2006

Comparing 4-D clearance formats during taxi operations

Jennifer Lynn Williams
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Williams, Jennifer Lynn, "Comparing 4-D clearance formats during taxi operations" (2006). *Master's Theses*. 2985.

DOI: <https://doi.org/10.31979/etd.28bc-ky9j>

https://scholarworks.sjsu.edu/etd_theses/2985

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

COMPARING 4-D CLEARANCE FORMATS DURING TAXI OPERATIONS

A Thesis

Presented to

The Faculty of the Department of Psychology

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Jennifer Lynn Williams

August 2006

UMI Number: 1438600

Copyright 2006 by
Williams, Jennifer Lynn

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 1438600

Copyright 2007 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2006

Jennifer Lynn Williams

ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

Dr. Kevin Jordan

Dr. David Foyle
NASA Ames Research Center

Dr. Tony Andre

APPROVED FOR THE UNIVERSITY

ABSTRACT

COMPARING 4-D CLEARANCE FORMATS DURING TAXI OPERATIONS

by Jennifer Lynn Williams

This study investigated pilots' abilities to comply with 4-D clearances during taxi operations. A 4-D taxi clearance, in which a pilot is required to be at a specific location at a specific time, allows an aircraft to make an active runway crossing without slowing or stopping, enabling surface traffic to flow smoothly and efficiently. Pilots were asked to comply with 4-D taxi commands that were presented in 3 different formats: speed, time, and "both." They taxied with these formats across different route distances and with various commanded speeds and times. Analyses evaluated the accuracy and consistency with which pilots complied with the clearances. Display usage was also evaluated with an eye tracker. Results indicate that pilots were most accurate and more consistent when presented with both speed and time information. Pilots dwell patterns indicate that they used the speed and time information during different phases of taxiing.

Acknowledgments

I would first like to thank my committee members, Dave Foyle, Kevin Jordan, and Tony Andre, for their guidance, encouragement, and patience throughout this project. Many thanks also go to my co-workers Becky Hooey, Glenn Meyer, and Ron Miller for offering help when it was needed and humoring me when it wasn't. My appreciation goes to all of my friends who inspired me and helped me keep my sense of humor. I could not have completed this process without the daily love and support from Sam Gharabally, who always knew how to encourage me and when to simply listen. Finally, I owe all of my accomplishments to my parents, Shirley and Cliff Williams, whose love and hard work have made my hard work possible. Thank you to my entire family for always believing in me.

TABLE OF CONTENTS

SECTION	PAGE
INTRODUCTION	1
METHODS	16
Participants	16
Apparatus	16
Displays	20
Experimental Design	24
Procedures	27
RESULTS	29
Time of arrival absolute error	31
Time of arrival error bias	37
Standard deviation of speed	44
Percentage of eye dwell time	52
DISCUSSION	62
Compliance Accuracy	62
Consistency	66
Strategy	68
Conclusions	70
Implications for Design	71
REFERENCES	72

APPENDICES

Appendix A: Post block and post study questionnaires.....	74
Appendix B: Post block response frequencies.....	81
Appendix C: ANOVA with quartiles for TOA absolute error.....	85
Appendix D: ANOVA with quartiles for TOA error bias.....	86
Appendix E: ANOVA with quartiles for SD of speed.....	87
Appendix F: ANOVA with quartiles for percentage of dwell time.....	88
Appendix G: Human Subjects Approval.....	89

LIST OF TABLES

TABLE	PAGE
1. Commanded Time of Arrivals as a Function of Route Distance and Commanded Speed.....	4

LIST OF FIGURES

FIGURE	PAGE
1. Actual out-the-window view.....	18
2. Applied Science Laboratories head mounted eye tracker.....	19
3. Graphical depiction of the speed format display.....	21
4. Graphical depiction of the time format display.....	22
5. Graphical depiction of the speed/time format display.....	23
6. Electronic map with highlighted taxi clearance.....	25
7. Electronic map without highlighted taxi clearance.....	26
8. Mean time of arrival (TOA) absolute error for display formats.....	33
9. Mean time of arrival (TOA) absolute error as a function of the interaction between distance and 4-D command.....	34
10. Mean time of arrival (TOA) absolute error as a function of the interaction between format and quartile.....	36
11. Mean time of arrival (TOA) error bias for route distances.....	38
12. Mean time of arrival (TOA) error bias for 4-D command.....	39
13. Mean time of arrival (TOA) error bias as a function of the interaction between distance and 4-D command.....	41
14. Mean time of arrival (TOA) error bias as a function of the interaction between format, distance, and quartile.....	43
15. Mean standard deviation of speed for format.....	45
16. Mean standard deviation of speed for route distance.....	46
17. Mean standard deviation of speed for 4-D command.....	47
18. Mean standard deviation of speed as a function of the interaction between format and distance.....	49
19. Mean standard deviation of speed as a function of the interaction between format and quartile.....	51
20. Mean percentage of eye dwell time for formats.....	54
21. Mean percentage of eye dwell time for route distances.....	55
22. Mean percentage of eye dwell time as a function of the interaction between format, distance and quartile.....	57
23. Comparison of the mean percentage of eye dwell time, between the speed and time formats and the speed and time information within the both format.....	59
24. Comparison of the mean percentage of eye dwell time, between the speed format, the time format, and the speed information and time information within the both format as a function of the interaction with quartile.....	61

Introduction

Delays caused by airport surface congestion, such as those incurred by aircraft waiting to cross active runways, result in the largest delay cost in the US airspace system (Glass & Gawdiak, 1997). These delays result from a competition for resources such as runways and taxiways (Glass & Gawdiak, 1997). Cheng, Sharma, and Foyle (2001) point out that reducing airport surface congestion requires not only increasing the amount of usable airport surfaces (i.e., adding runways and taxiways), but also increasing the efficiency with which existing runways and taxiways are utilized through more precise taxiing and improved runway coordination. This may be accomplished with the development of surface management systems (SMS) that provide information to air traffic controllers that helps them direct traffic in a more efficient and safe manner than that of current day operations (Glass & Gawdiak, 1997). Although most of the SMSs under development provide only advisory information, automated systems are being developed that utilize dynamic algorithms to calculate the most efficient movement of all surface traffic based on changing conditions (Hecker, Doehler, Korn, & Ludwig, 2001; Cheng & Foyle, 2002). In addition to physical information about surface traffic that these systems can utilize, such as location, heading, and speed; time can be included as an additional component of movement in order to coordinate traffic. This fourth dimension (time) could allow the system to predict future traffic locations and enable four-dimensional (4-D) coordination of surface traffic as aircraft move forward in time. This information can be presented to air traffic controllers and pilots in order to coordinate runway crossings, and it could potentially lead to shorter and fewer taxi delays. If these

systems are to be implemented, pilots will be required to comply with 4-D taxi clearances provided by the automation. A 4-D taxi clearance keeps traffic flowing smoothly by requiring pilots to reach a particular runway crossing or taxiway intersection by a specified time in order to ensure that they do not have to stop to wait for cross traffic. Research has shown that aircraft take twice as long to cross active runways when starting from a standstill compared to crossing without having to stop (Cheng, Sharma, & Foyle, 2001). If pilots can reach a runway crossing within a window of time that allows them to cross without stopping, they can potentially achieve shorter taxi times, increase fuel efficiency, and increase traffic throughput of the airport (Cheng, Sharma, & Foyle, 2001). Future implementation of automated 4-D taxi clearances would likely require advanced displays to support pilots during taxi operations. The purpose of the present study was to determine the basic information required for those displays and the fundamental abilities that pilots have in complying with 4-D taxi clearances.

Although 4-D taxi clearances introduce the dimension of time into surface operations, speed is also an important component, since it interacts with time and distance as dimensions of movement. Thus, this study compared taxi commands presented in speed and time formats because both speed and time were considered relevant pieces of information for complying with 4-D taxi clearances. The general task that pilots performed was to taxi to an active runway crossing by a given time, or at an average speed, with as little error as possible. For the speed format, pilots were instructed to taxi to a specific runway crossing at a commanded average speed. With the time format, pilots were instructed to taxi to a specific runway crossing by a commanded

time of arrival (TOA). There was also a speed/time format in which they were given both a commanded speed and a commanded TOA that corresponded with each other and pilots were instructed to comply with both of them.

The 4-D commands were based on combinations of four commanded speeds (in knots) and three route distances. In order to calculate commanded TOAs for the time and speed/time formats, the given route distance was divided by the commanded speed (converted to equivalent feet per minute). Thus, when given a commanded TOA, it was based on an implicit speed/distance combination. The 4 commanded speeds for the 3 route distances were equivalent to 12 commanded TOAs, listed in Table 1. For presentation and discussion purposes, data are presented in terms of the four commanded speeds, since the equivalent commanded times were linked to the underlying route distance and could not be reduced to four distinct categories.

The main research question was whether pilots are able to comply more accurately with 4-D commands given in a speed, time, or speed/time format. The variability of their ground speed was also of interest in order to determine how consistent pilots' speed was within the routes. Additionally the proportion of time that they looked at the different displays was assessed in order to determine how they used the information in the different formats. In order for 4-D taxi clearances to increase the efficiency of taxi operations, not only is compliance at the final destination (a runway crossing) important, but compliance throughout a taxi clearance is also necessary for the coordination of numerous aircraft and for an automated system to be able to accurately predict where various aircraft will be in the future. Thus, in addition to analyzing performance for

Table 1

*Commanded Time of Arrival as a Function of Route**Distance and Commanded Speed*

Route distance	Commanded speed			
	10 kts	14 kts	18 kts	22 kts
3,000 ft	2:58	2:07	1:39	1:21
6,000 ft	5:55	4:14	3:17	2:41
12,000 ft	11:51	8:28	6:35	5:23

Note. Actual route distances were approximate and within 5% of target distance.

whole clearances, it was also of interest to analyze performance during different phases of taxi clearances. Each route was divided into four equal-distance quartiles (e.g., for a 3,000 ft route, quartiles were approximately every 750 ft). Analyzing accuracy, variance, and eye dwell time during different quartiles provided a more detailed assessment of how consistent pilots were and what their strategies might have been within the taxi routes.

Because multiple tasks are involved in taxing an aircraft, such as watching for approaching traffic and navigating the airport surface, pilots cannot focus all of their attention on complying with a taxi command and therefore, must estimate their speed or elapsed time when not looking at the display. In a review of the literature, Wickens and Hollands (2000) explain that people operate as “intuitive statisticians” by estimating the state of their environment from a sample of observations. Research has shown that people estimate the mean of a set of observations well, but have a conservative tendency to disregard extremely high or low values (Wickens & Hollands, 2000). Studies of the perception of speed indicate that in general, people tend to underestimate the speed at which they are traveling (Denton, 1966; Recarte & Nunes, 1996). This phenomenon is more pronounced at lower speeds than higher speeds (Recarte & Nunes, 1996). In other words, drivers typically estimate that they are going slower than they actually are, and this error tends to be larger at lower speeds. However, studies also show that speed estimation generally improves at higher speeds. This pattern is similar for both passive verbal estimations and active speed productions while driving (Recarte & Nunes, 1996). In this experiment, pilots’ speed estimations were in the form of an active speed production task while taxiing a simulated aircraft on an airport surface. However,

participants were not required to estimate their actual speed, as they had a ground speed indicator telling them their current speed at any given moment. Instead, they had to estimate their *average* speed across the route. Based on the pattern of general underestimation of speeds and conservative tendency for mean estimation found in the literature, we might expect pilots to accumulate error as they traverse a route and that longer routes would accumulate more error than shorter routes because there is no inherent feedback as to their average performance and thus, no way to correct their error. This lack of feedback for average performance, in combination with the availability of information about their current performance, should also lead to a particular pattern of variability across the quartiles, in which pilots would quickly attempt to reach the commanded speed at the beginning of the trial and maintain the speed throughout the trial without being prompted to correct their error at the end, making their speed variability higher at the beginning of the routes.

There are various theories about the underlying processes of time estimation. However, despite different explanations, there appears to be a general finding that prospective time estimations tend to be negatively related to task demands (Zakay, Nitzan, & Glicksohn, 1983; Zakay, 1992). The more difficult or demanding a task is, the shorter an individual's time estimation of that task will be. This trend is more pronounced for longer durations (Brown, 1985). Additionally, "empty" time spans are typically judged as longer than "filled" time spans (Zakay, Nitzan, & Glicksohn, 1983). These findings are often explained by attentional cognitive counter models that propose that time estimation is associated with a "cognitive counter" that requires mental resources for

its operation (Zakay, 1992). These models consider time estimation to be a function of the amount of attention that is allocated for processing temporal information. When an event is presented, the counter ticks off an increment and time estimation is then based on the value of increments processed by the counter. The more mental resources that are available to operate the cognitive counter, the more increments can be recorded, producing the negative relationship between prospective time estimations and task demand (Zakay, 1992). This study used a prospective time production task in which participants knew in advance that they needed to estimate a certain time by taxiing to their cleared destination as close to the commanded TOA as possible. Again, this was not a pure estimation task because they had a clock display that counted up as they traversed the route. However, because the task involved filled time spans during which pilots were also required to taxi the aircraft, navigate the airport, and remember their taxi clearance, they were likely unable to attend to the clock display constantly and would have to estimate the elapsed time while attending to other information. Thus, the general task was considered to be somewhat demanding, and it was expected that overall, participants would underestimate the elapsed time. Specifically, longer commanded times and taxi routes were expected to result in less accurate time estimations than shorter commanded times and routes due to the increased mental demand associated with remembering longer clearances and navigating more extensive routes. The variability of the time format was expected to produce a quartile pattern opposite that of the speed format. The time format did not provide current performance information, but did provide feedback at the end of the trial in the form of the clock displaying their TOA (and

whether it was early or late compared to the commanded TOA), thus making time more relevant at the end of the trial. Research has found evidence indicating that attention to time increases whenever the dimension of time becomes more relevant for adaptation to the environment (Zakay, 1992). It can be argued that the time display would be the most relevant near the runway crossing and this “temporal relevance” at the ends of routes should lead pilots to attempt to correct their error at the ends of the routes, creating greater variability in the final quartile.

These anticipated time estimations are similar to the anticipated speed estimations in that it was expected that pilots would underestimate both speed and time. However, because the task is a production task as opposed to a simple estimation task, the two formats should produce opposite results with regards to the direction of the error. In other words, if pilots do underestimate their average speed they will travel too quickly and arrive too soon, causing their average speed to be greater than the commanded speed. However, if pilots underestimate time, they will travel too slowly and arrive later than the commanded TOA. The general trend of the accuracy data for distance and 4-D command in the speed and time formats should be the same. In each format, participants will perform less accurately with longer times and/or slower equivalent speeds, but their estimations will be more accurate for shorter times and/or faster equivalent speeds. They will also perform more accurately with short routes compared with long routes.

Despite the anticipation of similar estimation patterns with the speed and time formats, it was still expected that there would be a difference between performance accuracy in the speed format compared with the time format. Although there is a large

body of research that investigates how humans perceive and estimate speed and time, few studies directly compare the two modalities. There is some psychophysical research that examines the differences between different sensory modalities. This research has demonstrated that during tasks that require cross-modality attention between vision and other senses, visual dominance occurs (Wickens & Hollands, 2000). For example, participants asked to respond to either a light or a tone (with different hands) failed to notice the tone when the two stimuli were presented together. A more real world example occurs when sitting in a car at a stop light, one might experience the illusion that your car is moving backwards, because you see the car next to you moving forward. The visual stimulus of the other car's motion overrides the proprioceptive sensation that your car is not moving (Wickens & Hollands, 2000). Although having participants respond in the modality that is in conflict with vision can moderate this conflict, biases toward the visual stimuli are likely if the task itself is visual (Wickens & Hollands, 2000). Visual dominance has primarily been studied in comparison with auditory and proprioceptive modalities, but it is possible that it would also show dominance over the perception of time. Vision might dominate one's sense of time due to the salience of visual cues compared to temporal time cues. While vision is a direct percept, time is inferred from other cognitive processes and is not directly perceived (Coren, Ward, & Enns, 2004). The temporal nature of this inference might make cues for the passage of time less salient than directly perceived visual cues that are used during speed estimation. Research has shown that more salient cues tend to be more attended to and weighted more heavily when integrating information (Wickens & Hollands, 2000). With regards to this study,

the visual cues associated with taxiing around an airport surface might be more salient and thus would visually dominate the more indirect sense of time when participants are presented with a taxi command in the time format. Therefore, it was expected that participants' compliance would be more accurate for the speed format than for the time format because of the visual salience of the cues used during speed perception.

Of the three formats, the speed/time format, in which pilots were presented with both speed and time information, provided the most information for taxiing. More information is not always better due to confusion that can result from display clutter or conflicts between competing stimuli (Wickens & Hollands, 2000; Selcon, 1999). However, if the different pieces of information are congruent and complementary to each other, they can facilitate more accurate performance (Selcon, 1999). This phenomenon is known as a redundancy gain, which is a reduction in reaction time or an improvement in performance accuracy due to multiple presentations of the same information (Jackson & Selcon, 1997). Examples of redundancy gain include presenting instructions with both pictures and text, or giving a warning signal verbally and spatially. Presenting pilots with both speed and time information should produce a redundancy gain by giving participants the information needed to comply with taxi commands in two complementary formats. Research has shown that people use a subjective form of the equation for speed ($\text{velocity} = \text{distance} / \text{time}$) to relate estimates of speed, time, and distance (Awe & Johnson, 1993). By providing pilots with speed and time information, two pieces of that equation are available to be compared and checked against each other. Thus, the speed/time format, in which they receive both speed and time information, should produce the most accurate

compliance with taxi commands due to the redundancy of the information, compared with the speed or time formats individually. As with the speed and time formats, in the speed/time format, pilots were expected to comply more accurately with short routes and commanded TOAs and fast commanded speeds. Despite potentially resulting in more accuracy than the other formats, the speed/time format was expected to produce a combination of the speed and time quartile patterns for speed variability. Specifically, pilots were expected to use the speed information to correct their initial error at the beginning of the trial and maintain it throughout; they would then switch to the time information at the end of the trial in order to correct their accumulated error.

The main research question was how accurately pilots could comply with different 4-D display formats. Even though some trials presented speed information only, a TOA can be inferred based on the commanded speed and route distance. Thus, for the purposes of continuity, accuracy for all trials was measured as TOA absolute error, by converting actual average speeds and commanded speeds to TOA absolute error for the trials in the speed format. A main effect of format was expected for TOA absolute error. Specifically, because the temporal nature of time is considered to be less salient than visual speed cues, it was expected that performance with the speed format would be more accurate than with the time format. However, because the speed/time format could produce a redundancy gain, it was expected that performance with the speed/time format would be more accurate than with either the speed or time format. Additionally, it was expected that, as in other studies, participants would underestimate their speeds in the speed format, causing them to travel at an average speed that is faster than the

commanded speed. Alternatively, in the time format, participants' time estimations would be significantly shorter than the commanded times, resulting in arrival times that are later than the commanded arrival times. A main effect of route distance was also expected to show that regardless of format, accuracy would decrease as distance increased because error would accumulate as pilots progressed through the route, making short routes (3,000 ft) the most accurate and long routes (12,000 ft) the least accurate, and that increase in error would be due to pilots arriving earlier as distance increased. It was also expected that there would be a main effect of 4-D command, resulting in more accurate compliance and later arrival times as commanded speed increased. However, because the commanded speed is an average across the trial, pilots might have difficulty safely compensating for slower spool-up and turn speeds during short routes with faster than average speeds, thus causing them to taxi at an average speed slower than the commanded speed. Thus, an interaction between distance and 4-D command was expected, which would show a pattern of less error as commanded speeds increased, for medium (6,000 ft) and long routes (12,000 ft), but not for short routes (3,000 ft).

When analyzing performance by quartiles, TOA accuracy must be interpreted carefully, because despite being instructed to attempt to be consistent throughout the taxi routes, pilots were not explicitly instructed to get to certain points of the routes at specific times. However, TOA absolute error at each quartile could still provide useful information about taxi performance. As in the main analysis of compliance accuracy, pilots were expected to accumulate error across the trials, making longer routes less accurate than shorter routes. However, when examining accuracy by quartiles, it was

expected that the formats would produce different patterns as well. Research on order effects for mean velocity and total time estimation has shown that mean velocity estimation is subject to both primacy and recency effects, meaning that perception of speed during the first and last segments of a route is more heavily weighted when estimating the mean for the entire route. Alternatively, for total time estimation, results indicated that participants did not recall durations of previous segments and instead, relied on the cumulative advancement of the cognitive counter (Leiser, Stern, & Meyer, 1991). In the context of this study, the primacy effect for mean speed estimation could lead pilots to weight the first quartile, which is expected to be the most accurate, more heavily than the other quartiles, leading them to estimate that their average speed was closer to the commanded speed than it actually was and accumulate error across the route and arrive early. Because the speed format did not provide inherent performance feedback, pilots were not cued to correct their error at the end of the route, making the last quartile the least accurate. However, it was expected that for the time and speed/time formats, error would accumulate up to the third quartile and then, when time information became more relevant in the final quartile, pilots would use the time information to correct their error. With the time format, pilots' would taxi increasingly slowly, and then attempt to catch up at the end of the route. In contrast, with the speed/time format pilots would taxi slightly fast as with the speed format, until they slowed down near the end of the route to avoid arriving early. Thus, an interaction among format, distance, and quartile was also expected to show differences between quartiles for the formats with medium and long routes, but not for short routes because they will not provide enough

distance/time to accumulate substantial error. Specifically, for medium and long routes with the speed format, TOA error was expected to accumulate as pilots advanced through the quartiles with no correction at the end of the route. For the time and speed/time formats, TOA error was expected to accumulate as pilots advanced through the quartiles and then decrease as they corrected their error near the end of the route.

Standard Deviation (SD) of speed was analyzed in order to assess pilots' speed consistency throughout trials. Because the time format provided only a status display of a clock, whereas the speed and speed/time formats provided performance/control information that could be used to regulate speed maintenance, variability for time was expected to be greater than speed and both, which should not differ. A main effect of route distance was expected to show that variability would increase with shorter routes because there was less distance for maintenance between the spool up, at the beginning of a route, and the potential error correction, at the end of a route. A main effect of command was also expected, with variance increasing as 4-D commanded speed increased. Since the 4-D commands were actually commanded average speeds for the whole route, faster speeds would require pilots to slow down during turns and then overcompensate with above average speeds during straight-aways. However, if the 4-D command was a speed closer to a safe turning speed, pilots would not need to slow down for turns and speed up afterwards, thus reducing the fluctuations of the average speed. For the quartile analyses, an interaction was expected among format, distance, and quartile, with differences between quartiles for medium and long routes, but not for short routes because they would not provide enough distance or time to simply maintain a taxi

command. The quartile pattern for medium and long routes in the speed format was expected to show that variability would be large during the initial quartiles and decrease as quartile increased. In the time format, the pattern should show that variability remains at a constant level until the last quartile where it would increase due to error correction at the end of the route. The speed/time format was expected to produce a combination of the quartile patterns expected from the speed and time formats. Specifically, the first and fourth quartiles would have greater variability than the other quartiles during medium and long routes.

In order to gain insight into how pilots used the different formats, a head-mounted eye tracker was used to collect information about what percentage of the trial pilots looked at display information. In general, pilots were expected to look at display information more when correcting error than when simply maintaining their performance. Because the time format provided status information as opposed to control information, it was expected that pilots would look at the time display less than the speed display. Alternatively, the speed/time format provided a combination of status and control information from the speed and time formats, making it likely that pilots would look at the speed/time display more than the time or speed displays. It was also expected that there would be a main effect of route distance, which would reveal that pilots used the display proportionally the least during long routes (12,000 ft) and the most during short routes (3,000 ft), again, because longer routes have more distance during which pilots can maintain performance between referring to the display frequently at the beginnings and ends of routes for error correction. Similar to the SD of speed pattern, a main effect of

4-D command was expected, with pilots looking at displays less with slow commanded speeds than with fast commanded speeds, because they would have to refer to the displays more often when slowing down for turns and speeding up for straight-aways for faster commanded speeds. As with the other variables, an interaction was expected among format, distance, and quartile, for which display use would differ between quartiles for formats for medium and long routes, but not for short routes. For medium and long routes with the speed format, pilots should look at the display more during the first quartile compared with the other three quartiles, whereas with the time format they should look at the display more during the fourth quartile compared with the other quartiles. Again, the speed/time format should show a combination of the speed and time patterns, with more use in the first and fourth quartiles compared with the other two quartiles, because it is during those phases of the routes that the displays are most salient for error correction.

Methods

Participants

Eighteen instrument-rated, current commercial airline captains or first officers, who had flown a minimum of 500 hours as captains within the last 4 years, participated in the experiment. The mean age of the pilots was 49 years ($SD = 8$ years) and the mean number of flight hours as a commercial captain was 8,383 hrs ($SD = 8,031$ hrs) with a range of 500 to 30,000 hrs. All participants had normal vision or vision corrected to normal with soft contact lenses. Pilots were compensated for their time.

Apparatus

The experiment was conducted at the NASA Ames Research Center in a medium-fidelity surface operations simulator with Boeing 737 modeled dynamics. The simulated environment modeled Dallas Fort-Worth International Airport at 1,200 Runway Visual Range. The main out-the-window scene was projected at approximately 3,500 Lumens by an LCD projector onto a 6 ft by 8 ft screen that was approximately 8 ft away from the eye point of the participant (see Figure 1). Peripheral views were displayed on two 16 in. (H, 31.89 deg visual angle) by 12 in. (V, 24.19 deg) monitors located approximately 28 in. to the left and right of the participant, respectively. An airport diagram and text taxi clearance were displayed on a 13 in. (H, 18.46 deg visual angle) by 9.5 in. (V, 13.54 deg visual angle) monitor located approximately 40 in. in front of the pilot, below the out-the-window scene. Pilots controlled the simulated aircraft using a tiller, throttle, and set of toe-brakes. While taxiing, pilots wore an Applied Science Laboratory head-mounted eye tracker, Model 501, which was attached to an adjustable headband (see Figure 2). The eye tracker and headband combined weighed 8 ounces.

At the beginning of the experiment, pilots filled out a demographics questionnaire and signed an informed consent form. At the end of each of the three blocks of trials, participants completed a written questionnaire assessing their subjective measures of workload and of their own performance for each block. The questionnaire for the final block also included questions assessing their preferences for the command formats (see Appendix A).

Figure 1. Actual out-the-window view of the speed/time display. Note that all text and lines in white were actually displayed in green.

Figure 2. Applied Science Laboratories head mounted eye tracker, Model 501.

Displays

The current instantaneous ground speed indicator, elapsed time, and the taxi commands were superimposed graphically over the out-the-window screen simulating a Head-Up Display (HUD). When the command was in the speed format, the current ground speed was displayed inside a line box in the upper left corner of the main screen and the commanded speed was displayed 6 in. below it (see Figure 3). The text of the ground speed measured 2 in. (H, 1.15 deg) by 1 in. (V, 0.57 deg), and the commanded speed measured 2 in. (H, 1.16 deg) by 1 in. (V, 0.58 deg). For the time format, elapsed time was displayed inside a line box in the upper right corner of the out the window screen and the commanded time was displayed 6 in. below it (see Figure 4). The text of the elapsed time measured 2 in. (H, 1.14 deg) by 1 in. (V, 0.58 deg) and the commanded time measured 2 in. (H, 1.15 deg) by 1 in. (V, 0.57 deg). When the command was in the speed/time format, current ground speed, commanded ground speed, elapsed time, and commanded time were displayed in their respective locations simultaneously (see Figure 5).

During all trials, the HUD displayed a “text triad” in the upper center of the out-the-window screen that was 18 in. (H, 10.71 deg) by 4 in. (V, 2.39 deg). The text triad indicated the current taxiway on which the aircraft was traveling and the upcoming taxiways on the left and right respectively. In order to aid with navigation in the absence of a first officer, the HUD also displayed a highlighted centerline superimposed over the out-the-window scene along the cleared taxi route. This “scene-linked” centerline

Figure 3. Graphic depiction of the speed format display, with the highlighted centerline and text triad (not to scale).

Figure 4. Graphic depiction of the time format display with the highlighted centerline and text triad (not to scale).

Figure 5. Graphic depiction of the speed/time format display with the highlighted centerline and text triad (not to scale).

symbolology conformed to the out-the-window scene as though it was part of the simulated environment (Foyle, McCann, & Shelden, 1995).

An electronic, north-up diagram of the airport and a text taxi clearance were also present during all of the trials. At the beginning of each trial, while the pilot received and read back the taxi clearance, the map displayed an ownship icon at the starting point of the clearance and a magenta line highlighting the cleared taxi route (see Figure 6). After the pilot repeated the clearance back to ground control (the experimenter), the trial started and the ownship icon and highlighted route disappeared. The text taxi clearance remained at the bottom of the diagram throughout the trials (see Figure 7).

Experimental Design

This experiment was a 3 (format) x 3 (distance) x 4 (4-D command), within subjects design. The three levels of command format were speed, time, and speed/time. The levels of route distance were short (3,000 ft), medium (6,000 ft), and long (12,000 ft) and consisted of 45 unique routes that averaged 1 turn per 3,000 feet across routes. Taxi commands consisted of 10, 14, 18, and 22 knots or the equivalent times based on the distance of each route, which ranged from 1 min 21 sec to 11 min 51 sec. Additional analyses were conducted with a 3 (format) x 3 (distance) x 4 (4-D command) x 4 (quartile), within subjects design. The four levels of quartile were based on data markers positioned proportionally to the distance of each route. The data markers were not visible to the participants and were used only for data analysis.

The experiment was comprised of three blocks of 15 trials each. Each block consisted of one of the three command formats. At the beginning of each block

Figure 6. Electronic map of Dallas-Fort Worth Airport with the taxi clearance highlighted in magenta and the ownship icon at the beginning of the taxi clearance. The text clearance is displayed below the map.

Figure 7. Electronic map of Dallas-Fort Worth Airport with the text clearance displayed below the map. The highlighted taxi clearance and ownship icon were removed after the pilot confirmed the taxi clearance.

participants taxied three practice trials of the given format, after which they completed 12 experimental trials, resulting from the combination of the 3 route distances and 4 commanded speeds. The presentation order of the blocks was counterbalanced across subjects. Taxi commands and route distances were randomized within each block and matched for each command format group.

Although participants were given commands in both the speed and/or time formats, performance data from the speed format was converted from average ground speed to TOA error in order to compare the different command formats. The TOA absolute error was analyzed to determine the magnitude of the error compared to the commanded TOA. Average TOA error bias was also analyzed in order to determine the direction of the error. The dependent variable used to assess variability in pilots' taxi performance was the SD of speed. The percent of dwell time spent looking at specific eye tracker zones was also analyzed to determine how much the participants used the different displays.

Procedures

The participants first completed a minimal risk consent form, a demographics questionnaire, and read general instructions for the simulation. Next, the participants were taken into the simulation room and seated at the controls. Before beginning the first experimental block, participants completed three initial phases: familiarization, training, and introduction. The participants first taxied through three extended trials in order to become familiar with the simulator and the airport. During the extended trials the pilots were given a taxi clearance and a ground speed indicator and were told to taxi at a variety

of speeds at which they were comfortable, while staying on route. In order to ensure that participants were calibrated to the simulator and were able to estimate speed based on the cues from the simulator, the participants completed three to four training trials during which they were asked to taxi at specific speeds without using a ground speed indicator. Once they indicated, by pressing a button on the joystick, that they thought they had reached the target speed, they were shown their actual speed on the main screen. If their actual speed was within 20% of the target speed, they were instructed to come to a complete stop. If their actual speed was more than 20% above or below the target speed, they then had an opportunity to correct their actual speed and maintain the target speed for five seconds, after which they were instructed to come to a complete stop. This process was repeated for three speeds (10, 20, and 30 kts) until they had successfully taxied at each target speed without a ground speed indicator two times. Finally, in order to ensure that they understood the different conditions that they would see during the experimental trials, the participants taxied through one introductory trial for each command format, each of which consisted of the same displays and procedures as the upcoming experimental trials. The participants were then given a 10-minute break.

Following the break, the experimenter gave the participants detailed instructions for the experimental trials and fit the eye tracker for calibration. When calibration was complete, the experimenter left the room and the first block of experimental trials began. The experimenter told the participant what his first command format was and he then completed 3 practice trials and 12 experimental trials for that format. For each trial, the participants received a verbal taxi clearance from ground control via an intercom, a text

taxi clearance at the bottom of the airport diagram, and a taxi command in the speed, time, or speed/time format. The experimenter, acting as ground control, read the taxi clearance to the participants and confirmed that their read back of the clearance was correct. Once they received this information, a tone was sounded and the trial began. The pilots' task was to maintain the commanded average speed, or arrive at the destination at the commanded time, as closely as possible without going over or under the commanded speed/time. Trials contained a small amount of realistic traffic around the airport, but there were no potential incursions or close calls with any of the traffic present. After each trial they were shown their actual average speed, or TOA, as well as the commanded speed, commanded TOA, or both (depending on the condition).

The participants completed one block, consisting of 15 trials of the first command format, in approximately 90 min. After they finished the first block they completed a questionnaire and then took a 45-min lunch break. After the lunch break, the practice, experimental, and questionnaire process was repeated for the two remaining blocks until they had completed the experiment. The participants took another 10-min break between the second and third blocks. After completing all of the blocks, the participants were debriefed and thanked for their participation. Pilots participated individually and each pilot completed the entire experiment in one day, which lasted approximately 9 hrs including instruction, breaks, and lunch.

Results

A three-way, repeated measures analysis of variance (ANOVA) with format (speed, time, and speed/time), distance (3,000; 6,000; and 12,000 ft), and 4-D command

(10, 14, 18, and 22 kts) was performed on TOA absolute error, TOA error bias, SD of speed, and percentage of dwell time, in order to assess the accuracy, consistency, and strategy with which pilots complied with taxi commands presented in different formats. These analyses also investigated differences between performance with the different 4-D commands and different distances and the interactions among the three variables. Additional 4-way, repeated measures ANOVAs were also performed on the same dependent variables with format (speed, time, and speed/time), distance (3,000; 6,000; and 12,000 ft), 4-D command (10, 14, 18, and 22 kts), and quartile (1, 2, 3, 4), in order to investigate pilots' performance during different phases of the routes. Because the quartiles were based on distance (proportionate to the distance of the route), variation in pilots' taxi rates during the trials resulted in different quartile durations and unequal data sampling between quartiles. The main effects from these analyses were not calculated because the resulting quartile means would not be weighted equally in the calculation of an overall mean. Consequently, only the results of interactions among quartile, format, distance, and 4-D command were analyzed because the calculation of main effects would be inaccurate. The dependent variables were calculated in two different ways for the quartile analyses. The TOA absolute error and TOA error bias data were calculated as proportions of the route (e.g., TOA for Quartile 2 was calculated from the beginning of the route until the second data marker; equivalent to 50% of the route). SD of speed and percentage of dwell time were calculated as independent segments of the routes (e.g., Quartile 2 was calculated from the first data marker up to the second data marker). In order to obtain pilot feedback and guidance for future research, subjective data were

collected for completeness. The responses for the subjective workload questionnaires are presented in Appendix B.

Of the 18 participants, 2 pilots made a wrong turn and went off of their cleared route during one trial. Both participants stopped and contacted ground control (the experimenter). They were instructed to hold while the next trial was started and that the trial would be replaced with a new trial at the end of the block. The trials were replaced with new clearances with equivalent turns, formats, distances, and 4-D commands. A trial was also replaced for a third participant due to technical difficulties with the simulator. This trial was also replaced with an equivalent clearance.

TOA Absolute Error

TOA absolute error was calculated as the absolute value of the difference between the commanded TOA and pilots' actual TOA at the designated runway crossing at the end of each route. This measure of error magnitude was used to assess the accuracy of compliance with the 4-D commands. In order to test the hypotheses that pilots would taxi most accurately with the speed/time format and that their accuracy would improve as route distances and 4-D commanded speeds increased, a 3 (format) x 3 (distance) x 4 (4-D command) repeated measures ANOVA was conducted on TOA absolute error. Overall, with a Huynh-Feldt correction of sphericity, the analysis revealed that there was a main effect of format, $F(1,21) = 5.95$, $p < .05$, and a significant interaction between distance and command, $F(3,54) = 4.00$, $p < .01$. Pair-wise comparisons of the formats, with a Bonferroni correction, indicated that the TOA absolute error for the speed/time format ($M = 8.87$ s) was significantly lower than TOA absolute error for the speed format

($M = 15.54$ s), $t(17) = 3.72$, $p < .01$; and the time format ($M = 20.28$ s), $t(17) = 3.43$, $p < .01$. Pilots' TOA compliance was more accurate with the speed/time format than with either the speed or time formats (see Figure 8).

Simple effects analyses for the interaction between distance and 4-D command revealed a significant effect of distance for the commanded speed (CS) 10, $F(2,53) = 5.81$, $p < .01$; the CS 14, $F(2,53) = 11.08$, $p < .01$; the CS 18, $F(2,53) = 4.56$, $p < .05$; and the CS 22, $F(2,53) = 4.79$, $p < .05$. There was also a significant effect of 4-D command for short, $F(3,71) = 6.76$, $p < .01$, medium, $F(3,71) = 11.30$, $p < .01$; and long, $F(3,71) = 7.57$, $p < .01$, routes. A trend analysis revealed a significant linear by quadratic interaction between distance and 4-D command, $F(1,17) = 6.35$, $p < .05$. As shown in Figure 9, TOA absolute error increased as routes lengthened and this increase in error was larger as 4-D commanded speed was slower (i.e. the slope of this trend became more negative as 4-D commanded speed decreased).

A 3 (format) x 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was conducted on TOA absolute error. There were significant interactions among quartile and each of the other factors (see Appendix C), but further analyses focused specifically on the interaction between format and quartile in order to address the a priori hypothesis that pilots' accuracy patterns across quartiles would differ among formats. There was a significant interaction between format and quartile, with a Huynh-Feldt correction, $F(3,47) = 7.30$, $p < .01$. Based on predictions about the pattern of accuracy within trials, simple effects analyses were performed on quartile at each level of format. The analyses revealed that TOA absolute error differed between quartiles with

Figure 8. Mean time of arrival (TOA) absolute error for display formats.

Vertical lines indicate ± 1 standard error of the mean.

Figure 9. Mean time of arrival (TOA) absolute error as a function of the interaction between distance and 4-D command. 4-D command is plotted in terms of speed, but the 4-D commands were also presented as TOAs depending on the display format. Vertical lines indicate +/- 1 standard error of the mean.

the speed, $F(3,71) = 8.18$, $p < .01$; time, $F(3,71) = 4.22$, $p < .01$; and speed/time formats, $F(3,71) = 5.23$, $p < .01$. In order to determine if the shapes of the quartile trends differed between formats, separate trend analyses were conducted on each level of format. These analyses revealed that there was a significant linear trend across quartiles for the speed format, $t(68) = 4.95$, $p < .001$; indicating that TOA absolute error increased as quartile increased; and a significant quadratic trend for the time and speed/time formats, $t(68) = -3.39$, $p < .01$, and $t(68) = -3.24$, $p < .01$, respectively. This indicates that TOA absolute error increased from early to later quartiles, but that increase diminished for later quartiles (see Figure 10). Because the time and speed/time formats had quadratic trends, a 2 (format) x 3 (distance) x 4 (command) x 4 (quartile) repeated measures ANOVA was performed on TOA absolute error, without the speed format, in order to determine if the quadratic trends differed. There was still a significant interaction between format and quartile, $F(3,51) = 3.44$, $p < .05$; indicating that the quadratic trend was larger for the time format than for the speed/time format.

It was also of interest to examine the simple effects of format, which revealed that TOA absolute error differed between formats for Quartile 1, $F(2,53) = 29.89$, $p < .001$; Quartile 2, $F(2,53) = 55.83$, $p < .001$; Quartile 3, $F(2,53) = 26.33$, $p < .001$; and Quartile 4 $F(2,53) = 5.04$, $p < .01$. Despite the main effect of format showing no differences between the time and speed formats at the ends of the routes (see Figure 8 or Quartile 4 in Figure 10), TOA absolute error was higher for the time format than for the speed format during Quartile 1, $t(17) = 5.78$, $p < .001$; during Quartile 2, $t(17) = 8.52$, $p < .001$; and

Figure 10. Mean time of arrival (TOA) absolute error as a function of the interaction between format and quartile. Quartile was measured as a proportion of the distance of the taxi route. Note that the data for Quartile 4 are equivalent to those in Figure 8. Vertical lines indicate +/- 1 standard error of the mean.

during Quartile 3, $t(17) = 5.71$, $p < .001$: but the speed and speed/time formats did not differ. This analysis reveals that the speed and speed/time formats were actually more comparable, with regards to TOA accuracy during the first three quartiles, than the speed and time formats were. These results are also supported by the subjective data, which indicate that pilots rated the speed and speed/time formats as easier to use than the time format (see Appendix B, items 22 to 24).

TOA Error Bias

TOA error bias was measured as the signed difference between the commanded TOA and the pilots' actual TOA. This was calculated by subtracting the commanded TOA from the actual TOA. Thus, negative values indicate an early TOA and positive values indicate a late TOA. This measure of the error direction was analyzed in order to investigate how early or late pilots' compliance was and also indicated if they were taxiing, on average, faster or slower than the 4-D commanded speed. In order to test the hypotheses that pilots would arrive early with the speed format and late with the time format, a 3 (format) x 3 (distance) x 4 (4-D command) repeated measures ANOVA was conducted on TOA error bias. Overall, with a Huynh-Feldt correction of sphericity, the analyses revealed that there was a main effect of distance, $F(1,23) = 9.50$, $p < .01$; a main effect of 4-D command, $F(2,27) = 29.36$, $p < .001$; and a significant 2-way interaction between distance and 4-D command, $F(5,78) = 7.22$, $p < .001$. Trend analyses revealed a significant linear trend for distance, $F(1,17) = 11.56$, $p < .01$, and for 4-D command, $F(1,17) = 40.23$, $p < .001$. As seen in Figure 11 and Figure 12, these linear trends indicate

Figure 11. Mean time of arrival (TOA) error bias for route distances. TOA error bias was measured as the actual value of the difference between the commanded TOA and the actual TOA; vertical lines indicate +/- 1 standard error of the mean.

Figure 12. Mean time of arrival (TOA) error bias for 4-D command. 4-D command is plotted in terms of speed, but the 4-D commands were also presented as TOAs depending on the display format. Vertical lines indicate ± 1 standard error of the mean.

that pilots' accuracy increased with shorter route distances and decreased with slower 4-D commanded speeds.

Simple effects analyses of the interaction between distance and 4-D command revealed that TOA error bias for 4-D commands differed for short, $F(3,71) = 13.19$, $p < .001$; medium, $F(3,71) = 22.43$, $p < .001$; and long, $F(3,71) = 14.26$, $p < .001$, routes. Simple effects analyses also revealed that TOA error bias differed between route distances for CS 10, $F(2,53) = 4.65$, $p < .05$; CS 14, $F(2,53) = 8.33$, $p < .01$; and CS 18, $F(2,53) = 7.23$, $p < .01$; but not for CS 22. A trend analysis revealed a significant linear by linear interaction, $F(1,17) = 18.97$, $p < .001$. As seen in Figure 13, with the exception of CS 22, pilots arrived at the runway crossings earlier as distance increased, and across distances pilots arrived earlier as 4-D commanded speed decreased.

A 3 (format) x 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was conducted on TOA error bias. There were significant interactions among quartile and each of the other factors (see Appendix D), but further analyses focus specifically on the interaction among format, distance, and quartile in order to address the hypotheses that pilots' would arrive increasingly early with the speed format and increasingly later with the time format across quartile for medium and long distances. There was a significant 3-way interaction among format, distance, and quartile, $F(5,79) = 4.09$, $p < .01$. Based on predictions about the pattern of accuracy within trials, simple effects analyses were performed on quartiles at each level of distance within each format. Within the speed format, TOA error bias differed among quartiles for short routes, $F(3,71) = 7.48$, $p < .001$; medium routes, $F(3,71) = 11.20$, $p < .001$; and long routes,

Figure 13. Mean time of arrival (TOA) error bias as a function of the interaction between distance and 4-D command. 4-D command is plotted in terms of speed, but the 4-D commands were also presented as TOAs depending on the display format. Vertical lines indicate ± 1 standard error of the mean.

$F(3,71) = 16.01, p < .001$. Within the speed/time format, data differed among quartiles for medium routes, $F(3,71) = 5.46, p < .01$; and long routes, $F(3,71) = 7.60, p < .001$; but there were no differences between quartiles for short routes. There were no differences among quartiles within the time format (see Figure 14).

Separate, post hoc ANOVAs were conducted within the speed and speed/time formats in order to specifically assess the trends within the individual formats. A 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was conducted on TOA error bias within the speed format. There was still a significant interaction between distance and quartile, $F(3,55) = 32.09, p < .001$. The trend analysis revealed a significant quadratic by linear trend, $F(1,17) = 9.70, p < .01$; indicating that pilots arrived earlier as quartile increased, and that negative error was larger as distance increased (see Figure 14). Because there was no effect of quartile for the short distance in the speed/time format, a post hoc, 2 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was performed with the medium and long distances within the speed/time format, in order to determine if the effect of quartile differed between the two remaining distances. There was still a significant interaction between distance and quartile within the speed/time format, $F(2,37) = 6.64, p < .01$. The post hoc trend analysis also revealed a significant linear by cubic trend, $F(1,17) = 4.74, p < .05$. Pilots arrived earlier as distance increased, and that negative error increased and then reversed and decreased as quartile increased. The interaction indicates that the cubic trend of quartile was larger for the long distance than for the medium distance, as seen in Figure 14. These results reveal that pilots had a general tendency to taxi too fast and arrive early

Figure 14. Mean time of arrival (TOA) error bias as a function of the interaction among format, distance, and quartile. The panels present the interaction within the speed format, the time format, and the both format. Vertical lines indicate ± 1 standard error of the mean.

in relation to their 4-D command. In most cases, when provided with time information pilots used it to correct their error by slowing down before they reached the final runway crossing.

SD of Speed

The SD of speed was calculated as the SD of the ground speed of each trial. This variable indicated how consistent pilots' speed maintenance was around their average speed, regardless of the accuracy of their compliance. In order to test the hypotheses that pilots' speed maintenance would be more consistent with the speed and speed/time formats and with slower 4-D commands and longer route distances, a 3 (format) x 3 (distance) x 4 (4-D command) repeated measures ANOVA was conducted on the SD of speed. Overall, with a Huyhn-Feldt correction of sphericity, the analysis revealed that there was a significant main effect of format, $F(2,33) = 92.60$, $p < .001$; distance, $F(1,24) = 15.41$, $p < .001$; and 4-D command, $F(3,51) = 155.11$, $p < .001$. There was also a significant interaction between format and distance, $F(2,41) = 8.70$, $p < .001$. Pair-wise comparisons of format, with a Bonferroni correction, indicated that the SD of speed in the time format ($M = 7.23$ kts) was significantly higher than in the speed format ($M = 4.27$ kts), $t(17) = -10.96$, $p < .001$; and in the speed/time format ($M = 4.74$ kts), $t(17) = -10.75$, $p < .001$; but the speed and speed/time formats did not differ (see Figure 15). A trend analysis revealed a significant quadratic trend for distance, $F(1,17) = 15.30$, $p < .01$ (see Figure 16). It also revealed a significant linear trend for 4-D command, $F(1,17) = 264.52$, $p < .001$; which indicated that SD of speed increased with faster commanded speeds, as seen in Figure 17.

Figure 15. Mean standard deviation of speed for format. Vertical lines indicate ± 1 standard error of the mean.

Figure 16. Mean standard deviation of speed for route distance. Vertical lines indicate +/- 1 standard error of the mean.

Figure 17. Mean standard deviation of speed for 4-D command. Vertical lines indicate +/- 1 standard error of the mean.

Simple effects analyses of the interaction between format and distance revealed that SD of speed differed for route distances in the speed format, $F(2,53) = 9.70$, $p < .001$; and in the speed/time format, $F(2,53) = 3.97$, $p < .05$; but there were no differences in the time format. Because there was no effect of distance in the time format, a post hoc, 2 (format) $\times$ 3 (distance) $\times$ 4 (command) repeated measures ANOVA was performed with the speed and speed/time formats only, in order to determine if the effect of distance differed between the two formats. There was still a significant interaction between format and distance, $F(2,34) = 3.44$, $p < .05$; indicating that the quadratic trend of distance was larger with the speed format than with the speed/time format, as seen in Figure 18. For the speed and speed/time formats, pilots' compliance throughout each trial was more consistent as route distance got longer, and this increase in consistency was more pronounced in the speed format. The time format was less consistent than the speed and speed/time formats, and showed no improvement with longer distances.

A 3 (format) $\times$ 3 (distance) $\times$ 4 (4-D command) $\times$ 4 (quartile) repeated measures ANOVA was conducted on SD of speed. There were significant interactions among quartile and each of the other factors (see Appendix E), but further analyses focus specifically on the interaction between format and quartile in order to address a priori hypotheses that pilots' speed variability across quartiles would result in different patterns among formats for medium and long routes. Whereas there was no expected interaction among format, distance, and quartile, results did reveal a significant interaction between format and quartile, $F(3,59) = 8.54$, $p < .001$. Based on predictions about the pattern of accuracy within trials, simple effects analyses were performed on quartiles at each level

Figure 18. Mean standard deviation of speed as a function of the interaction between format and distance. Vertical lines indicate ± 1 standard error of the mean.

of format. The analyses revealed that SD of speed differed among quartiles with the speed, $F(3,71) = 98.48$, $p < .001$; time, $F(3,71) = 50.73$, $p < .001$; and speed/time formats, $F(3,71) = 74.08$, $p < .001$. Post hoc trend analyses were conducted at each level of format in order to determine what the shapes of the quartile patterns were. These analyses revealed that there was a significant cubic trend for the speed format, $F(1,68) = 18.93$, $p < .001$; for the time format, $F(1,68) = 11.11$, $p < .01$; and for the speed/time format, $F(1,68) = 15.50$, $p < .001$.

In order to determine if these cubic trends differed from each other, repeated measures ANOVAs were performed with two combinations of two of the three formats. A 2 (format) x 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was performed on SD of speed, with the time and speed/time formats only. This analysis revealed a significant interaction between format and quartile, $F(3,51) = 4.85$, $p < .01$; indicating that the cubic trend of quartile was larger for the speed/time format. As shown in Figure 19, for time and speed/time, SD of speed decreased between Quartile 1 and Quartile 2, and then increased as quartile increased; but the initial decrease in SD of speed was larger for the speed/time format than for the time format. Another 2 (format) x 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was performed on SD of speed, with the speed/time and speed formats only. This analysis also revealed a significant interaction between format and quartile, $F(2,38) = 8.06$, $p < .01$; indicating that the cubic trend of quartile was again larger for the speed/time format. However, in this case, the initial decrease in SD of speed is equivalent, whereas the

Figure 19. Mean standard deviation of speed as a function of the interaction between format and quartile. Vertical lines indicate ± 1 standard error of the mean.

subsequent increase of SD of speed as quartile increased was larger for the speed/time format than for the speed format.

Percent Dwell Time

Percentage of dwell time was calculated as the proportion of dwell time that pilots were looking at various display information. Total dwell time was calculated as the cumulative time during which the eye tracker detected a minimum pupil diameter of 30 mm while the eye was fixated in one location for a minimum of 0.25 seconds. Due to technical difficulties, no valid eye tracking data was collected for one participant. Consequently, analyses of dwell time information consisted of valid eye tracking data from 17 of the 18 participants. Data for 10 individual trials were replaced with the mean of the appropriate conditions because the eye calibration was lost due to shifting of the headband or technical difficulties with the eye tracker. These trials were distributed among five participants. Of the 10 trials, 6 trials were from the speed format, 1 trial was from the time format, and 3 trials were from the speed/time format. Additionally, data from an entire block (the time format) from one subject was replaced with means from the appropriate conditions due to technical difficulties.

In order to test the hypotheses that pilots would look at the speed/time display the most and that they would look at the displays proportionally more as 4-D commanded speed increased and distance decreased, a 3 (format) x 3 (distance) x 4 (4-D command) repeated measures ANOVA was conducted on the percentage of dwell time on display information. Overall, the analysis revealed that there was a significant main effect of format, $F(2,32) = 14.88$, $p < .001$; and distance (Huynh-Feldt correction), $F(1,23) =$

50.66, $p < .001$. Pair-wise comparisons of format, with a Bonferroni correction, indicated that the percentage of dwell time was significantly lower for the time format ($M = 9.6\%$) than for the speed ($M = 19.8\%$), $t(16) = 5.10$, $p < .001$; and the speed/time ($M = 19.9\%$) formats, $t(16) = 4.48$, $p < .01$ (see Figure 20). Pilots' subjective judgments that they referred to the speed information more than time information in both the speed and speed/time formats support this finding (see Appendix B, items 1 to 4). Pair-wise comparisons of distance, with a Bonferroni correction, indicated that the percentage of dwell time was significantly higher for short routes ($M = 19.3\%$) than for medium routes ($M = 15.7\%$), $t(16) = 7.12$, $p < .001$; and for long routes ($M = 14.4\%$), $t(16) = 8.07$, $p < .001$. The percentage of dwell times was also higher for medium routes than for long routes $t(16) = 4.26$, $p < .01$. A trend analysis revealed a significant quadratic trend for distance, $F(1,16) = 14.16$, $p < .01$. As seen in Figure 21, the percentage of time that pilots looked at the displays increased as distance decreased, and that increase was larger between short and medium routes than between medium and long routes.

A 3 (format) x 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was conducted on percentage of dwell time on display information. There were significant interactions among quartile and each of the other factors (see Appendix F), but further analyses focus specifically on the interaction among format, distance, and quartile in order to address the hypothesis that the pattern of display usage across quartiles would differ among formats for medium and long routes. Results revealed a significant interaction among format, distance, and quartile, $F(12,192) = 3.00$, $p < .01$. Based on predictions about the pattern of accuracy within trials, simple effects analyses

Figure 20. Mean percentage of eye dwell time for formats. Vertical lines indicate ± 1 standard error of the mean.

Figure 21. Mean percentage of eye dwell time for route distances. Vertical lines indicate +/- 1 standard error of the mean.

were performed on quartiles at each level of distance within each level of format. Simple effects analyses revealed that within the speed format, percentage of dwell time differed among quartiles for short routes, $F(3,67) = 12.68, p < .001$; medium routes, $F(3,67) = 5.65, p < .01$; and long routes, $F(3,67) = 3.70, p < .05$. Within the time format, percentage of dwell time differed among quartiles for short routes, $F(3,67) = 6.38, p < .01$, medium routes, $F(3,67) = 9.48, p < .001$, and long routes, $F(3,67) = 19.36, p < .001$. Within the speed/time format, percentage of dwell time differed among quartiles for short routes, $F(3,67) = 3.90, p < .05$; medium routes, $F(3,67) = 3.80, p < .05$; and long routes, $F(3,67) = 5.21, p < .01$.

Separate, post hoc ANOVAs were conducted within the different levels of format in order to specifically assess the trends within the individual formats. A 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was performed on percentage of dwell time within the speed format. There was still a significant interaction between distance and quartile, with a Huynh-Feldt correction, $F(4,67) = 13.06, p < .001$. The trend analysis revealed a significant quadratic by quadratic trend, $F(1,16) = 7.81, p < .05$; indicating that percentage of dwell time decreased as distance increased from short to medium routes and as quartile increased from Quartile 1 to Quartile 2, but that decrease in percentage of dwell time leveled off between medium and long routes and between the remaining three quartiles, as seen in Figure 22. There was also a significant linear by cubic trend, $F(1,16) = 15.70, p < .01$; which appears to indicate that instead of leveling off, the percentage of dwell time increased near the end of the route with the short distance. A 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA

Figure 22. Mean percentage of eye dwell time as a function of the interaction among format, distance and quartile. The panels present the interaction within the speed format, the time format, and the speed/time format. Vertical lines indicate +/- 1 standard error of the mean.

was performed on percentage of dwell time within the speed/time format. There was also a significant interaction between distance and quartile, with a Huyhn-Feldt correction, $F(4,60) = 5.30$, $p < .01$. The trend analysis revealed a significant linear by cubic trend, $F(1,16) = 23.90$, $p < .001$; indicating that percentage of dwell time increased as distance decreased, and the pattern of that increase across quartiles was such that percentage of dwell time decreased between Quartile 1 and Quartile 2, and then increased as the remaining quartiles increased (see Figure 22). A 3 (distance) $\times$ 4 (command) $\times$ 4 (quartile) repeated measures ANOVA was conducted on percentage of dwell time within the time format, which revealed that there was no significant interaction between distance and quartile (see Figure 22).

It was also of interest to compare the percentage of dwell time on speed and time information within the speed/time format in order to determine how pilots used the information when presented together. Pair-wise comparisons revealed that percentage of dwell time on speed information in the speed/time format ($M = 12.8\%$) was significantly lower than the speed format ($M = 19.8\%$), $t(16) = 3.72$, $p < .05$. The percentage of dwell time on time information ($M = 7.4\%$) was lower than the percentage of time on speed information ($M = 12.8\%$) in the speed/time format, $t(16) = 2.69$, $p < .05$. However, percentage of dwell time did not differ between time information in the speed/time format and the time format, as seen in Figure 23.

In order to investigate these differences within the trials, a 2 (format) $\times$ 3 (distance) $\times$ 4 (4-D command) $\times$ 4 (quartile) repeated measures ANOVA was conducted on percentage of dwell time, which compared the speed format with the speed

Figure 23. Comparison of the mean percentage of eye dwell time, between the speed and time formats and the speed and time information within the speed/time format. Vertical lines indicate ± 1 standard error of the mean.

information in the speed/time format only. There was a significant interaction between format and quartile, with a Huynh-Feldt correction, $F(2,24) = 4.52$, $p < .05$. Trend analyses revealed a significant linear trend for speed information in the speed/time format, $t(64) = -2.85$, $p < .01$; and a significant quadratic trend for the speed format, $t(64) = 3.33$, $p < .01$. Whereas percentage of dwell time in the speed format decreased at the beginning of the trial and then leveled off, in the speed/time format the percentage of dwell time for the speed information declined at the beginning of the routes and continued to decline slightly near the end of the routes as well, as seen in Figure 24.

A 2 (format) x 3 (distance) x 4 (4-D command) x 4 (quartile) repeated measures ANOVA was also conducted on percentage of dwell time, which compared the time format with the time information in the speed/time format only. There was no significant interaction between format and quartile, and the trend analyses revealed a significant quadratic trend for time information in the speed/time format and in the time format, $t(64) = 3.98$, $p < .001$; and $t(64) = 2.93$, $p < .01$, respectively. Regardless of format, when using time information, pilots' percentage of dwell time increased as quartile increased, and that change got larger near the end of the routes (see Figure 24).

Overall, results indicate that pilots' TOA absolute error was the lowest with the speed/time format, but when analyzed throughout the trials, performance with the speed format was equivalent to the speed/time format during the first three quartiles and did not diverge until the ends of the routes. Pilots also taxied more accurately with shorter routes and faster 4-D commanded speeds. TOA error bias data revealed that pilots arrived equally early regardless of format, but arrived earlier with slower 4-D commanded speeds

Figure 24. Comparison of the mean percentage of eye dwell time, among the speed format (SF), the time format (TF), and the speed information (S/T) and time information (T/S) within the speed/time format as a function of the interaction with quartile. Vertical lines indicate ± 1 standard error of the mean.

and longer distances. The results also revealed that pilots taxied the least consistently with the time format and as 4-D commanded speed increased and as distance decreased. Finally, percentage of dwell time data indicate that pilots looked at the speed and speed/time displays more than the time display. Additionally, quartile analyses revealed that pilots' patterns of display usage when given both speed and time information were equivalent to the patterns of usage with the individual speed and time formats in that they looked at the speed information primarily at the beginning of the route and looked at the time information primarily near the end of the route.

Discussion

Compliance Accuracy

Results supported the hypothesis that pilots' compliance would be most accurate when taxiing with the speed/time format. However, the prediction that pilots' compliance with the speed format would be more accurate than with the time format was not supported. Results did not support the hypothesis that there would be an interaction among format, distance, and quartile, but the predictions about quartile patterns differing among formats were accurate. Figure 10 reveals that pilots did accumulate error throughout the speed format trials, whereas with the time and speed/time formats pilots accumulated error and then corrected that error, typically by slowing down, near the end of the trials. Despite the main effect of format showing no difference between pilots' accuracy with the speed and time formats at the end of the routes, the interaction between format and quartile indicates that the speed format was more accurate than the time format during the first three quartiles. Not only were pilots more accurate with the speed

format than the time format during the first three quartiles, but they also rated the speed format as easier to use than the time format as well (see Appendix B, items 22 to 23). It was only at the end of the trials that the combination of the quadratic trend of the time format and the increasing linear trend of the speed format led to equivalent error with the two formats. The equivalent compliance accuracy between the speed and speed/time formats during the first three quartiles indicates that even when presented with the speed/time format, pilots likely used the speed information as a guidance indicator during those taxi phases and incorporated the time information for conformance monitoring primarily near the end of the route. As seen in Figure 10, pilots' compliance with the speed/time format demonstrated high accuracy as in the speed format and error correction near the end of the route as in the time format. This indicates that presenting both speed, as a guidance indicator, and time information, for conformance monitoring, did produce a redundancy gain. Incorporating both speed and time information in more advanced future displays could potentially facilitate general accuracy throughout the clearance and focused error correction near critical locations or destinations.

Results partially supported the hypothesis that there would be an interaction between distance and 4-D command for TOA absolute error. The linear by quadratic trend indicates that pilots' compliance accuracy did decline as distance increased, and that decline was more pronounced as 4-D commanded speed decreased. When taxiing short routes, pilots were most accurate and did not experience the more drastic declines in accuracy when given slower 4-D commanded speeds, as they did with longer routes. This interaction can be explained by previous research, which indicates that people

estimate slower speeds less accurately than faster speeds, and that people estimate long durations less accurately than short durations. As hypothesized, shorter routes provided less distance and time during which estimation error could accumulate. Such an interaction between route distance and commanded speed could guide the definition of performance limitations in the presence of certain clearance characteristics. Dynamic algorithms could incorporate this interaction to set limitations on one variable when the other variable falls outside certain parameters.

In addition to investigating the degree to which pilots were accurate at complying with 4-D commands, determining the bias of that compliance provided a more complete picture of their performance. Results for TOA error bias did not support the hypothesis that pilots would tend to arrive late with the time format and early with the speed format. Pilots arrived equally early with all three formats. The linear trends of distance and 4-D command did support the hypotheses that pilots would arrive earlier as distance increased, and as 4-D commanded speed decreased. The results also confirmed the prediction that distance would interact with 4-D command and pilots would arrive earlier as 4-D commanded speed decreased, but that pattern would not be as pronounced for short routes. It appears that pilots had a tendency to arrive early, possibly because arriving early could be easily corrected by slowing down, but arriving late could be impossible to correct if the error was detected too late. However, it is also possible that the simulated environment was not as realistic as taxiing in the real world, causing pilots to taxi at speeds faster than they would feel comfortable with in a real aircraft or causing them to disregard the 4-D command because they found it excessively slow.

Additionally, whereas pilots arrived earlier with longer distances for the slower 4-D commanded speeds, they arrived equally late regardless of distance for CS 22. This pattern might indicate that taxiing at an average of 22 kts was too difficult and pilots had to attempt to catch up when complying with that 4-D command. If pilots taxied the simulated aircraft faster than they would a real aircraft, it is likely that they might find it even more difficult to taxi at an average of 22 kts in a real world situation.

Quartile analyses for TOA error bias partially supported the hypothesis that there would be an interaction among format, distance, and quartile. As predicted, with the speed format pilots taxied increasingly early as distance and quartiles increased, whereas with the speed/time format pilots taxied increasingly early as distance and quartile increased, but then slowed down to correct their early arrival at the end of the route. Also, as predicted, there was no effect of quartile for short routes in the speed/time format. Contrary to the hypothesis was the result that there was no effect of quartile for any distances in the time format. It appears that the trend of the time format was similar to the speed/time format, but pilots' performance was so variable that there was no clear pattern for the quartiles. These results indicate that regardless of the differences between accuracy quartile patterns for the different formats, the decline in accuracy was due to pilots taxiing too fast (the result of underestimating their speed and/or overestimating elapsed time) and consequently arriving too early relative to the 4-D command. Unlike the speed format, the speed/time format enabled pilots to assess how early they were near the end of the routes, at which point they would slow down to attempt to correct their error. Because pilots' accuracy error was primarily due to early arrivals, this might be a

compliance issue that SMSs could adjust for by giving faster/shorter 4-D commands and shorter taxi clearances.

Consistency

As predicted, pilots' speed maintenance was less consistent with the time format than with the speed or speed/time formats. The presence of control information in the speed and speed/time formats enabled pilots to taxi within a smaller range of speeds around their average speed. With the time format pilots did not have any feedback or explicit guidance regarding control information, making their speed maintenance less consistent. Results also supported the hypotheses that SD of speed would increase as distance decreased and as 4-D commanded speed increased. As hypothesized, the linear decrease in variability of speed with slower 4-D commanded speeds is likely due to the ability of pilots to taxi through turns without slowing down. Whereas with faster 4-D commanded speeds they would have to slow down to turn and then overcompensate with faster than average speeds. The unexpected interaction between format and distance revealed that pilots taxied more consistently as distance increased in the speed and speed/time formats, but in the time format pilots' speed control was equally less consistent regardless of distance. It appears that the presence of speed information helps pilots regulate their taxi speed throughout the route, and the increase in consistency with longer distances was the result of more maintenance time between the initial spool up and the final runway crossing at the end. In order to aid dynamic algorithms in prediction of traffic movement, the use of speed and time based commands could facilitate more consistent speed control. Because it might not always be feasible to give a longer taxi

clearance, the larger speed variability associated with shorter taxi clearance could also be counterbalanced with the use of slower 4-D commanded speeds.

The prediction that there would be an interaction among format, distance, and quartile was not supported, but the interaction between format and quartile partially supported the hypotheses about the quartile patterns for the different formats. The cubic trends revealed that speed variance was large at the beginning of the routes regardless of format because the initial speed/engine spool up required pilots to overcompensate in order to reach their 4-D command. The interaction resulted from a larger increase in consistency after the spool up with the speed and speed/time formats than with the time format. Additionally, with the time and speed/time formats pilots taxied less consistently towards the end of the routes when they corrected their error, whereas with the speed format their speed maintenance remained consistent because there was no time information to cue them to correct their error. Thus, as predicted, pilots' speed maintenance was least consistent at the beginning of the routes. Contrary to predictions, pilots' speed maintenance in the time format was also highly variable at the beginning of the routes, but did increase again at the end of the routes when time became more salient for error correction. As predicted, pilots' speed maintenance with the speed/time format appeared to be a combination of the speed and time trends, with the least consistency near the beginnings and ends of the routes. Given that the increase in speed variability near the beginning and end of the routes with the speed/time format was due to error correction, a predictive algorithm could incorporate that information into a profile of taxi phases and increase the error buffer in those areas accordingly.

Strategy

As expected, pilots looked at the time display less than the speed and speed/time displays. However, they did not look at the speed/time display more than the speed display. Although pilots did not use the speed/time format more than the speed format, they did use each of these two formats twice as much as the time format. Subjective data indicate at least half of the pilots stated that they found themselves looking at the speed information when they should have paying attention to the outside world during both the speed and speed/time formats (see Appendix B, items 11 to 14). Given that the displays were simplistic text displays, this problem could potentially be alleviated through the use of a more advanced display. However, more research is warranted in order to determine if pilots experience cognitive tunneling when provided with speed information during taxi operations. Results supported the prediction that pilots would look at the displays proportionally less as distance increased. This is likely due to the longer distances having more maintenance time during which it was not necessary to check the displays as often. The prediction of a main effect of 4-D command was not supported. As expected, the interaction among format, distance, and quartile revealed different quartile and distance patterns for the three formats. Within the speed format, pilots looked at the display the most at the beginning of the trial while correcting their initial spool-up error. Contrary to the prediction that this pattern would be less pronounced for short routes, it was actually more pronounced for short routes in that the initial decrease in display usage was larger than for medium and long routes. Within the time format, pilots looked at the display consistently during the first three quartiles, but used the display most near the end of the

routes while they corrected their error before the runway crossing. This pattern was consistent regardless of distance. With the speed/time format pilots demonstrated the predicted combination of usage from the speed and time formats in that they looked at the display frequently at the beginning and at the end of the routes when they would have been correcting error.

In order to confirm that the increases in display usage were due to an actual combination of the speed and time patterns, analyses within the speed/time format examined the patterns in more detail. Comparisons of the main effect of format confirmed that pilots used the speed and time information proportionally within the speed/time format as they did in the speed and time formats. Comparing the quartile information within the speed/time format clearly revealed that when the information was displayed simultaneously, pilots used the speed information most at the beginning of the routes and used the time information most at the end of the routes, just as they did in the speed and time formats. The pattern for the use of speed information in the speed/time format was slightly different than in the speed format, in that usage of speed information in the speed/time format decreased near the end of the trial. However, the pattern for time information was the same within the speed/time format as in the time format. Both pieces of information appear to be useful for complying with 4-D commands and produce a redundancy gain by presenting relevant information in complementary formats that are useful during different taxi phases. Future displays could incorporate these patterns by providing speed information throughout a 4-D clearance, but only providing time information during relevant phases of taxi (right before a checkpoint or destination).

Conclusions

In order to taxi precisely, pilots must have relevant display information to help them comply with 4-D taxi commands. Combining speed and time clearances provides redundant information about movement and enables pilots to switch between relevant information during appropriate phases of taxiing. Utilizing the high taxi accuracy associated with speed information and the error correction facilitated by time information could enable pilots to safely comply with coordinated runway crossings. The consistent speed maintenance also associated with speed information could enable SMSs to accurately predict and coordinate traffic flow for surface operations. In addition to providing both speed and time information, shorter taxi clearances also promote more precise taxiing when paired with average taxi speeds and durations.

More research is required before this study can be generalized to real world situations. Given that the simulation provided no proprioceptive feedback and a smaller than average amount of airport traffic, pilots might have taxied faster and less cautiously than they would at a real airport. Additionally, the display was intentionally simplistic in order to capture the baseline nature of the information. The display was also presented on a HUD, and might not be as salient or accessible when displayed differently. However, these results could be useful in guiding the design of future advanced 4-D taxi displays. Designers might consider including explicit speed information for control and time information for implicit feedback. Additionally, when designing within the limitations of pilots' compliance with 4-D commands, clearances should likely remain

shorter in order to avoid error accumulation. If incorporated appropriately, such displays could become an invaluable tool for efficient taxi operations.

Implications for system design

- Providing speed and time information produces more accurate and consistent compliance.
- Shorter routes produce more accurate compliance.
- Faster 4-D commands produce more accurate compliance.
- Speed information is relevant during initial taxi phases for consistent speed control.
- Time information is relevant during critical checkpoints for precise error correction.

References

- Awe, C. A. & Johnson, W. W. (1993). The effect of scene content on speed, time, and distance perception. *International Symposium on Aviation Psychology*, 7, 281-285.
- Brown, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics*, 38, 115-124.
- Cheng, V.H.L., & Foyle, D.C. (2002). Automation tools for enhancing ground-operation situation awareness and flow efficiency. Proceedings of the AIAA Guidance, Navigation, and Control Conference, Paper AIAA 2002-4856.
- Cheng, V. H. L., Sharma, V., & Foyle, D. C. (2001). A study of aircraft taxi performance for enhancing airport surface traffic control. *IEEE Transactions on Intelligent Transportation Systems*, 2(2), 39 - 54.
- Coren, S., Ward, L. M., & Enns, J. T. (2004). Sensation and Perception (6th ed.). New Jersey: John Wiley & Sons, Inc.
- Denton, G. G. (1966). A subjective scale of speed when driving a motor vehicle. *Ergonomics*, 9, 203-210.
- Foyle, D. C., McCann, R. S., & Shelden, S. G. (1995). Attentional issues with superimposed symbology: Formats for scene-linked displays. In R. S. Jensen, & L. A. Rakovan (Eds.), Proceedings of the Eighth International Symposium on Aviation Psychology, 98-103. Columbus, Ohio: Ohio State University.
- Glass, B. J., & Gawdiak, Y. (1997). Integrated navigation applications in airport surface traffic management. *Saint Petersburg International Conference on Integrated Navigation Systems*, 4, 105-113.
- Hecker, P., Doehler, H. U., Korn, B., & Ludwig, T. (2001). Extending enhanced vision capabilities by integration of advanced surface movement guidance and control systems (A-SMGCS). *Proceedings of SPIE, the International Society for Optical Engineering*, 4363, 58-66.
- Jackson, M. & Selcon, S. J. (1997). A parallel distributed processing model of redundant information integration. *Engineering psychology and cognitive ergonomics: Vol. 2. Job design and product design (pp. 193-200); Proceedings of the International Conference*, Stratford upon Avon.

- Leiser, D., Stern, E., & Meyer, J. (1991). Mean velocity and total time estimation: Effects of order and proportions. *Journal of Environmental Psychology*, 11, 347-358.
- Recarte, M. A., & Nunes, L. M. (1996). Perception of speed in an automobile: Estimation and production. *Journal of Experimental Psychology: Applied*, 2, 291-304.
- Selcon, S. J. (1999). Using redundancy in the design of time-critical warnings: A theory-driven approach. *Human factors in auditory warnings* (pp. 201-224). Brookfield, VT: Ashgate.
- Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall Inc.
- Zakay, D. (1992). On prospective time estimation, temporal relevance and temporal uncertainty. *Time, Action, and Cognition*. Netherlands: Kluwer Academic Publishers.
- Zakay, D., Nitzan, D., & Glicksohn, J. (1983). The influence of task difficulty and external tempo on subjective time estimation. *Perception & Psychophysics*, 34, 451-456.

Appendix A

Post-Block Questionnaires for the Speed, Time, and Speed/Time Formats

Please take a moment to recall the display format that you taxied with during the previous block of trials and then answer the questions about this format.

The **speed format** display consisted of a current ground speed indicator in knots and a commanded average speed: picture

Please circle one choice for each question:

1. During the **speed** format trials, how often did you refer to the **speed** information?

Rarely	Seldom	Sometimes	Frequently	Most of the time
1	2	3	4	5

2. During the **speed** format trials, how easy/difficult was it to maintain an **average speed**?

Very easy	Easy	Neither easy nor difficult	Difficult	Very difficult
1	2	3	4	5

3. Did you ever find yourself focused on the **speed** information when you should have been paying attention to the external runway/taxiway or other external information?

Yes No

4. During the **speed** format trials, how often did you refer to the **map** information?

Rarely	Seldom	Sometimes	Frequently	Most of the time
1	2	3	4	5

5. Did you ever find yourself focused on the **map** information when you should have been paying attention to the external runway/taxiway or other external information?

Yes No

6. Please rate your **overall workload** in the speed format. In doing so, please consider how demanding was the task of taxiing while complying with your speed command.

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

7. Please rate the **overall ease of use** of the speed format display:

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

8. Please rate the degree to which the speed format display **enables compliance**:

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

Please take a moment to recall the display format that you taxied with during the previous block of trials and then answer the questions about this format.

The **time format** display consisted of elapsed time in minutes and seconds and a commanded time of arrival: picture.

Please circle one choice for each question:

1. During the **time** format trials, how often did you refer to the **time** information?

Rarely	Seldom	Sometimes	Frequently	Most of the time
1	2	3	4	5

2. During the **time** format trials, how easy/difficult was it to arrive at a **commanded time**?

Very easy	Easy	Neither easy nor difficult	Difficult	Very difficult
1	2	3	4	5

3. Did you ever find yourself focused on the **time** information when you should have been paying attention to the external runway/taxiway or other external information?

Yes No

4. During the **time** format trials, how often did you refer to the **map** information?

Rarely	Seldom	Sometimes	Frequently	Most of the time
1	2	3	4	5

5. Did you ever find yourself focused on the **map** information when you should have been paying attention to the external runway/taxiway or other external information?

Yes No

6. Please rate your **overall workload** in the time format. In doing so, please consider how demanding was the task of taxiing while complying with your time command.

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

7. Please rate the **overall ease of use** of the time format display:

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

8. Please rate the degree to which the time format display **enables compliance**:

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

Please take a moment to recall the display format that you taxied with during the previous block of trials and then answer the questions about this format.

The **“both” format** display consisted of both the speed display and the time display. On the left side current ground speed and commanded speed were displayed and on the right side elapsed time and commanded time of arrival were displayed: picture

Please circle one choice for each question:

1. During the **“both”** format trials, how often did you refer to the **speed** information?

Rarely	Seldom	Sometimes	Frequently	Most of the time
1	2	3	4	5

2. During the **“both”** format trials, how easy/difficult was it to maintain an **average speed**?

Very easy	Easy	Neither easy nor difficult	Difficult	Very difficult
1	2	3	4	5

3. Did you ever find yourself focused on the **speed** information when you should have been paying attention to the external runway/taxiway or other external information?

Yes	No
-----	----

4. Please rate your **overall workload** in the **“both”** format. In doing so, please consider how demanding was the task of taxiing while complying with your **speed** command.

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

5. During the **“both”** format trials, how often did you refer to the **time** information?

Rarely	Seldom	Sometimes	Frequently	Most of the time
1	2	3	4	5

6. During the **“both”** format trials, how easy/difficult was it to arrive at a **commanded time**?

Very easy	Easy	Neither easy nor difficult	Difficult	Very difficult
1	2	3	4	5

7. Did you ever find yourself focused on the **time** information when you should have been paying attention to the external runway/taxiway or other external information?

Yes No

8. During the “**both**” format trials, how often did you refer to the **map** information?

Rarely	Seldom	Sometimes	Frequently	Most of the time
1	2	3	4	5

9. Did you ever find yourself focused on the **map** information when you should have been paying attention to the external runway/taxiway or other external information?

Yes No

10. Please rate your **overall workload** in the “both” format. In doing so, please consider how demanding was the task of taxiing while complying with your **time** command.

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

11. Please rate the **overall ease of use** of the “both” format display:

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

12. Please rate the degree to which the “both” format display **enables compliance**:

Very Low	Low	Neutral	High	Very High
1	2	3	4	5

Post Study Questions

Please take a moment to recall the three different display formats that you experienced/taxied with during the experiment and then answer the questions about these formats.

The **speed format** display consisted of a current ground speed indicator in knots and a commanded average speed.

The **time format** display consisted of elapsed time in minutes and seconds and a commanded time of arrival.

The **“both” format** display consisted of both the speed display and the time display. On the left side current ground speed and commanded speed were displayed and on the right side elapsed time and commanded time of arrival were displayed.

9) Please rank the formats (speed, time, both) in order of preference:

Choice # 1: _____

Choice # 2: _____

Choice # 3: _____

10) Can you think of any situations or taxi environments where one command format may be preferred over another? Please explain:

11) Can you think of any situations or taxi environments where one command format might have more problems compared with another? Please explain:

Appendix B

Response Frequencies for Post Block and Post Study Questionnaires

Item	Questions	Response Frequencies					Most of the time
		Rarely	Seldom	Sometimes	Frequently		
1.	During the speed format trials, how often did you refer to the speed information?	0	0	0	11		7
2.	During the time format trials, how often did you refer to the time information?	0	1	1	12		4
3.	During the “ both ” format trials, how often did you refer to the speed information?	0	0	0	13		5
4.	During the “ both ” format trials, how often did you refer to the time information?	0	0	9	9		0
5.	During the speed format trials, how often did you refer to the map information?	1	6	6	5		0
6.	During the time format trials, how often did you refer to the map information?	1	0	2	14		1
7.	During the “ both ” format trials, how often did you refer to the map information?	1	2	8	7		0
		Very easy	Easy	Neither easy nor difficult	Difficult	Very difficult	
8.	During the speed format trials, how easy/difficult was it to maintain an average speed ?	1	8	6	3		0
9.	During the time format trials, how easy/difficult was it to arrive at a commanded time ?	0	0	4	11		3
10.	During the “ both ” format trials, how easy/difficult was it to maintain an average speed ?	3	5	7	3		0

Item	Questions	Response Frequencies	
		No	Yes
11.	Did you ever find yourself focused on the speed information when you should have been paying attention to the external runway/taxiway or other external information?	8	10
12.	Did you ever find yourself focused on the time information when you should have been paying attention to the external runway/taxiway or other external information?	12	6
13.	(During the "both" format trials) Did you ever find yourself focused on the speed information when you should have been paying attention to the external runway/taxiway or other external information?	9	9
14.	(During the "both" format trials) Did you ever find yourself focused on the time information when you should have been paying attention to the external runway/taxiway or other external information?	13	5
15.	(During the speed format trials) Did you ever find yourself focused on the map information when you should have been paying attention to the external runway/taxiway or other external information?	16	2
16.	(During the time format trials) Did you ever find yourself focused on the map information when you should have been paying attention to the external runway/taxiway or other external information?	14	4
17.	(During the "both" format trials) Did you ever find yourself focused on the map information when you should have been paying attention to the external runway/taxiway or other external information?	15	3

Item	Questions	Response Frequencies				
		Very low	Low	Neutral	High	Very high
18.	Please rate your overall workload in the speed format. In doing so, please consider how demanding was the task of taxiing while complying with your speed command:	0	8	7	3	0
19.	Please rate your overall workload in the time format. In doing so, please consider how demanding was the task of taxiing while complying with your time command:	0	0	5	11	2
20.	Please rate your overall workload in the “both” format. In doing so, please consider how demanding was the task of taxiing while complying with your speed command:	0	3	9	6	0
21.	Please rate your overall workload in the “both” format. In doing so, please consider how demanding was the task of taxiing while complying with your time command:	1	6	4	7	0
22.	Please rate the overall ease of use of the speed format display:	0	1	3	10	4
23.	Please rate the overall ease of use of the time format display:	3	6	2	7	0
24.	Please rate the overall ease of use of the “both” format display:	0	0	4	11	3
25.	Please rate the degree to which the speed format display enables compliance :	0	0	5	6	7
26.	Please rate the degree to which the time format display enables compliance :	2	6	7	2	1
27.	Please rate the degree to which the “both” format display enables compliance :	0	1	2	13	2

Item	Questions	Response Frequencies		
		First	Second	Third
	Please rank the three formats in order of preference:			
28.	Speed Format	1	15	2
29.	Time Format	0	2	16
30.	Speed/time format	17	1	0

Appendix C

Analysis of Variance with Quartiles for TOA Absolute Error

Source	SS	df	MS	F
Format (F) x Quartile (Q)	16159.45	3	5833.49	7.30**
Distance (D) x Q	8251.69	4	2293.82	6.08**
Command (C) x Q	19845.00	4	5068.34	6.52**
F x D x Q	5761.43	4	1439.98	2.24
F x C x Q	13794.12	4	3915.01	3.21*
D x C x Q	8317.91	4	1977.10	1.99
F x D x C x Q	13943.06	4	3890.91	1.80
Error F x Q	37657.20	47	799.65	
Error D x Q	23068.74	61	377.22	
Error C x Q	51742.35	67	777.34	
Error F x D x Q	43685.60	68	642.27	
Error F x C x Q	72955.05	60	1218.00	
Error D x C x Q	70901.45	72	991.34	
Error F x D x C x Q	131960.23	61	2166.15	

Note. Sphericity corrected with Huyhn-Feldt adjustment.

* $p < .05$. ** $p < .01$.

Appendix D

Analysis of Variance with Quartiles for TOA Error Bias

Source	SS	df	MS	F
Format (F) x Quartile (Q)	11027.32	3	3619.77	5.71**
Distance (D) x Q	5746.91	3	1854.21	4.36**
Command (C) x Q	15980.14	4	4476.15	6.77**
F x D x Q	8104.16	5	1744.11	4.09**
F x C x Q	9256.05	5	1927.56	2.45*
D x C x Q	11212.71	8	1437.92	3.75**
F x D x C x Q	12716.24	10	1223.67	2.49**
F x Q Error	32845.37	52	634.22	
D x Q Error	22393.93	53	425.02	
C x Q Error	40105.00	61	660.81	
F x D x Q Error	33678.82	79	426.36	
F x C x Q Error	64235.37	82	786.88	
D x C x Q Error	50890.05	133	383.89	
F x D x C x Q Error	86671.55	177	490.61	

Note. Sphericity corrected with Huyhn-Feldt adjustemnt.

*p < .05. **p < .01.

Appendix E

Analysis of Variance with Quartiles for SD of Speed

Source	SS	df	MS	F
Format (F) x Quartile (Q)	208.90	3	60.05	8.54**
Distance (D) x Q	726.10	5	150.37	38.14**
Command (C) x Q	700.00	7	94.51	25.45**
F x D x Q	44.92	9	4.79	1.26
F x C x Q	117.19	11	10.33	2.37**
D x C x Q	100.15	17	5.75	2.25**
F x D x C x Q	101.07	22	4.57	1.07
F x Q Error	415.81	59	7.03	
D x Q Error	323.64	82	3.94	
C x Q Error	607.16	159	3.81	
F x D x Q Error	467.53	126	3.71	
F x C x Q Error	840.75	193	4.36	
D x C x Q Error	757.36	296	2.56	
F x D x C x Q Error	1599.91	376	4.25	

Note. Sphericity corrected with Huyhn-Feldt adjustemnt.

*p < .05. **p < .01.

Appendix F

Analysis of Variance with Quartiles for Percent Dwell Time

Source	SS	df	MS	F
Format (F) x Quartile (Q)	0.148	3	0.052	23.80**
Distance (D) x Q	0.023	3	0.007	11.10**
Command (C) x Q	0.009	9	0.001	3.94**
F x D x Q	0.010	8	0.001	3.00**
F x C x Q	0.006	11	0.001	1.10
D x C x Q	0.007	11	0.001	1.69
F x D x C x Q	0.010	20	0.000	1.10
F x Q Error	0.100	46	0.002	
D x Q Error	0.034	52	0.001	
C x Q Error	0.054	129	0.000	
F x D x Q Error	0.037	136	0.000	
F x C x Q Error	0.081	172	0.000	
D x C x Q Error	0.068	169	0.000	
F x D x C x Q Error	0.141	315	0.000	

Note. Sphericity corrected with Huyhn-Feldt adjustemnt.

*p < .05. **p < .01.

Appendix G

Human Subjects Internal Review Board Approval

**San José State
UNIVERSITY**

**Office of the Academic
Vice President**
Academic Vice President
Graduate Studies and Research
One Washington Square
San Jose, CA 95192-0025
Voice: 408-283-7500
Fax: 408-924-2477
E-mail: gradstudies@sjsu.edu
http://www.sjsu.edu

To: Jennifer Williams
680 Garland Avenue, #4
Sunnyvale, CA 94086

From: Pam Stacks, *Pam Stacks*
Interim AVP, Graduate Studies & Research

Date: December 9, 2004

The Human Subjects-Institutional Review Board has approved your request to use human subjects in the study entitled:

"Comparing Speed and Time Command Formats During Taxi Operations."

This approval is contingent upon the subjects participating in your research project being appropriately protected from risk. This includes the protection of the anonymity of the subjects' identity when they participate in your research project, and with regard to all data that may be collected from the subjects. The approval includes continued monitoring of your research by the Board to assure that the subjects are being adequately and properly protected from such risks. If at any time a subject becomes injured or complains of injury, you must notify Pam Stacks, Ph.D. immediately. Injury includes but is not limited to bodily harm, psychological trauma, and release of potentially damaging personal information. This approval for the human subjects portion of your project is in effect for one year, and data collection beyond December 9, 2005 requires an extension request.

Please also be advised that all subjects need to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate, or withdrawal will not affect any services that the subject is receiving or will receive at the institution in which the research is being conducted.

If you have any questions, please contact me at (408) 924-2480.

cc: Kevin Jordan

The California State University:
Chico State's Office
Humboldt, Channel Islands, Chico
Dominatez Hills, Fresno, Fullerton
Hayward, Humboldt, Long Beach
Los Angeles, Maritime Academy
Monterey Bay, Northridge, Pomona
Sacramento, San Bernardino, San Diego
San Francisco, San Jose, San Luis Obispo
San Marcos, Stanislaus, Stanislaus