losing speed and altitude. Then follow three steps for safe handling of the engine failure.

**Power up.** You need all power that is available to you. Going from right to left across the throttle quadrant, move the mixture, propeller, and throttle levers full forward.

**Clean up.** Reduce possible drag to a minimum in order to gain speed. Retract gear and flaps.

*<Group 1, Traditional Method>*

**Identify-Verify-Feather.** To avoid additional drag, you must feather the engine. We will cover this procedure during the practice flight in the flight simulator. To identify the failed engine, use the “dead foot – dead engine” principle. Identify the engine by the force you are applying on the rudder. As you compensate for the yaw caused by the failing engine, one foot is applying significant force to the rudder pedal, while your other foot is resting. The foot which is
not pushing the rudder pedal ("dead" foot) is on the side of the engine that failed ("dead" engine). Next, you need to verify that you identified the correct (in this case, failing) engine. Pull back the throttle of that engine. You should hear no change in engine sound and feel no difference in rudder pressure. This confirms that you identified the engine correctly. Finally, feather that engine by retarding the prop lever (colored in blue) all the way back.

<Group 2, Alternative Method>

**Identify-Verify-Feather.** To avoid additional drag, you must feather the engine. We will cover this procedure during the practice flight in the flight simulator. To identify the failed engine, look at the Engine Status panel on the screen. You will see two lights, one for each engine. The green light indicates that your engine is properly working. The red light means that the engine failed. For example, if the left light is red and right light is green, then your left engine fails. Next, you need to verify that you identified the correct (in this case, failing) engine. Pull back the throttle of that engine. You should hear no change in engine sound and feel no difference in rudder pressure. This confirms that you identified the engine correctly. Finally, feather that engine by retarding the prop lever (colored in blue) all the way back.

<Both groups>

Finally, fly the airplane. Assess the situation and plan your return to field if possible. Climb to a safe altitude if you can. If unable, try to stay at your current altitude. Always remember the golden rule of flying – aviate, navigate, communicate!
Appendix E

Training Video Quiz
1. A failure of the critical engine will have __________ on the control of the airplane.
   a) Bigger effect
   b) Smaller effect
   c) No effect

2. P-factor is the force that creates a turning motion, and it is generated by the ________
   blade of the propeller.
   a) Ascending
   b) Descending

3. A propeller of a failed engine should be feathered because:
   a) This will reduce drag from the windmilling propeller
   b) This will let you fly faster
   c) This will protect you from a failure of the second engine
   d) None of the above

4. In the Baron 58, if the right engine fails, a pilot will feel yaw in which direction?
   a) Right
   b) Left
Appendix F

Participant Instructions
Beech Baron 58 speeds:

|  |  
|---|---|
| $V_R$ | 80 kts |
| $V_{MC}$ | 81 kts |
| $V_Y$ | 125 kts |

**Flight 1: KDAB – KDED**

KDAB DDHHMMZ 00000KT 10SM SCT100 27/21 A2992 RMK AO2

KDED DDHHMMZ AUTO 00000KT 10SM CLR 26/20 A2992 RMK AO2

Takeoff from runway 25R, fly runway heading, climb to 2000’, then follow the track on your navigation display. Maintain 175kts upon reaching the assigned altitude. Await further instructions.

**Flight 2: KSGJ – KDAB**

KDAB DDHHMMZ 00000KT 10SM SCT100 27/21 A2992 RMK AO2

KSGJ DDHHMMZ 00000KT 10SM SCT120 27/19 A2992 RMK AO2

Takeoff from runway 13, fly runway heading, climb to 3000’, then turn right heading 180. Maintain 175kts upon reaching the assigned altitude. Await further instructions.

**Flight 3: KSFB – KDAB**

KSFB DDHHMMZ 00000KT 10SM CLR 28/18 A2992 RMK AO2

KDAB DDHHMMZ 00000KT 10SM SCT100 27/21 A2992 RMK AO2

Takeoff from runway 9L, fly runway heading, climb to 4000’. Maintain 175kts upon reaching the assigned altitude. Await further instructions.