The Effect of Feedback Delays on the Next Generation Route Assessment Tool

A thesis submitted in partial fulfillment of the requirements

For the degree of Master of Arts in Psychology,

Human Factors and Applied Psychology

By

John O. Terceman

December 2012
The thesis of John Terceman is approved by:

________________________________________  _____________
Dr. Nhut Ho                                      Date

________________________________________  _____________
Dr. Walter Johnson                               Date

________________________________________  _____________
Dr. Tyler Blake, Chair                           Date

California State University, Northridge
# Table of Contents

Signature Page .............................................................................................................. ii

Table of Contents ......................................................................................................... iii

Glossary ............................................................................................................................ v

List of Figures .................................................................................................................. vi

List of Tables ................................................................................................................... vii

ABSTRACT ..................................................................................................................... viii

Chapter 1  Introduction .................................................................................................. 1

1.1 Background .............................................................................................................. 1
1.2 Trajectory-Based Operations .................................................................................. 2
1.3 Conflict Detection and Resolution System .............................................................. 4
1.4 Route Analysis Tool ............................................................................................... 6
1.5 Feedback Delays ...................................................................................................... 7
    1.5.1 Effects of Delays on Performance ................................................................. 7
    1.5.2 Delays in Tracking Tasks .............................................................................. 9
    1.5.3 User Tolerance of Feedback Delays ............................................................. 10
1.6 Human Memory Systems ....................................................................................... 11
    1.6.1 Relevance to Pilot Performance ................................................................. 12
1.7 User Behavior and Strategy Adoption .................................................................... 13
1.8 Task Difficulty and Time Pressure ........................................................................ 14
1.9 Traffic Density ....................................................................................................... 15

Chapter 2  Statement of the Problem ......................................................................... 19

2.1 Introduction ............................................................................................................ 19
2.2 Objectives ................................................................................................................ 19
2.3 Hypotheses ............................................................................................................. 20

Chapter 3  Methods ..................................................................................................... 21

3.1 Participants ............................................................................................................ 21
3.2 Facility ................................................................................................................... 21
3.3 Materials ............................................................................................................... 21
3.4 Experimental Design and Procedures .................................................................. 22
3.5 Independent Variables .......................................................................................... 23
    3.5.1 Feedback Delay ......................................................................................... 23
    3.5.2 Traffic Density ......................................................................................... 24
    3.5.3 Counterbalance Methodology ................................................................. 24
3.6 Dependent Variables ........................................................................................................ 25
3.7 Data Analysis .................................................................................................................. 26
  3.7.1 Quantitative Data Analysis ...................................................................................... 26
  3.7.2 Questionnaire Data Analysis ................................................................................... 27
  3.7.3 Video Analysis .......................................................................................................... 27
  3.7.4 Debrief Interview Analysis ..................................................................................... 27

Chapter 4 Results .................................................................................................................. 29

4.1 Decision Time .................................................................................................................. 29
4.2 Route Distance ............................................................................................................... 32
4.3 Failed Trials .................................................................................................................... 34
4.4 Questionnaire Data ........................................................................................................ 36
  4.4.1 Perceived Difficulty ................................................................................................. 36
  4.4.2 Perceived Performance ............................................................................................ 38
  4.4.3 Perceived Level of Feedback Delay ......................................................................... 39
  4.4.4 Perceived Level of Traffic Density .......................................................................... 40
4.5 Video Analysis ................................................................................................................ 41
  4.5.1 Move and Wait.......................................................................................................... 41
  4.5.2 Reverse search ....................................................................................................... 42
  4.5.3 Articulations ............................................................................................................ 42
  4.5.4. Conflict Test Speed ............................................................................................. 43
  4.5.5. Execution of a Conflicted Route ........................................................................... 44
4.6 Debrief Interview ............................................................................................................ 44
  4.6.1. Comfort Using Display ....................................................................................... 44
  4.6.2. Level of Delay ...................................................................................................... 45
  4.6.3. Difficulty Attributed to Delay and Traffic ............................................................. 45

Chapter 5 Discussion ............................................................................................................ 46

5.1 Feedback Delay and Traffic Density ............................................................................. 46
  5.1.1 Decreases in Route Efficiency ................................................................................. 46
  5.1.2 Increases in Time on Task ...................................................................................... 47
  5.1.3 Questionnaire Data ................................................................................................. 48
  5.1.4 What Have We Learned? ...................................................................................... 50

Chapter 6 Conclusion .......................................................................................................... 53

References ............................................................................................................................. 55
## Glossary

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4D</td>
<td>Four Dimensional</td>
</tr>
<tr>
<td>ADS-B</td>
<td>Automatic Dependent Surveillance Broadcast</td>
</tr>
<tr>
<td>ATC</td>
<td>Air Traffic Control</td>
</tr>
<tr>
<td>ATM</td>
<td>Air Traffic Management</td>
</tr>
<tr>
<td>CD &amp; R</td>
<td>Conflict Detection &amp; Resolution</td>
</tr>
<tr>
<td>CDTI</td>
<td>Cockpit Display of Traffic Information</td>
</tr>
<tr>
<td>CSD</td>
<td>Cockpit Situation Display</td>
</tr>
<tr>
<td>FAA</td>
<td>Federal Aviation Administration</td>
</tr>
<tr>
<td>FDDRL</td>
<td>Flight Deck Display Research Laboratory</td>
</tr>
<tr>
<td>IRB</td>
<td>Internal Review Board</td>
</tr>
<tr>
<td>JPDO</td>
<td>Joint Planning and Development Office</td>
</tr>
<tr>
<td>LCD</td>
<td>Liquid Crystal Display</td>
</tr>
<tr>
<td>LOS</td>
<td>Loss of Separation</td>
</tr>
<tr>
<td>NAS</td>
<td>National Airspace System</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>NextGen</td>
<td>Next Generation Air Transportation System</td>
</tr>
<tr>
<td>RAT</td>
<td>Route Assessment Tool</td>
</tr>
<tr>
<td>TBO</td>
<td>Trajectory-Based Operations</td>
</tr>
<tr>
<td>TOTE</td>
<td>Test-Operate-Test-Exit</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. Clearance-Based vs. Trajectory-Based Operations ................................................. 3
Figure 2. Cockpit Display of Traffic Information ................................................................. 4
Figure 3. Using the Route Assessment Tool (RAT) to avoid a conflict .......................... 5
Figure 4. Ownship Symbol Colors. ..................................................................................... 6
Figure 5. Human Memory System ....................................................................................... 12
Figure 6. Available Solution Spaces with No Delay .............................................................. 18
Figure 7. Limited Solution Spaces Due to Delay ................................................................. 18
Figure 8. Mean Averages of Decision Time (Feedback Delay x Traffic Density) ........ 30
Figure 9. Mean Averages of Feedback Delay (Feedback Delay x Traffic Density) ....... 32
Figure 10. Timed-out Trials ................................................................................................. 33
Figure 11. Conflict Executions ............................................................................................. 34
Figure 12. Perceived Difficulty .............................................................................................. 36
Figure 13. Perceived Performance ......................................................................................... 37
Figure 14. Perceived Level of Feedback Delay .................................................................. 38
Figure 15. Perceived Level of Traffic Density .................................................................. 40
List of Tables

Table 1. Summary Data of Decision Time ................................................................. 31
Table 2. Summary Data of Route Distance ............................................................... 32
Table 3. Questionnaire Data of Perceived Difficulty .............................................. 36
Table 4. Questionnaire Data of Perceived Performance .......................................... 38
Table 5. Questionnaire Data of Perceived Level of Feedback Delay ...................... 39
Table 6. Questionnaire Data of Perceived Level of Traffic Density ......................... 40
ABSTRACT

The Effect of Feedback Delays on the Next Generation Route Assessment Tool

By

John O. Terceman

Master of Arts in Psychology,
Human Factors and Applied Psychology

Predicted increases in air traffic are driving a paradigm shift in the Air Traffic Management system. The ATM research community has identified the adoption of Trajectory-Based Operations a viable concept of operation to accommodate the traffic growth while maintaining system safety and efficiency. This transition would require pilots and controllers to use trial planner tools to create, evaluate, modify trajectories that are exchanged between the ground and airborne systems. However, various lags (e.g., in communications or computation of trajectories) can induce delays in providing the feedback to the user during conflict search and trial planning processes. This research addresses this issue by conducting an experiment to study the effects of feedback delays on pilots during conflict detection and avoidance tasks. Twelve pilots participated in 40 en-route flight part-task simulations that required them to avoid separation violations with surrounding aircraft while maintaining flight safety and fuel efficiency. Actual air traffic over Kansas City airspace was simulated in the experiment and scaled to have moderate and higher density than current day level. The traffic was displayed on the NASA Ames Cockpit Situation Display, which featured a trial planner tool equipped with conflict alerting functionality. Feedback delays of 0, 2, 4 and 8 seconds were added to the conflict alerting functionality to observe any significant changes in pilot comfort as well as route efficiency and safety. Irrespective of the levels of feedback delays, pilots showed no significant decrease in performance in moderate traffic density simulations. However, the results were considerably different in the high traffic density condition, where the efficiency decreased significantly at levels of delay greater than 0 seconds and the pilot decision time increased significantly. In addition, pilots were more likely to commit to unsafe routes in high-density conditions with feedback delays greater than 0 seconds. It is reasonable to assume that delays are impossible to completely eradicate, but perhaps they need not be completely removed from these systems to successfully meet flight and business objectives. With thorough calculations stakeholders may be able to extrapolate the amount of feedback delay that can be deemed as acceptable in order to meet safety as well as business requirements.
Chapter 1 Introduction

1.1 Background
Pilots have traditionally relied on radio communications from Air Traffic Controllers (ATC) to receive necessary changes to their flight plans due to conflicting traffic or inclement weather. However, due to anticipated increases in air traffic, advanced systems have been in development. In 1980, the National Airspace System (NAS) carried 281 million passengers. That number has grown to currently exceed 650 million (JPDO, 2012). Numerous governmental agencies including NASA, the Federal Aviation Administration (FAA), the Department of Defense, and the Department of Transportation have created a partnership known as the Joint Planning and Development Office (JPDO). One of the major initiatives of the JPDO is to integrate a more sophisticated satellite surveillance approach that will no longer rely on traditional radar-based systems, referred to as Trajectory-Based Operations (TBO) (Battiste & Johnson, 2008).

The following experiment assessed the performance of pilots using advanced automation tools that are recommended for a TBO environment (FDDRL, 2008). Pilots were asked to complete flight plan modification exercises and were measured on decision time, relative efficiency, behavioral trends, and workload assessment. Along with the assessment of these advanced automation tools, this research also evaluates the effects of feedback delays within those systems. Inherent feedback delays exist within modern communication systems due to computational inadequacies and increased distances between satellites. Significant delays are believed to be potentially hazardous within a communication system that relies on rapid response rates. In order to accommodate the future needs of the NAS there must be a high level of precision when selecting trajectories, and, as air traffic grows, tools must also be more precise. Without the
successful design and implementation of these automation tools, the current ATM system will not be able to sustain this anticipated demand (Prevot et. al. 2008).

1.2 Trajectory-Based Operations

NASA and the FAA as well as other agencies are anticipating that a double or triple increase in air traffic demand will occur by the year 2025. Accommodating this demand will require considerable modernizations to the NAS. The Next Generation Air Transportation System (NextGen) conceived by the JPDO, proposes a shift away from traditional radar-dependent clearance-based traffic management to satellite surveillance-dependent TBO to accommodate this need (FAA, 2008). TBO designates four-dimensional virtual trajectories (3D space & time) to aircraft through the use of numerous advanced technologies, and is applied primarily to “en-route” portions of flight routes. The TBO concept requires aircraft to travel along negotiated routes as opposed to clearance-based flight that directs aircraft toward fixed waypoints in the sky (Figure 1). Pilots are then able to fly an optimized flight route by relying on aircraft separation automation, which assists in the avoidance of violating separation rules with neighboring aircraft. The shift towards TBO involves reallocating roles and responsibilities of pilots and ATC, as well as a considerable dependence on advanced automation tools, including Cockpit Situation Displays (CSD) and Route Assessment Tools (RAT) that allow for the continuous manipulation of trajectories (Johnson & Battiste, 2005). These tools assist pilots and controllers in deciding which trajectories to fly by calculating present aircraft position and as well as a future position in time.
The goal of commercial pilots is to safely navigate aircraft in an efficient manner while simultaneously providing passenger comfort. There is a definitive hierarchy in the priority of all of their objectives. Pilots differ from ATC in that they are highly motivated to conserve fuel and associated fuel costs, whereas ATC prioritize the general safety of all aircraft in their given sector. Additionally, studies have shown that ATC may be unable to safely direct air traffic densities that exceed a 33% increase of current day levels (Prevot et. al 2008) (Dwyer & Landry, 2009). Through 4D TBO, pilots and ATC are able to negotiate precise trajectories for aircraft, thus eliminating much of the difficulty in maintaining separation in high traffic airspace (FAA, 2007). An essential component of TBO is a shift from pilot-to-ATC voice communications to a digital data-link communication system. This enables the rapid sharing of detailed 4D information such as an aircraft’s current and future locations, speed, and altitude between controllers, pilots, and other aircraft (Erzberger, 2002). A primary example of a data-link enabled tool
is the CSD (Figure 2). The CSD is currently being researched and developed at NASA Ames’ Flight Deck Display Research Lab (FDDRL). The CSD is a type of Cockpit Display of Traffic Information (CDTI) that allows pilots to visualize as well as control active trajectories in real time. The current version of CSD is a high-fidelity graphical user interface that provides information such as data-linked traffic, weather and terrain information. CSD can be operated in a traditional 2D view, and a volumetric 3D/4D perspective.

![Figure 2. Cockpit Display of Traffic Information](image)

1.3 Conflict Detection and Resolution System
CSD allows users to view the traffic around them, and it equips them with Conflict Detection & Resolution (CD&R) capability, based on an automated resolution logic of the Advanced Airspace Concept (AAC) algorithm developed by Dr. Heinz Erzberger of NASA Ames (Erzberger, 2002). The AAC algorithm is able to generate countless conflict-free and efficient trajectories while accounting for the trajectories of surrounding
aircraft. CD&R is primarily an alerting function built into the RAT that pilots can use to modify their 4-D trajectory (Figure 3).

![Route Assessment Tool (RAT)](image)

**Figure 3.** Using the Route Assessment Tool (RAT) to avoid a conflict.

The CD&R alerts users of an impending conflict, or “loss of separation” between the user’s aircraft and an intruder aircraft. The user is notified if loss of separation will occur when their aircraft and the conflicting aircraft change in color from magenta to amber on the CSD (Figure 4).
1.4 Route Analysis Tool
The Route Analysis Tool (RAT) is a trial planner that enables pilots and controllers to search and discover conflict-free trajectories. Through the recent adoption of rapid satellite surveillance (e.g., Automatic Dependent Surveillance Broadcast) people conducting route discovery tasks may assume real-time update information. Pilots’ and ATCs’ use of advanced trial planning tools equipped with CD&R is the proposed method for maintaining efficient and safe flight plans (Johnson, 2003 and Prevot, 2008).

The RAT is the proposed trial planner tool developed by NASA Ames researchers. The RAT allows pilots to interactively manipulate flight plans in order to find safer and more efficient routes (FDDRL, 2008). Previous studies have shown the RAT to be an effective instrument for avoiding LOS and executing more direct flight plans (Johnson & Battiste, 2005 & Johnson, 2003). It enables users to maintain the safety of a flight through the assistance of a primary CD&R scan. When a conflict arises pilots can scan nearby airspace to find a safe alternative trajectory. The safety of a potential route is gauged via a secondary CD&R scan which alerts users of any conflicts along a potential new trajectory. Varying on how flight rules are distributed in the future, once pilots commit to a new route they may be required to either data-link ATC those flight plan changes for verification or independently execute the new route change and continue with their flight.
However, advanced automation tools such as these are subject to performance instability, due to computational feedback delays, which may significantly affect successful task completion (Bryson, 1993).

1.5 Feedback Delays
The literature discussed here includes various psychological theories, experimental findings, and expert opinions that have contributed to the research of visual feedback delays in human-computer systems. There has been virtually no previous research conducted to evaluate the effects of delays in the NextGen cockpit. However, it has been well established that feedback delays impair performance in a number of human-machine interactions (Kim et. al, 2004; Sheridan, 1963; Davis et. al, 2009; MacKenzie & Ware, 1993).

The main areas of study that are pertinent in the area of feedback delays within automation tools are as follows:

- Human performance effects of visual feedback delays
- User tolerance of feedback delays in computer systems
- Behavioral changes when exposed to feedback delays

1.5.1 Effects of Delays on Performance
Since there is no previous research on the effects of feedback delays during route modification tasks, we can only look to findings that have characterized the general and frequently encountered properties of visual feedback delays.

Robert Miller discusses the effects of delayed response times in human-computer communication loops. Miller establishes that human performance does not degrade linearly as delays increase, and that people have inherent psychological “step-down”
points (Miller, 1968). Step-down points are defined as lengths of visual feedback delay where people show a significant change in behavior. A drop in performance for a task requiring a rapid response typically characterizes a step-down point. There are three general step-down points identified by Miller and Nielsen, and they are said to occur at approximately 100 ms, 1 second, and 10 seconds (Nielsen, 1993). Delays that are longer than one second are easily noticeable, but there does not seem to be an interruption in users’ flow of thought if maintained below this level. A ten second delay is said to be the longest amount of delay acceptable before users lose focus and move onto other tasks.

Miller’s review of response times in human-computer conversational transactions also discusses major psychological phenomena that occur when people are forced to wait. People have expectations when communicating with individuals as well as computers, and delays in response time during these communication loops create interruptions, which often result in undesirable effects. The concept of “closure” is discussed as a cognitive need to mentally connect pieces of information into autonomous wholes in order to move on to something else. A delay in closure will make it difficult to advance from one task to another, and a common result is severe deterioration of human memory due to elapsed time. Closure is essential to maintaining effective communication and conducting efficient task-dependent problem solving. People group activities, information, or concepts into “clumps” that are defined by a purpose or sub-purpose, and they typically expect a “clump” to be completed before they can address the next clump. Miller states that the longer people store incomplete clumps in their short-term memory the more apt they are to forgetting or committing memory retrieval errors. Miller goes on to state that during complex problem solving short-term memory is heavily inundated
with information, and that closure within a timely manner is required to purge peoples’ short-term memory store. Other research on feedback delays suggests that it also leads to overcompensation in the user’s movements, lack of trust in the system, and disorientation and confusion of the user (Day, 1999).

1.5.2 Delays in Tracking Tasks
There is a considerable amount of research that has measured the negative effects of feedback delays on tracking and continuous movement task performance, e.g., operating aircraft, virtual reality and remote-surgery (Bryson, 1993 & Kim, 2005). This research suggests that human performance significantly breaks down in tracking and continuous movement tasks at approximately 200-300 ms. (3 to 5 frames per second) of feedback delay. This is the point at which objects no longer appear to be continuously in motion. Findings like these support the belief that visual feedback delays do indeed degrade performance and increase frustration in human-computer interactions.

Some research suggests that feedback delays should not always be circumvented, and that there are benefits to having longer wait times. Schooler and Anderson (1990) found that students being trained to learn a computer programming language benefited by having delays in response feedback. The results indicate that students in the immediate feedback condition completed problems faster, but with a higher frequency of errors compared to students who received feedback delays. They argued that longer wait times are attributable to more conscious and calculated decision-making. The RAT is a very highly responsive tool when operating free of feedback delays. It enables users to search for acceptable routes without the need of any significantly high-level thinking. It can be argued that in the presence of delays during route modification tasks, users can capitalize
on their extended waiting periods, thus making fewer failed attempts and engaging in a more deliberate route searching behavior.

1.5.3 User Tolerance of Feedback Delays
The following research was conducted mainly within the field of website usability. Researchers have performed numerous experiments in an attempt to identify acceptable webpage loading times. Lightner and Bose (1996) found that wait time is the most undesirable feature encountered when browsing the web, and Rose (2001) states that users will abandon a site if wait times are too long. These findings may help predict pilot behavior when faced with extreme levels of feedback delay. Users may perhaps abandon the use of the CD&R tool, and resort to manually identifying safe coordinates to fly towards. Additional research by Nielsen (1999) claims that two seconds is the “gold standard” when it comes to wait time for webpage loading, and anything beyond ten seconds usually leads to abandonment.

Miller (1960) discusses numerous strategies that people utilize when faced with a plethora of options, or when they may be overloaded with information. The filtration strategy allows people to analyze existing options and discard the information that seems unimportant, irrelevant, or nonessential. The next strategy is acceleration where people increase their cognitive analysis of the remaining options to make a decision as quickly as possible. Omission is the third strategy, and it is a way for people to disregard options that are perceived to be unrelated or worthy of abandonment. In the context of this study, a user could be overloaded with information due to an unexpected occurrence of feedback delay. The expected result would be for users to discard the inconsistent CD&R feedback information and use a more reliable and definitive method to avoid “loss of separation.”
Some researchers have associated increases in feedback delays with decreases in steps required to solve problems and increases in efficiency (Yntema, 1968). These results can be found in numerous studies that show delays in response time consistently lead to increases in errors, frustration, and time required to complete tasks (Goodman and Spence, 1978; Thadani, 1981). In a time sensitive task such as route modification the number of steps taken is not as significant as the amount of time required to commit to a safe and efficient trajectory.

1.6 Human Memory Systems

Human memory is a very complex and intricate system that is heavily researched by the international psychological community in order to understand how it operates. A brief explanation of short-term and iconic sensory memory is essential in understanding why human performance is believed to degrade when exposed to feedback delays.

Memory has been theorized to include long-term memory, short-term memory (working memory), and sensory memory. These three distinctly different memory stores all exist and operate independently to allow people to retain and recall important information. Working memory is the most relevant memory system in question for this research. Working memory is believed to be very limited in storage capacity and storage time. A component of working memory used in the temporary storage of visual and spatial information is the visuospatial sketchpad. The visuospatial sketchpad retains colors, shapes, motion, velocities, and other visual information in order to assist in humans’ ability to react to stimuli. It is an individual’s short-term memory store of spatial and visual information, and it is involved in tasks that require the planning of spatial movements (Baddeley, 2004).
Working memory is very limited in capacity and duration, being that it can only store a few autonomous items for only a few seconds. However, this capacity is heavily dependent on the complexity and familiarity of the items being remembered (Alvarez & Cavanaugh, 2003).

1.6.1 Relevance to Pilot Performance
As a pilot receives visual stimuli from the CSD they briefly store this information in their working memory in order for them to quickly formulate a plan of action as well as assist them in their decision making process. They typically continue to manipulate the CSD until they can identify an optimal route change. Once they find and execute a satisfactory trajectory they experience closure and are able to purge their working memory store. Their working memory store could potentially consist of a number of items such as areas of the display that were scanned and deemed unsafe, altitude and speed of surrounding aircraft, as well as numerous other data a pilot is trained to identify. The longer a pilot unsuccessfully searches for potential routes the more memory decay or memory replacement effect may take place. This can result in decreased task performance due to capacity limitations of working memory. Theoretically, if it were not for working memory, tasks would never reach completion because the user would constantly repeat
actions due to having forgotten receiving feedback, and perhaps even forgetting the main task at hand.

1.7 User Behavior and Strategy Adoption
Research in the area of human control and problem solving has brought forth numerous ideas about how humans respond and react to stimuli, and one popular example is the concept of TOTE units developed by Miller et. al (1960). TOTE stands for “Test-Operate-Test-Exit”, and it is the idea that people will proceed through a specific problem-solving loop to reach a goal. They initially “Test” a system to observe its current status, then “Operate” to create a change, followed by another “Test” to measure the effect of that change, and this loop will continue until the user has reached their goal at which point they will “Exit” the command loop.

An example of a TOTE unit observed in this study would be a user recognizing that ownership is in conflict (Test), then using the RAT to find an alternative trajectory (Operate), followed by observing to see if ownership is still in conflict (Test), and finally executing the route change if the user is satisfied with the result (Exit).

A study conducted by Teal and Rudnicky (1992) attempted to identify how system delays affect users’ selection of task strategy when engaging in keystroke tasks. They monitored user strategy selection when faced with feedback delays to understand the relationship between delay and user performance. Their results indicated that users tend to choose strategies that are best suited to system feedback delays.

It is expected that pilots will adopt numerous behavioral strategies when faced with delayed feedback of the RAT. A “move-and-wait” strategy has been seen in tele-operations, lunar rovers, and numerous other systems that have been experimentally
impacted with control delays (Day, 1999, MacKenzie and Ware, 1993; Kim 2005). Meticulous movements that are followed by waiting periods characterize a “move-and-wait” strategy. These waiting periods allow for the system to provide the operator with feedback. Once feedback is received the operator will once again make a calculated movement and then proceed to wait for a response. This strategy typically results in longer task completion times, however it diminishes the occurrence of delay induced user error. Another coping strategy that is expected to emerge as a result of delays is the execution of excessively safe route modifications. Pilots are expected to take more exaggerated and longer routes when searching for safe trajectories. This is especially true under high time-pressure scenarios where users need to respond rapidly, thus “playing it safe” and moving the RAT to a more unmistakably safe position. This adapted behavior could lead to the generation of safe yet highly inefficient flight routes.

1.8 Task Difficulty and Time Pressure
Feedback delays have proven to increase task difficulty and hinder human performance in a number of scenarios, but the added effect of time pressure is an important additional factor to consider when studying behavior under feedback delays. Wickens et. al, (1998) discusses how the complexity of a task has a direct effect on how much time an individual will need to complete that task. This is mainly due to an increase in cognitive processing required to make a desirable decision. This relates back to TOTE units, in that in a more complex task a higher frequency of Test-Operate-Test loops will occur in order to eventually exit the task.

Route modification tasks are fairly unique in nature because they require a high level of situational awareness, expansive technical knowledge and an ability to make crucial
decisions in a fairly limited amount of time. In addition, the inherent complexity of route modification tasks can be amplified due to a number of variables, e.g., increased time pressure, traffic density, number of conflicts, inadequate training, etc. NextGen concepts propose that pilots take on some of the tasks that controllers currently engage in, with the added assistance of advanced automation tools. The current system requires that controllers be able to rapidly respond to potential conflicts in order to address the following conflict and so forth. Since pilots primarily account for their own aircraft they may be allotted a few extra moments to consider the most efficient route modifications available. Thus, by creating a realistic experimental scenario, with the inclusion of fluctuating feedback delays, the full spectrum of effects under feedback delays can be identified.

1.9 Traffic Density
As highlighted in the above section, air traffic is expected to double or triple by 2025, and these increases will have a direct effect on pilots’ ability to safely and efficiently navigate aircraft (Prevot et. al. 2008). It is reasonable to expect the amount of available route options to be dependent on the number of proximal aircraft. Even under current-day levels traffic density drastically increases near or around airports. Available routes may decrease as the number of surrounding aircraft increases due to relatively scarce airspace. High traffic levels may result in pilots’ execution of undesirable deviations from the most efficient route in order to maintain safety. It can also be expected that pilots may require additional time to identify less obvious routes that are more efficient when flying amidst larger numbers of surrounding aircraft. Thus, when using the RAT and CD&R pilots may discover that under some circumstances there may be limited desirable route options.
Addressing the issue of varying traffic density is key to the successful design and implementation of the CSD. Additionally, the effect of reduced solution space on pilots’ ability to find safe and efficient routes may be exaggerated when there are lengthy feedback delays present in their navigation tools. It is crucial to identify whether there is an interaction between traffic density levels and amounts of feedback delay present in the conflict detection capability of the CSD. Understanding this relationship will allow system developers to mitigate potential risks associated with using advanced automation tools in high traffic airspace.

Pilots’ unawareness of available routes or solution spaces as well as the locations of potential conflicts is an important aspect to consider when evaluating the effects of feedback delays. When confronted with a conflict pilots may have numerous solution spaces available (illustrated in white) from which to choose (Figure 6). However, locating these spaces becomes a great challenge when systems fail to rapidly update their locations. A pilot is likely to receive a partial “image” of their environment while sweeping for potential safe routes (Figure 7).

Scenarios such as these may result in the masking of preferred or even acceptable route modifications. This effect can be described as a loss in situation awareness because pilots may be unaware of their surrounding environment, and they may not have adequate information to navigate safely nor efficiently. Therefore, in order for pilots to avoid LOS they may resort to extreme and often inefficient route modifications. It should be noted that this is based on the assumption that users will sweep for routes at equivalent rates. For users to be able to detect the same available solution spaces from a non-delayed condition they may be required to sweep at a significantly slower rate of speed. This is to
allow for pilots subjected to prolonged CD&R detection loops (feedback delays) to conduct an equally thorough scan of airspace. Another related expected result is for pilots to encounter a false-positive scenario where they may create a route modification that seems free of conflicts, but once they attempt to execute this proposed route the CD&R loop will complete its cycle resulting in conflict detection. This occurrence may result in a pilot either executing a conflicted, and potentially hazardous route, or having to return to the resolution search part of the task.
Figure 6. Available solution spaces with no delay.

Figure 7. Limited solution space due to delays.
Chapter 2  Statement of the Problem

2.1 Introduction
Current aircraft separation research conducted at NASA Ames Research Center has yet to address the effect of feedback delays on users. Pilot performance under various lengths of feedback delay must be evaluated prior to real world implementation to ensure the safe and efficient use of these systems. Feedback delays are present in numerous forms, and are a considerably disadvantageous factor within human-in-the-loop systems. They are often attributable to inadequately slow computer processing, sudden overloads in graphic rendering, high densities of air traffic and increased physical distances between communication satellites. The type of feedback delay addressed in this study is specifically computational feedback delay of the CD&R feature of the RAT. These computational feedback delay effects may be exacerbated in airspace that contains a higher level of traffic. Having numerous aircraft in the surrounding environment requires additional computing power and may often result in mild-to-severe system degradation.

The CSD is used to create conflict resolutions when conflicts have been brought to the attention of pilots. If the conflict detection feedback is delayed it may lead to users’ inability to maintain safe separation or do so efficiently.

2.2 Objectives
This research aims to identify the effects of computational feedback delay within conflict detection and resolution tasks on pilots using the NASA Ames CSD equipped with the RAT and CD&R system. There was also a strong interest in identifying strategies that pilots may adopt when faced with increasing levels of feedback delay within varying traffic densities. Results from this study will be used to influence design specifications for CD&R functions in CSDs.
2.3 Hypotheses

The term “performance” in the following hypotheses is used to identify decision time of as well as route distance measured in nautical miles. Performance is defined as both flight route distance as well as decision time since these two factors should be positively correlated. Route distance is used as a proxy for fuel efficiency. The primary hypotheses are:

H1. Feedback delays of two seconds or longer, in the CD&R function of the RAT, will result in a decrease in pilot performance regardless of traffic density.

H2. Pilot performance will be negatively affected in high-density traffic regardless of feedback delay length.

H3. There will be a greater detriment on performance when pilots are exposed to feedback delays in high-traffic density trials compared to low-traffic density trials.

H4. Participants will adopt various coping strategies, such as “Move and Wait”, when exposed to feedback delays.
Chapter 3  Methods

3.1 Participants
The experiment was performed on 12 pilots, 10 male and 2 female who were recruited by NASA Ames Research Center. All participants were United States citizens, and were paid $50 for approximately two hours of their time. None of the participants had previously used the NASA Ames CSD or any other CDTI system. This study operated under an existing IRB approval granted to NASA FDDRL.

3.2 Facility
The experiment was conducted in the FDDRL at NASA Ames Research Center in Mountain View, CA.

3.3 Materials
Participants interacted with a 27” LCD, QWERTY keyboard and optical mouse. The LCD presented the CSD software equipped with the RAT. With assistance from NASA Ames software engineers and researchers a customized version of the FDDRL CSD was used for this study. This experiment limited participants to an overhead 2D viewing perspective with viewable range set to a 160 nm radius.

Twelve traffic scenarios were adopted from a previous NASA experiment conducted in 2008. From those twelve trials a total of 64 unique trials were derived through the assignment of new aircraft as ownship. Forty of these trials were used for recording data and 24 were used for training purposes. These trials varied in altitude, number of conflicts, conflict heading, speed, traffic density, and overall difficulty. All 64 trials contained approximately 2x to 3x current-day traffic levels. Pilots were limited to solely creating horizontal route changes in order to effectively record efficiency in nautical miles. Camtasia Studio software was used to record data from the LCD monitor for later
analysis. During training a CSD training manual was provided to instruct participants about the relevant functions of the CSD. A post-trial questionnaire was administered via paper and pencil after each of the 40 trials and a post-experiment interview was conducted to assess user opinions.

3.4 Experimental Design and Procedures
This experiment was carried out as a within participants design, with the independent variables being feedback delay (0 seconds, 2 second, 4 seconds, 8 seconds) and traffic density (low-density and high-density). Participants engaged in training for this experiment to ensure that they were competent in using the CSD, CD&R, and RAT, and that they were able to make safe and efficient route modifications in a timely manner. Participants began the experiment once they successfully demonstrated the established criteria.

Once participants completed training, demonstrated a strong command of the tools, and completed all comprehension tests they were allowed to continue with the experiment. Participants were instructed to maintain safe separation throughout 40 unique flight scenarios while prioritizing passenger safety, route efficiency and completion time, respectively. Trials began with own-ship in conflict with one or more intruder aircraft. Trials also began with a 15-second scenario evaluation period to orient pilots with the proceeding air traffic scenario. After 15 seconds had elapsed the RAT was activated and the trial began. Once the RAT was activated participants were required to resolve the conflict(s). Trials were limited to a maximum time of two minutes after which time pilots were asked to complete a brief questionnaire. The entire experiment was approximately 1.5 to 2 hours in length, depending on the amount of training required and length of intermissions requested on behalf of pilots. All participants were debriefed at the close of
the experiment.

3.5 Independent Variables

3.5.1 Feedback Delay

Feedback delay intervals were 0 ms (0 Hz), 2 seconds (0.5 Hz), 4 seconds (0.25 Hz), and 8 seconds (0.125 Hz). Participants were exposed to these four levels of feedback delay in eight counterbalanced blocks of five trials to eliminate ordering effects. The two-second feedback delay level was an interval found to be a significant performance drop off point in previous studies of feedback delay. Additionally, measurements made during two pilot studies demonstrated a noticeable increase in time on task and route distance at feedback delay levels of four seconds or longer. Feedback delay was doubled to eight seconds to establish additional manipulations and was expected to push the boundaries of what pilots would be willing to tolerate during a simulation such as this.

There were specific details in the way in which delay was implemented in this study. The CD&R system contains an inherent amount of delay between route testing loops. This delay is estimated to be within 50 ms and 400 ms. Ideally any amount of delay would be a concern in a control condition, however, this inherent feedback delay is hardly perceptible and considered negligible. When feedback delays are set to 0.5 Hz or 0.25 Hz the CD&R will test a proposed route for conflict once every 2 or 4 seconds respectively. However, the conflict testing cycle is on an autonomous loop meaning it is not initiated with any action on behalf of the user. Thus the 2-second condition (0.5 Hz) may result in delays that range from 0 seconds and 2 seconds, and the 4-second (0.25 Hz) condition may result in delays that range from 0 ms and 4 seconds.

The behavioral changes between 2 seconds and 4 seconds was of significant interest due to the consistency of findings that suggest feedback delays within communication loops
significantly degrades performance at approximately 2 seconds (Miller, 1968). A delay of 4 seconds was predicted to result in significant behavioral differences such as increases in workload, increases in time required to complete a route modification, and decreases in route efficiency. Delays of 4 seconds or longer were also expected to encourage pilots to abandon the automation tool and to use higher-level cognitive skills to manually maintain safe separation. Through these multiple feedback delay manipulations a range of behaviors were identified and categorized.

3.5.2 Traffic Density
Participants were asked to navigate through 40 unique trials. Trials varied in numerous ways, but there were specifically two levels of traffic density that were controlled for. Trials were differentiated into two groups; low-density and high-density. These groupings were made quantitatively by assessing the proportion of airspace viable for a safe route modification. The amount of available solution space within a 160 mile diameter was measured horizontally from the point of initial conflict. The same measurement was made 20 miles ahead of initial conflict and 20 miles behind. A trial was automatically classified as low-density if it had a solution space greater than 5 miles available within a 40-mile diameter from the initial point of conflict. Additionally, a trial was classified as high-density if it had less than 20 miles of available solution space within an 80-mile diameter from the initial point of conflict.

3.5.3 Counterbalance Methodology
Forty trials were carefully selected from the original 64 to ensure consistency and to eliminate problematic or unusual flight scenarios. Trials were then grouped based on traffic density. The result was 20 trials of low-density traffic and 20 trials of high-density traffic. Trials were then placed into blocks of five trials of the same traffic density.
Feedback delay amounts were assigned to each block to ensure that each pilot saw each block only once. Every pilot received an equal amount of low and high-density blocks at each of four levels of feedback delay in random order. However, all participants completed the low and high-density 8-second blocks last to eliminate any expected carry-over effects due to severely long feedback delays.

3.6 Dependent Variables
To assess the effect of independent variables on pilot performance (i.e., efficiency and safety) the following quantitative measures were used:

1. The total distance of a modified route in order to compare the efficiency of route modifications across conditions. Maintaining more direct routes is a definitive way to conserve fuel consumption and save travel time.

2. Total time required to execute a route modification was measured. The time needed to execute a route is crucial because as aircraft near intruder aircraft, there are fewer conflict-free routes available. Additionally, route modification tasks are expected to be a small portion of the general duties pilots undertake while flying, and an excessive amount of time to complete them could have a negative effect on their other various tasks. Also, considerable increases in completion time may compound over the course of a flight leading to unsatisfactory travel delays.

3. The success rate of conflicts solved was measured to identify pilots’ ability in finding safe alternative routes throughout various experimental conditions.

4. The number of timed-out trials was also measured in order to identify pilots’ inability to successfully complete a trial within the allotted time of 2 minutes.

In addition to the previous numerical performance measurements, the following qualitative measures were analyzed:
1. Trends in behavior that emerged due to various levels of feedback delays were identified. Anticipated behaviors were reported on including “Move and Wait” or pilots’ use of multiple articulations, or flight legs in their route modifications.

2. A Likert rating scale was used to evaluate overall difficulty, overall frustration, perceived available solution space and perceived feedback delay. Values ranged from “very easy/very low” (1) to very difficult/very high (7). Pilots completed a questionnaire after each of the forty trials.

3. A debriefing interview was administered at the end of the experiment to gather suggestions, evaluate users’ trust in the automation, discuss pilots’ mental processes during the experiment, and document pilots’ opinions on the real world application of these tools.

3.7 Data Analysis
The data analysis followed numerous steps and consisted of various sets of data. This section is divided into four parts: (1) quantitative data analysis, (2) questionnaire data analysis, (3) video analysis, and (4) debrief interview analysis. Each type of data was collected to help illustrate a more complete picture of the effects of feedback delays and traffic densities on pilot performance.

3.7.1 Quantitative Data Analysis
Route distance (nm) and Decision Time (sec) were chosen as dependent measures of pilot performance. This data was compiled and analyzed using a repeated measures two-way ANOVA, also known as a factorial ANOVA, due to the presence of two independent factors (feedback delay and traffic density) as well as two dependent factors (route distance and decision time). The statistical analysis was a repeated measure ANOVA because all participants were given all eight conditions. The statistical significance of
main effects for both independent factors was analyzed as well as any significant interaction between independent factors. An alpha level of .05 was used for all statistical tests.

3.7.2 Questionnaire Data Analysis

Questionnaire data was collected at the end of each trial. The four questions were as follows:

- Rate the overall difficulty of this trial. (1=Very Easy; 7=Very Difficult)
- Rate your performance on this trial. (1=Very Low; 7=Very High)
- What was the approximate feedback delay level of this trial? (1=Very Low; 7=Very High)
- What was the traffic density level of this trial? (1=Very Low; 7=Very High)

This data was also analyzed with a repeated measures two-way ANOVA.

3.7.3 Video Analysis

Participants’ screens were recorded with Camtasia software and the resulting videos were later viewed to identify noticeable trends in behavior, failed trials, as well as any software issues which may have occurred. Videos were coded blind of condition to mitigate any bias in identifying behaviors in higher delay or higher traffic density conditions. The main behaviors identified were participants’ use of multiple articulations to modify their route, the presence of a “Move and Wait” approach to search for safe routes, any inability for pilots to resolve a conflict, the execution of conflicted routes, as well as the near execution of conflicted routes.

3.7.4 Debrief Interview Analysis

Participants were asked a series of questions at the end of the experiment to assess their general impressions of the tool as well as their opinions about the feedback delays they
encountered. Participants were generally asked a standard set of questions as well as additional ad-hoc probing questions. The audio from these interviews was recorded and later analyzed for common trends and insights that may have emerged from participants’ responses.
4.1 Decision Time

Participants were able to complete route modifications in the shortest overall time in the zero feedback delay condition, $M=26.2$ sec. Results indicated that decision time steadily increased with longer feedback delays. Time increased by more than 30% when feedback delay was set to only two seconds, $M=34.3$ sec. Time continued to increase when feedback delays were increased beyond four seconds. The additional increase in decision time for the 4-second condition was 15% greater than in the 2-second condition, $M=39.35$ sec, and the increase in time in the 8-second condition was approximately 6% longer than the 4-second condition, $M=41.89$ sec. Time increased a total of 60% between the zero and 8-second conditions. Data was also analyzed with a repeated measures two-way ANOVA and the results suggest a statistically significant main effect for Feedback Delay, $F(3,33)=10.33$, $p = 0.00$.

Pilots were also able to create safe routes more rapidly in the low-traffic density condition, $M=28.75$ sec., in comparison to the high-traffic density condition, $M=42.13$ sec. This represented a 46% increase in decision time between the low and high-density condition. Further analysis again suggested a significant main effect for Traffic Density, $F(1,11)=29.48$, $p = 0.00$.

Results indicated there was a marginally significant interaction suggesting that the effect of feedback delay was greater with high-density traffic than low-density traffic, $F(3,33)=2.476$, $p= 0.08$. In the high-density condition, decision times appeared to jump from $M=37.1$, $SD=13.25$ in the two-second delay condition to $M=47.7$, $SD=14.88$ in the 4-second delay condition. Trials were longest on average in the 8-second delay
condition, $M=50.46$, $SD=11.17$, and shortest in the 0-second delay condition, $M=33.2$, $SD=16.18$ (Figure 8).

When observing performance in the low-density condition there appeared to be more consistency in performance despite the amount of delay. However, decision time was noticeably lower when participants were faced with trials in the 0-second feedback delay condition, $M=19.2$, $SD=9.11$. 
Figure 8. Mean averages of Decision Time (Feedback Delay x Traffic Density)

![Decision Time (sec) diagram with bar charts showing Mean averages for Low Density and High Density across Feedback Delays of 0 sec, 2 sec, 4 sec, and 8 sec.](image)

<table>
<thead>
<tr>
<th>Feedback Delay (sec)</th>
<th>Low Density</th>
<th>High Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 sec</td>
<td>19.2</td>
<td>33.2</td>
</tr>
<tr>
<td>2 sec</td>
<td>31.5</td>
<td>37.1</td>
</tr>
<tr>
<td>4 sec</td>
<td>31</td>
<td>47.7</td>
</tr>
<tr>
<td>8 sec</td>
<td>33.3</td>
<td>50.46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feedback Delay (sec)</th>
<th>Low Density</th>
<th>High Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 sec</td>
<td>2 sec</td>
<td>4 sec</td>
</tr>
<tr>
<td>Mean</td>
<td>19.2</td>
<td>31.5</td>
</tr>
<tr>
<td>Std. Error</td>
<td>2.62</td>
<td>4.18</td>
</tr>
</tbody>
</table>

Table 1. Summary Data of Decision Time
4.2 Route Distance

Route distance data was analyzed with an identical approach as decision time. The average distance of a route completed in the 0-second condition was $M=11.25$ nm. This amount increased by 50% when feedback delay was set to two seconds, $M=16.90$ nm. Trials in the 4-second condition resulted in route distances that were approximately 29% longer than the 2-second condition, $M=21.80$ nm. The average length of a modified route in the 8-second condition was $M=20.35$ nm. ANOVA results indicated there was a statistically significant main effect for the feedback delay factor, $F(3,33)=3.94$, $p=0.02$.

Traffic Density also appeared to have a significant main effect on route distance, $F(1,11)=29.72$, $p=0.00$. The average distance of a route in the low-density condition was $M=11.02$ nm. The average distance of a route in the high-density condition was approximately 120% longer than the average route in the low-density condition, $M=24.12$ nm.

Results also indicated there was a marginally significant interaction suggesting that the effect of feedback delay was greater with high-density traffic than low-density traffic, $F(3,33)=2.58$, $p=0.07$. Pilots’ performance varied drastically between the low-density and high-density conditions. In the high-density condition there seemed to be a sharp increase in route distance between the 2-second delay condition, $M=21.30$, SD=9.45, and the 4-second delay condition, $M=32.09$, SD=20.50. Route distance was shortest in the 0-second delay condition, $M=16.12$, SD=13.39. There appeared to be a slight decrease in the 8-second condition, $M=26.97$, SD=10.49 (Figure 9).

The results from the low-density trials had a different pattern than the high-density trials addressed above. There seemed to be no drastic change in performance between the 2, 4
or 8-second delay conditions. However, pilots were able to complete considerably shorter route modifications in the 0-second feedback delay condition, $M=6.38, SD=4.13$ (Figure 8).

![Route Distance (nm)](image_url)

Figure 9. Mean Averages for Route Distance (Feedback Delay x Decision Time)

<table>
<thead>
<tr>
<th></th>
<th>Low Density</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 sec</td>
<td>2 sec</td>
<td>4 sec</td>
<td>8 sec</td>
<td>0 sec</td>
<td>2 sec</td>
<td>4 sec</td>
<td>8 sec</td>
<td>0 sec</td>
<td>2 sec</td>
<td>4 sec</td>
<td>8 sec</td>
</tr>
<tr>
<td>Mean</td>
<td>6.38</td>
<td>12.49</td>
<td>11.50</td>
<td>13.73</td>
<td>16.12</td>
<td>21.3</td>
<td>32.09</td>
<td>26.97</td>
<td>32.09</td>
<td>26.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. Error</td>
<td>1.20</td>
<td>2.76</td>
<td>2.77</td>
<td>3.32</td>
<td>3.87</td>
<td>2.73</td>
<td>5.92</td>
<td>3.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Summary Data of Route Distance
4.3 Failed Trials
Pilots were required to create a conflict-free route modification while doing so in less than 2 minutes. They were also required to only execute conflict-free routes. If pilots were unable to meet either of these two criteria the trial was deemed a failure. Out of the 480 total trials recorded 25 resulted in either type of failure. Eleven of these failures were timed-out trials and 14 were conflict executions.

Seven of the eleven timed-out trials occurred in the 8-second condition. The two-second and four-second conditions had 2 timed-out trials each, and the zero-second condition had no timeouts. Additionally, 9 out of 11 timed-out trials occurred in the high-density condition, including 5 trials in the 8-second delay/high-density condition (Figure 10).

![Timed-out Trials](image)

Figure 10. Timed-out Trials
There appeared to be a large concentration of conflict execution failures in the 8-second delay condition. Out of the 14 total conflict execution failures eight of them occurred in the longest delay condition. There were also four failures in the 4-second delay condition and two in the 2-second delay condition. There were no failures in the zero-second delay condition (Figure 10).

![Figure 11. Conflict Executions](image)

Figure 11. Conflict Executions
4.4 Questionnaire Data
Pilots were asked to answer four questions at the end of each trial in order to measure changes in their perception. There were each presented as a one to seven Likert scale.

The questions were as follows:

- Rate the overall difficulty of this trial. (1=Very Easy; 7=Very Difficult)
- Rate your performance on this trial. (1=Very Low; 7=Very High)
- What was the approximate feedback delay level of this trial? (Very Low=1; Very High=7)
- What was the level of traffic density in this trial? (1=Very Low; 7=Very High)

4.4.1 Perceived Difficulty
Questionnaire data suggests that pilots perceived trials to be more difficult as feedback delays increased, \(F(3, 33)= 18.073, p>.05\). Pilots were not given information as to when feedback delays changed. They were unaware of the number of intervals of feedback delay. The data also showed that pilots deemed trials to be more difficulty when traffic density was high, \(F(1, 11)= 13.11, p>.05\). There appeared to be a marginally significant interaction suggesting that the effect of feedback delay on perceived difficulty was greater with high density than lower density trials, \(F(3, 33)= 2.56, p=.07\). Pilots perceived trials to be more difficult as feedback delay increased, but their perception of difficulty was exaggerated in higher traffic density.
Figure 12. Perceived Difficulty

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low_0</td>
<td>2.55</td>
<td>0.80</td>
</tr>
<tr>
<td>High_0</td>
<td>2.78</td>
<td>0.70</td>
</tr>
<tr>
<td>Low_2</td>
<td>2.75</td>
<td>1.29</td>
</tr>
<tr>
<td>High_2</td>
<td>3.18</td>
<td>0.91</td>
</tr>
<tr>
<td>Low_4</td>
<td>2.82</td>
<td>1.14</td>
</tr>
<tr>
<td>High_4</td>
<td>3.95</td>
<td>0.80</td>
</tr>
<tr>
<td>Low_8</td>
<td>3.53</td>
<td>1.15</td>
</tr>
<tr>
<td>High_8</td>
<td>4.42</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 3. Questionnaire Data for Perceived Difficulty
4.4.2 Perceived Performance

Pilots were asked to rate their performance for the previous trial. Questionnaire data suggests that pilots perceived their performance to be better in low feedback delay trials, F(3, 33)= 3.98, p>.05. The data also showed that pilots rated their performance higher in low density trials compared to high density trials, F(1, 11)= 14.23, p>.05. There was no significant interaction between feedback delay and traffic density in terms of perceived performance, F(3, 33)= 1.52, p=.23.

![Figure 13. Perceived Performance](image-url)
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low_0</td>
<td>5.07</td>
<td>1.33</td>
</tr>
<tr>
<td>High_0</td>
<td>4.85</td>
<td>1.47</td>
</tr>
<tr>
<td>Low_2</td>
<td>4.77</td>
<td>1.36</td>
</tr>
<tr>
<td>High_2</td>
<td>4.50</td>
<td>1.33</td>
</tr>
<tr>
<td>Low_4</td>
<td>5.10</td>
<td>1.19</td>
</tr>
<tr>
<td>High_4</td>
<td>4.22</td>
<td>0.79</td>
</tr>
<tr>
<td>Low_8</td>
<td>4.65</td>
<td>1.13</td>
</tr>
<tr>
<td>High_8</td>
<td>3.87</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Table 4. Questionnaire Data for Perceived Performance

4.4.3 *Perceived Level of Feedback Delay*

Pilots were asked to estimate the severity of feedback delay in the previous trial. Questionnaire data suggests that pilots were able to successfully distinguish between low and high feedback delays within trials of varying delays, F(3, 33)= 5.25, p>.05. The data also showed that pilots were unable to distinguish between low and high feedback delays within trials of varying traffic density, F(1, 11)= 2.42, p=.15. There was no significant interaction between feedback delay and traffic density in terms of being able to identify the severity of feedback delays, F(3, 33)=1.59, p=.81.

![Figure 14. Perceived Level of Feedback Delays](image-url)
Table 5. Questionnaire Data for Perceived Level of Feedback Delay

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>St. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low_0</td>
<td>2.50</td>
<td>1.40</td>
</tr>
<tr>
<td>High_0</td>
<td>2.88</td>
<td>1.39</td>
</tr>
<tr>
<td>Low_2</td>
<td>3.12</td>
<td>1.49</td>
</tr>
<tr>
<td>High_2</td>
<td>3.20</td>
<td>0.81</td>
</tr>
<tr>
<td>Low_4</td>
<td>3.53</td>
<td>1.22</td>
</tr>
<tr>
<td>High_4</td>
<td>3.87</td>
<td>1.07</td>
</tr>
<tr>
<td>Low_8</td>
<td>3.97</td>
<td>1.24</td>
</tr>
<tr>
<td>High_8</td>
<td>4.45</td>
<td>1.05</td>
</tr>
</tbody>
</table>

4.4.4 Perceived Level of Traffic Density

Pilots were asked to estimate the traffic density of the previous trial. Questionnaire data suggests that pilots were able to successfully distinguish between low and high traffic densities within trials of varying delays, $F(3, 33)=10.23, p>.05$. The data also suggests that pilots were able to distinguish between low and high traffic densities within trials of varying traffic density, $F(1, 11)=10.61, p>.05$. There was no significant interaction between feedback delay and traffic density in terms of being able to identify changes in traffic density, $F(3, 33)=1.591, p=.21$. 
Figure 15. Perceived Level of Traffic Density

Table 6. Questionnaire Data for Perceived Level of Traffic Density

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low_0</td>
<td>3.12</td>
<td>0.58</td>
</tr>
<tr>
<td>High_0</td>
<td>3.35</td>
<td>0.66</td>
</tr>
<tr>
<td>Low_2</td>
<td>3.08</td>
<td>1.16</td>
</tr>
<tr>
<td>High_2</td>
<td>3.62</td>
<td>1.03</td>
</tr>
<tr>
<td>Low_4</td>
<td>3.07</td>
<td>0.98</td>
</tr>
<tr>
<td>High_4</td>
<td>3.90</td>
<td>0.91</td>
</tr>
<tr>
<td>Low_8</td>
<td>3.58</td>
<td>0.93</td>
</tr>
<tr>
<td>High_8</td>
<td>4.42</td>
<td>0.97</td>
</tr>
</tbody>
</table>

4.5 Video Analysis

4.5.1 Move and Wait
One of the more prevalent behaviors attributable to feedback delays was the adoption of a “Move and Wait” approach. A “Move and Wait” approach can be described as pilots moving the RAT in small increments then waiting at a position to allow for the conflict detection alerting to catch up and notify them on the status of a proposed route. Waiting periods varied in length from 2 seconds to as long as 13 seconds. Pilots who would begin
the study in the zero-second delay condition would typically not engage in a “Move and Wait” approach, but would rather move the RAT in a fluid motion as they received near instant feedback as to where conflicts lay. However, once a pilot was exposed to a two or four-second delay condition there would be the tendency of a carry-over effect into subsequent zero-feedback delay trials. “Move and Wait” cycles appeared to slow down with increases in feedback delay. This behavior would generally lead to a considerable sacrifice of time in an attempt to find a conflict-free route.

4.5.2 Reverse search
Another behavior that was noticed was pilots tendency to reverse search or backtrack their RAT movements during trials that had 4 or 8-second feedback delays. Since pilots were motivated to complete their trials within a certain time frame they would often move the RAT rapidly across the screen in an attempt to resolve their traffic conflict. During this rapid movement a safe route is detected and reported by the system, but since this information is delayed pilots would back track to their previous positions to find the true location of the conflict-free route.

4.5.3 Articulations
There appeared to be a drastic difference in the use of articulations, or legs, by pilots faced with trials that appeared to be difficult to solve and those which seemed easy. Low-density trials with minimal feedback delays would typically require a small adjustment of the route to avoid a conflict, whereas high-density and longer feedback delay trials required more time and accounted for more failures. In easier trials pilots generally created one additional articulation to avoid an impending conflict. However, some pilots would add two or three additional articulations to an already safe and
relatively efficient route in an attempt to create the most efficient flight path available. There movements were often meticulous and unrushed in an attempt to achieve the most desirable route modification for their aircraft.

The use of articulations emerged quite differently in high-density and high-delay trials. During trial development it was noted that each of the 40 trials could potentially be solved with a single articulation added to the existing path. In the case of more difficult trials pilots would often resort to creating multiple articulations after all single articulation paths were tested. If all direct paths appeared to be conflicted by one or more intruder aircraft, pilots tended to add additional articulations to angle their aircraft around potential conflicts. A single articulation was often used to solve for an initial conflict that was a more immediate threat. Additional articulations would be used beyond that position to avoid conflicts that were further out in the future. These types of multi-articulation routes were much more prevalent in high-density and high-delay trials.

4.5.4. Conflict Test Speed
It was generally noticeable that pilots were able to use the RAT in a smooth and rapid motion when feedback delay was set to zero seconds. They were able to test dozens of potential routes within a few seconds by dragging the RAT across the display. This being mainly due to the rapid communication loop established between the CD&R alert and the pilot. Visual feedback was received much more rapidly allowing pilots to experience closure on the viability of a potential route. Pilots would then decide if they should execute that route or continue their search for a safe and efficient alternative. As illustrated in numerous sections above, pilots exposed to delays received information updates less often, which resulted in slower conflict test speed.
4.5.5. Execution of a Conflicted Route
One of the most compelling behaviors noticed while observing video sessions was a pilot’s retreat from the route execution process due to the emergence of a previously undetected conflict. Typically once a pilot had identified a conflict-free route they would move their cursor toward the bottom of the display to initiate the route execution command. However, on some occasions pilots would retreat mid-process because a delayed conflict had not yet been made visible. Pilots would then terminate the route execution process and continue the search for a safe and efficient route. This behavior occurred mainly in the 4 or 8-second feedback delay condition. In some situations pilots executed a conflicted route before an alert was detected resulting in the execution of a conflicted route (Figure 11).

4.6 Debrief Interview
Various topics were discussed with pilots during the post-experiment interview and this section will address some of the major themes that emerged from their responses.

4.6.1. Comfort Using Display
All pilots stated that the display was easy to use and they could see this being a beneficial component in the cockpit of future aircraft. However a majority of pilots stated that they would be uncomfortable using this display in terminal airspace while operating numerous other instruments. Some suggested that the responsibility of route modifications be assigned to a secondary pilot, known as a “pilot not flying”. The pilot would simply approve or reject the proposed route. Others stated that this tool would be perfectly acceptable in cruise portions of flight that were further than 100 nm from terminal airspace, and others stated that it would be completely acceptable if it were not for lengthy delays.
4.6.2 Level of Delay
When pilots were asked if the level of delay seen was an acceptable amount they nearly all stated that it was a nuisance, and that there would need to be some modifications to be able to perform the task successfully. Some stated that it negatively affected their ability to complete the task rapidly, and it also had a negative effect on their ability to create the most efficient route. Nearly all pilots confirmed that delays longer than those experienced in this study would not have been preferable for real-world use. Pilots’ responses ranged when asked to guess the length of the longest delays. Some guessed 2 to 3 seconds and others assumed the delay was closer to 5 seconds, but on average pilots perceived delays to be at or near 3 to 4 seconds.

4.6.3 Difficulty Attributed to Delay and Traffic
Pilots unanimously acknowledged that some trials were more difficult than others, and that difficulty was attributable to both feedback delays and traffic densities. They generally stated that delay made an already difficult trial harder still. Some pilots stated that they would prefer trials that were low in delay and high in traffic compared to low in traffic and high in delay because they would still be able to rapidly scan congested airspace with a RAT that was very responsive. Numerous pilots stated that delays prevented them from creating route modifications in a rapid manner, and that they would have been able to locate more efficient paths if delay had not been a factor. It was also noted that pilots described feedback delays as “a nuisance”, “frustrating”, and “a hindrance”. Pilots stated that delays in the RAT “held them back” from performing the assigned tasks at a higher level.
Chapter 5 Discussion

The primary goal of this study was to examine the effects of visual feedback delays on pilots who engaged in a route research and modification task. Research in this specific subject area is a crucial component to the successful implementation of advanced automation tools in aircraft of the future. This experiment was specifically designed to record performance metrics of decision time and route distance in addition to various qualitative measures to illustrate the scope of effects feedback delays have on pilots. Numerous studies in various application areas have already suggested that feedback delays have negative effects on user performance. The results of this experiment are consistent with those findings and give valuable insights into the effects of feedback delays throughout varying levels of traffic density.

5.1 Feedback Delay and Traffic Density

Path efficiency is a critical factor in the successful transition to TBO. Unnecessarily lengthy routes may have detrimental effects to the airline industry, especially since fuel accounts for over 40% of a typical airline’s operating costs (Roberson, 2008). Unpredictable route distances would also have a direct effect on the ability for aircraft to reach destinations at their scheduled time. Ball et. al. (2010) estimates that flight delays currently cost airlines $8.3 billion annually and have a negative effect on U.S. GDP of $4 billion. The potential costs due to less than efficient flight plans could prove to be substantial.

5.1.1 Decreases in Route Efficiency

Feedback delays in this study appeared to have a masking type effect on efficient routes that were in a pilot’s surrounding airspace. Furthermore, delays caused available paths to appear fragmented and diminished. The implications of a system with high levels of
feedback delay are that pilots may be unaware of where efficient paths may exist. This factor may not only be attributable to lengthy routes, but may also lead to a potential risk to flight safety if additional measures such as automated route resolutions are not made available. An automated resolution option appears to be an essential component for a system that may be burdened by feedback delays in high-traffic density airspace.

The negative effect of delays on path efficiency were apparent even with delays of two seconds suggesting that engineers should consider preventing delays that exceed that level. Flight paths maintained a similar length in low-density traffic as feedback was increased to four and eight seconds. Additionally, route distances in high-density traffic increased noticeably as feedback delays were increased to four and eight seconds. When focusing on low-traffic density conditions, route distance appeared to double from the zero-second to two-second delay conditions, and maintained a similar length in the 4 and 8-second delay conditions. This may suggest that feedback delays have a uniform effect on route distance regardless of the length of delay, between two and eight seconds, when pilots have an abundance of available routes to choose from. However, in high-traffic density conditions there appears to be a sharp increase in route distance as feedback delay increases from two seconds to four seconds. It appears that high feedback delays are not as manageable in high-density traffic than they are in low-density traffic.

5.1.2 Increases in Time on Task
The data also showed that pilots were able to complete route modifications fastest in the 0-second delay condition, regardless of traffic density level. Decision time rose sharply when feedback delays of two seconds were added. Similar to route distance, decision times in the low-density conditions appeared to level off when delays exceeded two seconds. These effects are expressed quite differently in low and high-density traffic.
Low-traffic density trials appear to experience an increase in decision time at 2-seconds of delay, although decision time stays at a relatively constant level in the 4 and 8-second conditions. High-traffic density trials also showed an increase in decision time at 2-seconds of delay, but continued to increase as feedback delay was set to longer intervals (Figure 5). This is an interesting trend because it suggests that feedback delays cause a relatively similar interruption on pilot performance in low-traffic areas, but can have an exponentially damaging effect as traffic density is increased. Small increases in decision time may seem incremental in the scope of a single flight modification, however these delays draw pilots’ attention from numerous other tasks they may need to complete during flight.

Video data showed that pilots did not commit “timeouts” or conflict executions when feedback delays were absent, resulting in a 100% success rate in trial completion. Interviews with participants clearly indicated that it was easiest to accomplish their assigned tasks of preserving passenger safety, maintaining route efficiency and solving conflicts expeditiously with shorter feedback delays. When examining the data as a whole it is clear that participants performed at the highest level and had the highest level of satisfaction when feedback delays were removed.

5.1.3 Questionnaire Data
Pilots were asked to rate each of forty trials on four different factors. These factors were of importance in order to compare pilots’ self-report of trial difficulty, level of feedback delay, level of traffic density and task success to their quantitative performance data.

Pilots perceived trials to be more difficult as feedback delays were increased. These increases appeared to be more severe in high-density traffic compared to low-density traffic as scores jumped by larger margins between each level of delay. The data
essentially showed that feedback delays made the task appear to be more difficult. The same effect was attributable to traffic density (Figure 12). This data is consistent with quantitative performance data and further supports the case that feedback delays result in detriments to pilot performance.

In addition to trial difficulty, pilots were asked to rate their own performance after each trial. Pilots believed they performed at a less optimal level as feedback delays were increased. A similar trend was present when traffic density was increased (Figure 13). This data is parallel to quantitative performance metrics, which suggests that pilots performed at a less than optimal level when feedback delay or traffic densities were introduced.

Pilots were also asked to rate the amount of feedback delay and traffic density they felt was present in each trial. Results indicated that pilots were able to successfully distinguish changes in feedback delay as well as changes in traffic density. This data supports pilot accounts of feedback delay and traffic density as having distinct negative effects on their task performance. Pilots consistently mentioned that long feedback delays were one of the two main contributing factors to a difficult trial. The secondary factor mentioned was traffic density. Information gathered from interviews with pilots suggests that feedback delays were the more burdensome of the two factors because route searching could still be performed at a high level when the RAT was more responsive. Also, questionnaire data showed that feedback delays appear to have a direct relationship on users’ perception of traffic density and trial difficulty.

This research was an attempt to investigate the characteristics of feedback delays within a pilot’s task of modifying a conflicted route. It should be noted that not all levels of
feedback delay have an equal effect, and that more delay does not necessarily suggest a consistent decrease in route efficiency or decision time. Two seconds appears to be a threshold for pilot performance in a route modification task such as the conducted in this experiment. Thorough cost-benefit analyses should be conducted to measure the potential gains of minimizing feedback delays to near zero levels.

5.1.4 What Have We Learned?
All but one of the hypotheses, H3, were supported from the data analyzed in this study. One of the most significant findings from this study was that feedback delays of two-seconds or greater, within the Conflict Detection & Resolution (CD&R) feature of the Route Assessment Tool (RAT), appear to have a negative effect on pilot performance. Pilots were indeed able to successfully complete trials in the 4 or 8-second delay conditions within the given parameters. However, they consistently reported that delays were problematic, annoying, and that they would have performed better had delays not been present. It is reasonable to assume that delays are impossible to completely eradicate, but perhaps they need not be completely removed from these systems to successfully meet flight and business objectives. With thorough calculations stakeholders may be able to extrapolate the amount of feedback delay that can be deemed as acceptable in order to meet safety as well as business requirements.

We have also learned that the effects of delays on pilots vary depending on traffic density. Higher levels of delay (greater than two seconds) have a stronger effect on pilot performance in high-density airspace. Additionally, there does not appear to be any noticeable benefit in mitigating feedback delays greater than two seconds in low-density traffic. This suggests that there may need to be rules or precautionary measures established for aircraft operating within high-density areas such as terminal airspace.
Numerous pilots reported that feedback delays made their task difficult, and that they would have preferred a system that had little or no delays. Additionally, pilots were able to successfully discern which factors were attributable to decreases in their performance. Pilots were more than willing to offer suggestions on how to alleviate the problems faced with feedback delays. One pilot suggested a meter that displayed the length of the feedback delay to help gauge how long of a waiting period was necessary. Another pilot suggested first moving the RAT to the furthest corners of the display to find any conflict-free route, but then moving the route closer to the center of the display to find an efficient alternative. Other pilots suggested a “loading” icon similar to those found on web browsers would have been helpful in communicating the system was experiencing some sort of delay. The exploration of these and other methods is an important part of addressing feedback delays within TBO. These additions could be very effective at notifying pilots of the status of the system, and could potentially alleviate some of the stress and uncertainty that comes with ambiguous amounts of feedback delay.

The implications of lengthy feedback within the CSD could have strong negative effects on ATC. Pilots who are unable to locate and execute a safe and efficient route in a timely manner may be overly dependent on a route modification data-linked from ATC. Numerous simultaneous route modification requests could inundate controllers and cause an overload of their workload capacities. This type of scenario would essentially render the proposed system infeasible due to inefficiencies and safety risks. Air-Traffic Controllers would face similar challenges if they were faced with feedback delays. They would probably prefer the absence of any amount of feedback delay, and it is reasonable to expect similar performance decreases as pilots. However, it could be
argued that it would be more difficulty to successfully complete route modifications of numerous aircraft when feedback delays are present. Delays to one aircraft could have a compounding effect on proceeding aircraft due to pilots’ inability to attend to more than one aircraft at a time. Additionally, monetary losses are exponentially greater if delays are excessive on the groundside primarily due to the large volume of route modifications a controller would make relative to a pilot. ATC would need to be evaluated while exposed to feedback delays in order to get a true understanding of how their performance would be affected.
Chapter 6 Conclusion
Feedback delays have proven to be problematic to users in a number of application areas including manned and unmanned aircraft, medical devices, web, gaming as well as consumer electronics. The findings in this thesis will provide the basis for a series of new research initiatives in the area of feedback delays within NextGen systems. It is safe to assume that feedback delays will be a very real problem for ATC operating within a trajectory-based traffic management system. The effects of delays on ATC could in fact be exponentially worse due to the numerous aircraft they oversee. The effects of delays on performance were noticeably worse in higher density traffic and it can be expected that ATC will experience the same difficulties. Research shows that ATC are unable to safely manage aircraft when traffic reaches approximately triple today’s levels (Prevot, 2008). The assumption is that since ATC rely on trial planner tools more so in high-densities they will struggle if those tools are subjected to feedback delays. Thus, it is even more crucial to mitigate feedback delays within groundsie trial planner tools.

There are other aspects of feedback delays that are important to consider for individuals responsible for the implementation of NextGen systems such as the CSD. The potential financial costs are immense if delays are not managed below acceptable levels. A typical Boeing 747 burns approximately 3200 gallons of fuel per hour while traveling at cruise speed, i.e. 585 mph. At this rate the aircraft would consume 5.5 gallons per mile. If fuel costs around $5 per gallon, the average mile at cruise speed would cost around $27. This research shows that the average route distance for a trial in the zero-second and high-density condition was 16 miles. Thus, the average route would cost about $432 in fuel. If feedback delay was increased to 2-seconds, fuel costs per route would be about $570, and if set to 4-seconds it would be approximately $860 per route. These amounts may seem
trivial but they are actually quite substantial when considering the number of route changes one aircraft may complete in the course of a year. The financial feasibility of TBO related systems could arguably be dependent on the mitigation of feedback delays in order for pilots and controllers to identify the most efficient flight paths.

Therefore the successful design and implementation of tools such as the NASA Ames CSD will be tremendously beneficial to the JPDO’s proposed method of conducting TBO. Through continued research and analysis of system performance criteria, tools such as these may eventually prove to be effective at accommodating high levels of air traffic while improving overall safety and efficiency.
References


