Aircraft Landing Lights Enhance Runway Traffic Safety

DOT/FAA/CT-TN05/18

Federal Aviation Administration
William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

Karen Buondonno, FAA Simulation and Analysis Group
Kimberlea Bender, Titan Corporation
Nicole Racine, Titan Corporation
Edmundo Sierra, Titan Corporation

March, 2005

DOT/FAA/CT-TN05/18

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/CT-TN05/18	**2. Government Accession No.**	**3. Recipient's Catalog No.**

4. Title and Subtitle Aircraft Landing Lights Enhance Runway Traffic Safety (AL^{2}ERTS) Final Report	**5. Report Date** March 2005
	6. Performing Organization Code ACB-330
7. Author(s) Karen Buondonno (FAA Simulation and Analysis Group), and Kimberlea Bender, Nicole Racine, Edmundo Sierra (Titan Corporation)	**8. Performing Organization Report No.** DOT/FAA/CT-TN05/18
9. Performing Organization Name and Address Federal Aviation Administration, ACB 330 William J. Hughes Technical Center Atlantic City International Airport, NJ 08405	**10. Work Unit No. (TRAIS)**
	11. Contract or Grant No.
12. Sponsoring Agency Name and Address Federal Aviation Administration Office of Runway Safety and Operational Services 490 L'Enfant Plaza Suite 725 Washington, DC 20024	**13. Type of Report and Period Covered** Technical Note
	14. Sponsoring Agency Code Office of Runway Safety and Operational Services

15. Supplementary Notes

16. Abstract

The Federal Aviation Administration Office of Runway Safety and Operational Services formed a simulation team to investigate the safety effects of standardizing the use of aircraft landing lights in the airport environment. Specifically, the simulation team explored the procedural use of landing lights as a direct message to other pilots indicating that aircraft were cleared to depart. Thirty-two pilots participated in the study as either the Captain or First Officer of a B747-400 simulator crew. The simulator crews were divided into two groups or crews. Each crew flew either a set of 16 scenarios in an environment with a standardized use of landing lights or scenarios using current practices. In four of the scenarios in each environment, a confederate aircraft made an error that resulted in a runway incursion (RI) that could have resulted in an accident with the B747-400 simulator if not detected by the subject crews. Multidimensional measures of RI severity and situation awareness (SA) were made after each scenario. In general, the pattern of results suggest that standardizing the use of aircraft landing lights to indicate that aircraft were cleared to depart prevented or reduced the severity of RIs or accidents, and increased pilot SA. The data shows that crews in the standard condition held-short more frequently, generally experienced less severe incursions, initiated a response to RIs significantly faster, used the landing lights effectively as a first cue, and unanimously felt that safety was increased because of the standardized procedures. Further studies are recommended to determine the effects of other factors such as consistency of the message, message conspicuity, and effects of the message on other human system elements. Evaluating alternatives such as pulse lighting and the potential value of cues to Air Traffic Control may reveal additional benefits.

17. Key Words AL^{2}ERTS, Aircraft Landing Lights, Aircraft Exterior Lights, Standard Operating Procedures (SOP), Runway Incursion, Virtual Reality Simulation, Real-Time Simulation	**18. Distribution Statement** This document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161

19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 110	**22. Price**

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

ACKNOWLEDGMENTS

The authors wish to acknowledge several people who contributed their expert talent and many long hours to this study. Without their hard work and outstanding support, this study would not have been a success.

FAA Office of Runway Safety and Operational Services: Will Swank

FAA Innovations Division and NASA Ames Research Center Liaison Office: Parimal Kopardekar, Loren (Rod) Ketchum, Barry Scott

Study Development, Execution, Analysis, Facility and Laboratory Support at the NASA Ames Research Center: Terry Rager, Linton Kypta, and Bob Shipley (NASA), Diane Carpenter, Jim Miller, Charley Ross, Bob Cornell, Jerry Jones, Ramesh Panda, Ghislain Saillant, Ian Maclure, Dave Brown, Gary Uyehara, Jeff Dewey, Tom Standifur, and Norm Gray (Northrop Grumman Information Technology), Jan Aubright, and Wendy Krikorian (QSS Recruitment Office)

Study Analysis Support at the FAA Technical Center: Doug Baart (ACB-330) and George Smith (Titan Corporation)

Airline Participants: Airline Pilots Association, Allied Pilots Association, United Airlines, Northwest Airlines, Cathay Pacific Airways, Atlas Air Incorporated

We would also like to acknowledge and thank Dr. Paul Krois and Dr. Kim Cardosi for their contributions to this effort.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

EXECUTIVE SUMMARY

The Federal Aviation Administration (FAA) Office of Runway Safety and Operational Services formed a simulation team to investigate the safety effects of standardizing the use of aircraft landing lights in the airport environment. This document describes the simulation, which was a proof-of-concept study, termed Aircraft Landing Lights Enhance Runway Traffic Safety (AL^2ERTS). The purpose of this study was to gather subjective and performance data from flight crews as they operated in scenarios with and without standard exterior lighting procedures. Specifically, the simulation team explored the procedural use of landing lights as a direct message to other pilots indicating that aircraft were cleared to depart. The necessary data included a measure of Runway Incursions (RI), accidents, and pilot situation awareness (SA).

The simulation team, comprised of researchers from the FAA William J. Hughes Technical Center (WJHTC) and the National Aeronautics and Space Administration (NASA) Ames Research Center (ARC), conducted a real-time, human-in-the-loop simulation in October 2003-January 2004. The simulation utilized NASA ARC's Crew Vehicle Systems Research Facility (CVSRF) level D certified, Boeing 747-400 simulator.

Sixteen crews composed of a Captain and First Officer participated in this study in which they were instructed to taxi, depart, or land in 16 scenarios. Half of the crews operated in a baseline condition that had no standard procedures for using landing lights to indicate that aircraft are cleared to depart (no standard condition); the other half operated in an environment with standard procedures (standard condition). Both conditions included four scenarios in which a scripted confederate aircraft committed an error or followed erroneous instructions that resulted in a RI with the potential to result in an accident if not detected by the subject crew. We compared crews in the no standard condition and the standard condition in terms of their response to these scripted RIs.

In general, the pattern of results supports the standardized use of aircraft landing lights to indicate that aircraft are cleared to depart. The data showed that crews taxiing in an environment with a standard use of landing lights held-short more frequently (thereby preventing more incursions) than those with no standard. Crews with no standard crossed the runway with greater frequency and were involved in more collisions. Crews generally experienced incursions that were less severe when operating with a standard use of landing lights. Overall, landing lights provided a faster cue that there was a potential for a collision than movement alone, and crews in the standard condition reported that their first cue of an impending incursion were the landing lights.

Standardization of the use of landing lights also showed some benefits for SA. The 3D Situation Awareness Rating Technique (SART) rating trend of responses showed a slight increase in SA for Captains. Initial response times to departing aircraft were significantly faster for crews taxiing in the standard condition. Given accurate knowledge of events in the environment (namely, an aircraft departing), a faster response means greater safety. Finally, all of the pilots in the standard condition indicated that the standard use of landing lights increased safety.

In this simulation, we demonstrated the benefits of the procedure in an ideal environment where the complexity was relatively low and the lights were always visible. Further studies are suggested to determine the effects of other factors such as consistency of the message, message conspicuity, and effects of the message on other human system elements. Evaluating alternatives such as pulse lighting and the potential value of cues to Air Traffic Control may reveal additional benefits.

ACRONYMS

AC	Advisory Circular
AL^2ERTS	Aircraft Landing Lights Enhance Runway Traffic Safety
ARC	Ames Research Center
ATC	Air Traffic Control
CAST	Commercial Aviation Safety Team
CVSRF	Crew Vehicle Systems Research Facility
E/O	Expert Observer
EOS	Experiment Operator Station
FAA	Federal Aviation Administration
NASA	National Aeronautics and Space Administration
NWA	Northwest Airlines
OM	Operations Manager
ORD	Chicago O'Hare International Airport
PI	Principal Investigator
RI	Runway Incursion
RI JSIT	Runway Incursion Joint Safety Implementation Team
SA	Situation Awareness
SART	Situation Awareness Rating Technique
SME	Subject Matter Expert
SFO	San Francisco International Airport
SOP	Standard Operating Procedure
TIPH	Taxi Into Position and Hold
UAL	United Airlines
VCR	Videocassette Recorder
WJHTC	William J. Hughes Technical Center

1. Introduction

1.1 Background

The Runway Incursion Joint Safety Implementation Team (RI JSIT) was chartered by the Commercial Aviation Safety Team (CAST) and General Aviation Joint Steering Committee to develop a plan to effectively reduce the severe threat of fatalities and loss caused by commercial and general aviation runway incursion (RI) accidents/incidents. CAST's goal is to reduce the US commercial aviation fatal accident rate by 80% by the end of the year 2007. To help accomplish this goal, the RI JSIT brought together expert representatives from across the aviation community including participants from government, industry, and pilot and controller unions. These experts developed, prioritized, and coordinated a plan to implement the most effective analytically data-driven intervention strategies recommended by the Runway Incursion Joint Safety Analysis Team. RI JSIT analyzed those intervention strategies to determine the feasibility of gaining significant safety benefits through implementation. They incorporated twenty-two Safety Enhancements into seven Detailed Implementation Plans (FAA, 2002). One of these plans is to develop Standard Operating Procedures (SOPs) for ground operations; more specifically, SOPs relating to aircraft taxi operations and use of aircraft lighting during taxi operations.

Industry wide, SOPs are among the highest scoring safety enhancements across five accident categories including Controlled Flight into Terrain, Approach and Landing, Loss of Control, RIs, and Turbulence. The RI JSIT views the implementation of SOPs for aircraft taxi operations as one of the most powerful near-term interventions, as well as a low-cost option, in mitigating the occurrence and severity of RIs[1].

The Federal Aviation Administration (FAA) Office of Runway Safety and Operational Services formed a simulation team to investigate the safety effects of standardizing the use of aircraft landing lights in the airport environment. Researchers from the FAA William J. Hughes Technical Center (WJHTC) and the National Aeronautics and Space Administration (NASA) Ames Research Center (ARC) comprised the simulation team. This simulation team conducted a real-time, human-in-the-loop simulation in October 2003-January 2004.

This document describes the simulation, which was a proof-of-concept study, termed Aircraft Landing Lights Enhance Runway Traffic Safety (AL^{2}ERTS). In this study, we explored a baseline condition representing current use of procedures and a condition with the new SOP. In each condition, the four[2] data collection scenarios included a scripted error by a confederate pilot (a simulated aircraft whose crew was under our control) that induced a potential RI/accident. The simulation utilized the Crew Vehicle Systems Research Facility (CVSRF) at NASA ARC. In particular, the study employed NASA's level D certified, Boeing 747-400 simulator. The

[1] In the U.S., a runway incursion is defined as "any occurrence in the airport runway environment involving an aircraft, vehicle, person, or object on the ground that creates a collision hazard or results in a loss of required separation with an aircraft taking off, intending to take off, landing, or intending to land."

[2] In point of fact, in 8 of the 16 scenarios that every crew experienced, a confederate aircraft made an error that resulted in a RI and could have resulted in an accident with the B747-400 simulator if not detected and prevented by the subject crews. However, only four of those scenarios were included in the analysis because they represented precisely the situations of interest to the sponsor.

simulation team and sponsor selected San Francisco International Airport (SFO) and Chicago O'Hare International Airport (ORD) as the emulated airports.

1.2 Objectives

The purpose of this study was to gather subjective and performance data from flight crews as they operated in scenarios with and without standard exterior lighting procedures. Specifically, the simulation team explored the procedural use of landing lights as a direct message to other pilots indicating that aircraft were cleared to depart. The necessary data included a measure of RIs, accidents, and pilot situation awareness (SA).

The goal of this study was to investigate whether standardizing the use of aircraft landing lights to indicate that aircraft were cleared to depart (1) prevented or reduced the severity of RIs or accidents, and (2) increased pilot SA.

2. Empirical Evaluation

2.1 Organizational Roles and Responsibilities

2.1.1 Simulation Team

The simulation team consisted of the individuals who developed and conducted the simulation. The team consisted of an Operations Manager (OM), the Principal Investigator (PI), human factors researchers, Subject Matter Experts (SMEs), and laboratory personnel.

The OM and PI were responsible for the overall management of the simulation and were vested with the authority to direct the activities of all members of the team as they related to the simulation. The OM was a representative from NASA ARC's CVSRF. The PI was a representative from the FAA WJHTC's Simulation and Analysis Group.

Human factors researchers and SMEs had specified roles in the simulation, administered forms, and conducted participant briefings and debriefings. These members of the team (from FAA WJHTC and NASA ARC) were continuously available in the test areas to support the OM and PI.

Laboratory personnel (from NASA ARC) operated, monitored, and maintained the laboratory systems used in the simulation. They were continuously available in the test areas to support the OM and PI.

2.1.2 Sponsoring Organization

The FAA Office of Runway Safety and Operational Services was the sponsoring organization of the AL^{2}ERTS simulation. In addition, the Office of Runway Safety and Operational Services was directly involved with the effort by providing support in areas such as requirements guidance, simulation planning and scenario development, data reduction activities, analysis, as well as assisting in the participant recruitment effort. The Office of Runway Safety and Operational Services maintained overall authority for all aspects of the study.

2.2 Participants

Participation in this study was strictly voluntary, and the privacy of all participants will be protected. We did not record individual names or identities nor will we release them in any reports. We assigned each participant a participant code (e.g., T_1C, T_1F, T_2C, T_2F, etc.) that remained the same throughout the study.

The simulation team maintained strict adherence to all federal, union, and ethical guidelines throughout the study. The Human Research Institutional Review Board officially approved AL^2ERTS on October 2, 2003. Participants were subjected to minimal risk. Per the definition of minimal risk, the probabilities of harm or discomfort anticipated in this simulation were not greater than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.

Thirty-two pilots participated in the study as either the Captain or First Officer of the B747-400 simulator during the five weeks of simulation. The simulation team recruited pilots from sources such as the airlines, the Airline Pilots Association, and the Allied Pilots Association. Participants were current or retired/furloughed (9 months or less) B747-400 type-rated Captains or First Officers.

It was desirable that both pilots of each crew were from the same company; however, 4 of the 16 crews were from different or 'mixed' companies. Twelve of the crews were from United Airlines (UAL); therefore, we instructed the crews to use the B747-400 cockpit configuration and a checklist from UAL. The mixed crews used the Captain's configuration and checklist, when possible, and we provided additional training to the First Officer to mitigate any potential effects. If the Captain's configuration and/or checklist were not available, the pilots used UAL's and we trained one or both members of the crew (as appropriate) to mitigate any potential effects. The FAA WJHTC and NASA ARC personnel provided familiarity and procedural training as needed.

2.3 Airports

The simulation team and sponsor selected to emulate SFO and ORD for this simulation. Jeppesen® Airport diagrams can be found in Appendix A[3]. Selection criteria included runway configuration and recent (within five years) RI historical data.

2.4 Scenario Characteristics

A total of 32 scenarios representing two airports (i.e., SFO and ORD) were developed and utilized for the simulation. Each crew flew either the set of 16 standard scenarios or the set of 16 no standard scenarios. Scenarios varied as to whether the B747-400 simulator was an arrival, a departure, or a taxiing aircraft. In 8 of the 16 scenarios that every crew experienced, a confederate aircraft made an error that resulted in a RI and could have resulted in an accident with the B747-400 simulator if not detected and prevented by the subject crews (7 of these 8 scenarios were based on the events of actual RIs that occurred in the field). The other eight scenarios were similar; however, we did not induce an incursion.

[3] The Jeppesen® airport diagrams of ORD and SFO are for simulation purposes only. They are not intended for commercial use or navigation purposes.

The simulation team used 8 of the 32 scenarios to achieve the objectives of this study and the other 24 were used to control order effects. In the eight scenarios that we included in the analysis, an incursion was planned and the B747-400 crews were always crossing the runway as the confederate departed.

Some characteristics of the scenarios were varied to represent potential differences in the overall environment. Half of the planned incursions were at SFO and half were at ORD. In half of the incursions, the aircraft were situated at intersecting runways and/or taxiways that met at an angle of approximately 90°. The other half of the incursions began at obtuse angles (about 135°). We also balanced the side of the incursions; that is, whether they occurred on the left side of the B747-400 or the right. Appendix B summarizes the characteristics of the eight scenarios that we used for analysis. Appendix C describes the detailed events of these scenarios.

For the 24 scenarios used as controls, the confederate aircraft and the B747-400 crews performed taxi operations in other combinations (i.e., each could have been arriving, departing, or taxiing). One-third of these scenarios had planned incursions; however, these were not used in the analysis and we did not balance these incursions by airport, angle, or direction of occurrence.

All scenarios occurred in daylight conditions and in visual meteorological conditions. The weather conditions were clear and invariable, with unlimited visibility, and winds light and variable.

2.5 Facility

2.5.1 NASA B747-400 Simulator

The CVSRF is a unique national research facility dedicated to studies of aviation human factors and airspace operations and their impact upon aviation safety. An integral component of the CVSRF is the B747-400 simulator.

CAE Electronics built NASA's B747-400 simulator to meet the FAA Level D certification requirements (Sullivan & Soukup, 1996). The Boeing 747-400 has an advanced level of automation available to the pilots. The visual system uses photo texturing and offers superior scene quality depicting out the window scenes in night, day, dusk, or dawn conditions. In addition, the simulator has an advanced digital control loading and a six degree-of-freedom motion system. Data collection was available for user interaction with all subsystems, including the autopilot system and communication devices.

2.5.2 Pseudo-Pilot

One trained pseudo-pilot supported the operation by emulating all pilot communications (other than the B747-700 simulator). The pseudo-pilot used a voice disguiser to enhance realism. This individual was an SME with commercial pilot experience and was familiar with the airports and their operating procedures.

2.5.3 Experiment Operator Station

The Experiment Operator Station (EOS) operator was responsible for setting up the simulation environment (e.g., starting each run, selecting the proper airport, setting the weather conditions, aircraft weight, aircraft start position, etc.). This individual was an SME with commercial pilot experience and was familiar with the airports and their operating procedures.

2.5.4 Air Traffic Controller Support

One individual acted as the controller for the ground and tower operations. This individual performed air traffic control (ATC) ground-to-air and ground-to-ground communications as required. He used a voice disguiser, as necessary, to enhance realism. This individual was an SME with ATC experience and was familiar with the airports and their operating procedures.

2.5.5 Expert Observer

One Expert Observer (E/O) was stationed inside the cockpit of the B747-400 simulator. The E/O was an SME with commercial pilot experience and was familiar with the airports and operating procedures. The E/O recorded data on the E/O Over-the-Shoulder Form (see Appendix D) and the E/O Post-Run Form (see Appendix E).

2.6 Experimental Design

In this study, we instructed the crews to taxi, depart, or land in 16 scenarios. Sixteen crews composed of a Captain and First Officer participated in this study. Eight of the crews operated in an environment that had no standard procedures for using landing lights to indicate that aircraft are cleared to depart; the other eight operated in an environment with standard procedures. In half of the scenarios that every crew experienced, a confederate aircraft made an error that resulted in an incursion with the potential to result in an accident if not detected by the subject crews. Four of these scenarios for each crew were included in the final analysis.

This study design included the *Standardization* factor and it has two levels. The first level, *no standard*, represented a taxi environment in which the aircraft used the present policies and culture of selected airlines for aircraft lighting during taxi operations. The second level, *standard*, represented a taxi, position and hold, and takeoff environment in which the recommended standardization of procedures for the use of aircraft lighting were in effect per the Advisory Circular (AC) 120-74A (see applicable excerpt in Appendix F)[4]. All aircraft were scripted to comply correctly with the SOPs contained in AC 120-74A during the standard condition without exception.

The study included the Standardization factor in a design of independent groups. We chose this design instead of a within-subjects design because it was considered unreasonable to expect flight crews to switch from using one set of rules during taxi to the current practice without being influenced by the preceding conditions they received. The levels of Standardization (no standard and standard) constitute the independent (between-subjects) groups. We divided the study crews into two groups: half of the crews were placed in the no standard condition and the other half were placed in the standard condition.

The simulation team used counterbalancing to neutralize the effects of order as well as to provide a washout period after incursions. Practice effects (e.g., carryover, sensitization, and practice effects) were not the objects of study and the presence of practice effects could reduce the

[4] While all elements of the SOP pertaining to exterior lighting were adhered to, the objective of interest for this study was to evaluate the benefit derived from the particular section referring to the use of landing lights.

sensitivity of the design. Therefore, we used a Latin square arrangement of the airports (namely, SFO & ORD) and a random arrangement of the scenarios.

Table 1 shows how the factors were counterbalanced. Half of the crews were in the no standard condition and other half were in the standard condition. Confederates in scenarios 1 to 8 and 17 to 24 committed no incursions, but they did commit incursions in scenarios 9 to 16 and 25 to 32 (shown as underlined). Run order is also apparent from this table. For example, the first crew (c_1) encountered Scenario 16 first, Scenario 7 second, Scenario 14 third, etc. We only used the shaded scenarios in the analysis. We presented all others to control for order effects and to provide variability in the pilots' tasks.

Table 1. Counterbalancing of Scenarios for Each Group

		Run															
	Crew	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}	r_{16}
No Standard	c_1	16	7	14	10	13	1	12	11	15	6	8	3	2	4	9	5
	c_3	15	16	6	5	12	1	3	2	7	14	8	4	11	10	13	9
	c_5	6	15	8	12	9	11	2	13	7	14	16	5	3	1	10	4
	c_7	7	16	14	5	2	9	12	3	15	8	6	1	4	13	11	10
	c_9	10	4	13	3	1	15	16	6	2	12	11	9	5	14	7	8
	c_{11}	3	10	13	9	4	14	16	7	12	5	2	1	11	6	15	8
	c_{13}	5	9	11	10	12	14	15	16	2	13	4	3	1	8	6	7
	c_{15}	13	9	2	11	12	6	16	7	10	5	4	3	1	8	15	14
Standard	c_2	30	31	24	22	32	23	26	29	17	20	19	18	25	28	27	21
	c_4	31	32	22	24	30	23	18	27	29	28	17	19	25	20	21	26
	c_6	31	24	22	32	23	30	25	27	18	20	29	26	19	28	21	17
	c_8	23	30	32	22	31	24	27	25	18	21	28	26	17	29	20	19
	c_{10}	18	29	20	27	17	30	31	24	21	26	28	19	25	32	22	23
	c_{12}	17	20	29	27	18	31	24	22	21	26	19	25	28	32	23	30
	c_{14}	20	18	25	29	27	32	31	22	21	17	26	28	19	23	30	24
	c_{16}	29	25	18	20	27	22	32	23	28	19	26	17	21	30	24	31

2.7 Procedures

One crew, consisting of a Captain and First Officer, participated at the NASA Ames facility each day of simulation. We briefed the crews on relevant experiment information and gave them a safety briefing in the simulator. The crews then ran two training runs to gain familiarity with the simulator. Training runs also served to instruct crews in the standard group on new lighting procedures. After the crews indicated that they were familiar with the cockpit and expected procedures, data collection runs were conducted. The crew ran a series of 16 scenarios at two different airports.

2.7.1 Crew Briefing and Training

During the initial briefing, we provided crews with minimal information about the study objectives. In particular, they were not made aware of planned incursions, nor were they given feedback regarding whether an incursion occurred unless they elicited feedback about an event as they would during actual operations. We instructed crews to taxi, depart, and land aircraft using the same consideration for efficiency and vigilance for safety that they would in a real-world environment. We then trained crews to use the aircraft lighting procedures corresponding to their condition assignment (standard or no standard). We also reviewed the procedures during the simulator orientation. We instructed eight of the crews (c_1, c_3, c_5, etc.), those in the no standard condition, to follow the current aircraft policies (including lighting policies) of the Captain's company (See Appendix G). All other aircraft (i.e., other than the Boeing 747-400) in the no standard conditions also complied with their appropriate company policies. We instructed the other eight crews (c_2, c_4, c_6, etc.), those in the standard condition, to follow the SOP for light usage as written in AC 120-74A (also reference Appendix F). After completion of two training scenarios, we assessed the s' understanding of the lighting SOP with a written test (see Appendix H). All other aircraft in the standard conditions followed the same SOP.

2.7.2 Data Collection

After the initial briefing, the participants signed an Informed Consent Form (see Appendix I) and completed a Background Questionnaire (see Appendix J). The participants experienced two training runs and then completed the 16 test scenarios. During the scenarios, SMEs collected event information. The participants completed a Situation Awareness Rating Technique (SART) Questionnaire (Taylor, 1990; see Appendix K) at the end of each scenario. The SMEs completed an Observer Rating Form (Appendix E) after each test scenario. After completion of the last scenario, the participants completed a Post-Simulation Questionnaire (see Appendix L) and participated in a final debriefing. The participants worked from 8:00 AM to 4:00 PM each day with three breaks after a series of scenarios and a lunch break (see Appendix M for schedule).

2.8 Dependent Measures

2.8.1 Subjective Measures

The simulation team collected subjective data from participants using questionnaires, SA ratings, and debriefing sessions (see Table 2). In addition, an E/O collected observer data using an over-the-shoulder rating form, and the Office of Runway Safety and Operational Services performed subjective severity evaluations of RIs.

Table 2. Subjective Data Collection Methods

Instrument	Source	Frequency	Objective
Background Questionnaire	Participants	Once	Gather participant demographic information
Situation Awareness Ratings Technique	Participants	Every run	Record participant situation awareness ratings
E/O Over-the-Shoulder Form	E/O	Every run	To record real-time event data
E/O Post-Run Form	E/O and Participants	Every run	To question participants about unusual/unforeseen events during the run (e.g. incursions)
Post-Simulation Questionnaire	Participants	Once	Gather information regarding impact of the use of landing lights as a means of communicating aircraft intent (if applicable), training adequacy, workload, user acceptance, simulation fidelity, etc.
Debriefing	Participants	Once	Gather information that was not previously acquired

2.8.1.1 Questionnaires

During the initial briefing session, participants completed a Background Questionnaire (see Appendix J). The Background Questionnaire's purpose was to obtain demographic information about our sample and solicit information related to pilot experience.

At the end of each run, the E/O asked participants specific questions per the E/O Post-Run Form. The E/O Post-Run Form solicited participant responses regarding information about the occurrence of unforeseen events (namely, to draw out information about incursions and which crewmember detected the incurring aircraft).

At the end of all runs, participants completed a Post-Simulation Questionnaire (see Appendix L). This questionnaire solicited participant responses regarding information such as simulation fidelity, adequacy of training for the simulation, user acceptance of the procedures, and the impact of the use of the standardized procedures in the simulation.

All questionnaires contained space to provide additional comments or information as appropriate.

2.8.1.2 Situation Awareness

One of the study's objectives was to determine whether crews had greater SA in scenarios where there were standard aircraft lighting procedures than those where there were not. Gawron's (2000) SA measure selection process indicated that the SART (Taylor, 1990) is most appropriate for studies in which an SA measure needs to be collected unobtrusively. The SA measure can be taken after the scenario's completion without much effect on the participant's rating because of the time delay between the event and the measure. We used the rating scale to gather subjective data regarding crewmembers' SA on three dimensions: demand of their attention, the amount of attention available to deal with the situation, and their understanding of the situation (see Appendix K). The E/O administered the measure to each crewmember after each run. We used

SART ratings data in combination with objective data to determine differences in SA (see Table 3).

Table 3. Data Collection Requirement for Situation Awareness Analysis (Objective 2)

Measure	Description	Data Collection
SART rating	Crewmember 1 to 7 rating on a multi-dimensional situation awareness scale	• Administration of SART • After every run • Each crewmember
Initial Response	Time difference between the incurring target's turning-on of landing lights (or moving) and the crews first control manipulation or verbal response (e.g. query to ATC, discussion among the crew)	• Time target begins moving. The target moved and turned on landing lights at the same time in the standard condition. • E/O Event marker will flag response • Measured to the nearest millisecond

2.8.1.3 Expert Observer Over-the-Shoulder Forms

The E/O collected data about events that occurred during each scenario such as information about incursion detection, pilot motivation, equipment errors, etc. The E/O was also tasked with pressing an event button as soon as the crew indicated (via spoken word or action) that they had detected an incursion. Since the data sets were very large, pressing the event button inserted a flag in the data set that cued the researchers where to begin looking for initial response time data.

2.8.1.4 Severity of Incursions

In addition to collecting simulator data, the scenarios each crew navigated were video recorded to assist in the analysis of the occurrence and severity of RIs. The FAA systematically categorizes each RI in terms of severity to determine the margin of safety associated with each event. To determine severity in this simulation, human factors researchers provided the Office of Runway Safety and Operational Services with verbal descriptions of the incursions. The verbal descriptions, simulator output data, and video recordings were available for review to determine the margin of safety associated with each scenario. Factors considered in the severity categorization included speed and performance of the aircraft, aircraft type, the extent of evasive action taken, visibility conditions, and distance between parties (horizontal and/or vertical). The Office of Runway Safety and Operational Services used the information given in the narrative for each scenario to determine a severity rating. Figure 1 describes the severity ratings by category.

It is important to note that in real world operations the severity of the scripted incursions may have been much less severe or may not have resulted in an incursion at all. In the simulation, the incursion scenarios were scripted to have the incurring aircraft continue its departure to determine the closest horizontal or vertical separation between aircraft for that particular scenario without any corrective action being taken by either the flight crew or air traffic controller. For example, the flight crew of a departing aircraft would in all probability see a crossing Boeing 747 aircraft and abort its takeoff, or the crossing could be aborted by a controller.

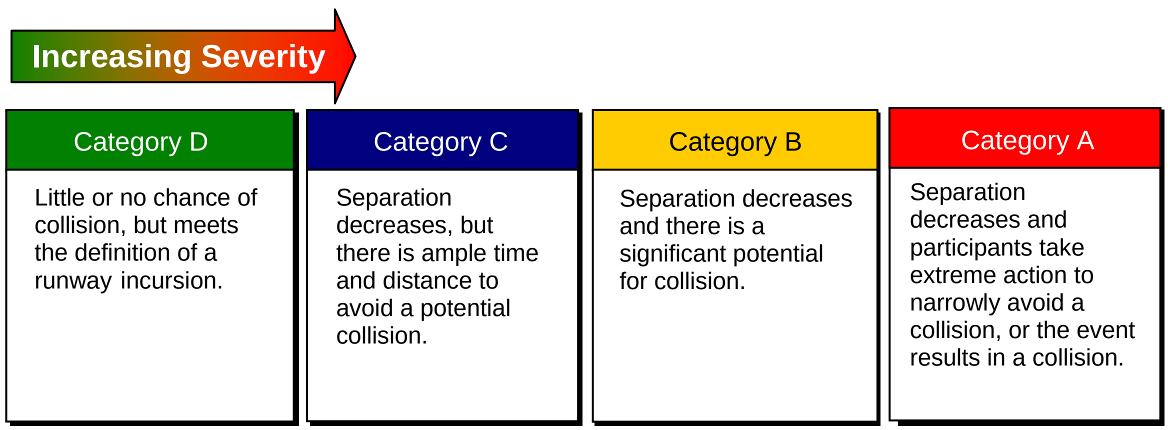


Figure 1. Runway Incursion Severity Categories.

2.8.2 Objective Measures

The objectives of the simulation were partially satisfied by measuring the incidence of Severity Categories in each standardization condition; however, other measures were obtained to support any conclusions drawn.

Table 4 lists the data collection instruments that we used and Table 5 lists the objective data that we collected, its description, and the source of the data.

Each run was video and audio recorded to capture the interaction between the crewmembers and between the crewmembers and ATC. The purpose was to gather supplemental data and to substantiate other subjective and objective data. Audio recordings captured all communications on the frequency and the ambient conversations between the pilots operating the B747-400 simulator. Two views were video recorded (a) an over-the-shoulder view of the pilots together in the cockpit, and (b) a view of the out-the-window screen. Since only one of the three out-the-window screens can be captured at a time, the EOS operator toggled to the appropriate screen. Aircraft lights were not visible on the video because the calligraphic lighting does not translate to the Videocassette Recorder (VCR) taping.

Table 4. Real-Time Data Collection Instruments

Instrument	Objective
Flight Data Recording	To collect data relating to flight operations such as navigation, accuracy, speed, and response time
Audio Recordings	For communication analyses, post simulation replay, and to provide backup source for data
Video Recordings	For post simulation replay and to provide backup source for data

E/O Over-the-Shoulder Form	Gather information regarding simulation events

Table 5. Objective Data, Its Description and Source

Data	Description	Source of Data
Closest Proximity	Distance (metal-to-metal) between the B747-400 simulator and confederate aircraft during planned incursion scenarios as measured by horizontal and/or vertical separation at defined points in time. The points in time were when NASA747 stopped before entering the runway or completely exited the runway.	Flight data recording
Runway Incursions	Scenarios were evaluated post-simulation to determine whether an incursion occurred and, if so, into what severity category it belongs	E/O, flight data, audio/video recording
False Alarm Rate	Proportion of no incursion trials on which the crews incorrectly detected an incursion	E/O, flight data recording
Probability of Correct Detection	Proportion of incursion trials on which the crews correctly detected an incursion	E/O
Initial Response Time	Measured from lights on and movement (this occurred at the same time in standard scenarios) (or from movement when no standard) to participant initial response (control action or verbalization) cued by E/O event marker	Audio/video recording, flight data recording
Altitude/Speed/Heading	Ground speed, time, position data for all aircraft including B747-400 simulator	Flight data recording
Aircraft Light States	State of all aircraft exterior lights including landing lights (on/off) for all aircraft including B747-400 simulator	Flight data recording

2.9 Treatment of Data

2.9.1 Descriptive Statistics

The simulation team selected the most appropriate descriptive statistical procedure for the data. We described nominal data (e.g., yes/no responses, categories, etc.) using the frequency counts or ratios and ordinal data (e.g., severity categories, ratings on questionnaires, etc.) using the median. We used the mean or median for interval data, depending on the shape of the distribution.

2.9.2 Statistical Treatment of the Data

We evaluated two research questions and corresponding hypotheses in this simulation to achieve the objectives. The first research question asked whether RIs or accidents occurred most in scenarios with no standard procedures or in those with standard procedures. The corresponding null hypothesis was that the number and severity of RIs and accidents that occurred were not different between the levels of the Standardization factor. The second research question asked whether pilot SA was higher in scenarios with no standard procedures or in those with standard procedures. The null hypothesis was that SA was not different between the levels of the Standardization factor.

We analyzed the principal measures using a Fisher exact probability test, Kolmogorov-Smirnov two-sample test, Robust Rank-Order Test, or a Wilcoxon-Mann-Whitney test. All statistically significant results reported in this document are significant at $p \leq .05$ unless stated otherwise. Because of our relatively low statistical power to detect significant effects, we also use marginal effects ($p \leq .10$) in some cases to help explain trends in the data. These marginal effects are important for the interpretation of the data even though they do not provide the same strength of support as the effects we arbitrarily call "significant." Whereas there is a 5% probability that a significant result was due to chance when using a criteria of $p \leq .05$, the probability of finding a significant result by chance increases to 10% when using a criteria of $p \leq .10$.

2.9.2.1 The Fisher Exact Probability Test

We used the Fisher exact probability test to determine whether there was a difference in terms of the number of prevented incursions between the Standardization conditions. The null hypothesis H_0 was that the crews did not hold short as a function of standardization group. The alternative hypothesis H_1 was that the crews using standard lighting procedures were able to prevent more incursions.

We also used the test to determine whether there were differences in the first reported cue between the Standardization conditions. The null hypothesis H_0 was that the crews did not differ in the first reported cue as a function of standardization group. The alternative hypothesis H_1 was that the crews using standard lighting procedures reported that their first cue was the lights more often.

2.9.2.2 Kolmogorov-Smirnov Two-Sample Test

We used the Kolmogorov-Smirnov test to determine whether there was a difference in terms of severity between the Standardization conditions in each scenario. We made the comparison between the levels of Standardization, no standard and standard. The null hypothesis H_0 was that there was no difference in severity experienced between the groups. The alternative hypothesis H_1 was that crews with standard procedures experience less severity than those with no standard procedures. We used the median of the severity categories for each crew for the analysis.

2.9.2.3 Robust Rank Order Test

We used the *robust rank-order test* to determine if the severity of incursions in the no standard group was higher than those in the standard group. The null hypothesis H_0 was that the sum of the severity ratings across the four scenarios was the same for the standard and no standard groups. The alternative hypothesis H_1 was that the sum of the standard group's ratings was higher[5] than those of the no standard group.

---

[5] Higher ratings are less severe. For example, an A rating was labeled with a 1, a B rating was a 2, a C was 3, and so on.

2.9.2.4 Wilcoxon-Mann-Whitney Test

We used the Wilcoxon-Mann-Whitney test to determine if SA was improved by the new lighting SOPs. The null hypothesis was that the standard and no standard conditions have the same distribution. The alternative hypothesis H_1 was that the median of the standard condition is greater than (H_1:$\theta_{\text{standard}} > \theta_{\text{no standard}}$) or less than ($H_1$:$\theta_{\text{standard}} < \theta_{\text{no standard}}$) the median of the condition that has no standard.

3. Simulation Assumptions and Limitations

It was beyond the scope of this study to determine changes in the conspicuity of aircraft resulting from different lighting configurations. The simulator cannot produce the photometric and colorimetric quantities of the actual lighting system with the fidelity required to determine the extent to which the human visual system can detect and identify the stimulus. All other aircraft presented to the crew of the B747-400 simulator correctly followed the prescribed lighting configuration of their group (refer to Appendix G), and their lights were visible. Further studies are necessary to address these and any other illumination issues.

There were no general aviation aircraft or ground vehicles in the scenarios.

Pilot reactions to incursions are presumed to have been better in the simulation than they would have been in the real world. The researchers assume that the crews were vigilant during the runs, which may have improved performance and SA. If no control measures were taken, the level of vigilance could have been so high that the advantage of the landing light message could have been suppressed or exaggerated. The between-groups design, runs conducted without incursions, scenarios added for the washout periods, and the counterbalances were expected to lessen the effect of vigilance on performance and SA. Even with the control measures in place, the level of vigilance in this setting may have limited our ability to use the SART such that finding statistically significant differences would require an impractical increase in sample size.

Some unforeseen technical problems and unusual simulator behavior did occur in the simulation. These occurrences were rare and largely minor issues. However, we repeated data collection runs when necessary to ensure data integrity. Therefore, it was determined that there was no observable impact on results.

4. Results

Sixteen flight crews participated in the simulation for one day each. Each crew flew 16 runs (see Appendix M for the Daily Pilot Schedule). Of the 16 runs, we used 4 for data analyses; the other 12 runs were implemented to provide variability in the pilots' tasks and for counterbalancing to neutralize the effect of order.

As previously described, scenarios varied by the presence and absence of standard procedures for exterior lighting. The primary focus of this simulation was to provide objective data as evidence to support or disprove the value of implementing a standard use of exterior aircraft lighting, particularly the use of aircraft landing lights. In addition, an objective of this study was the impact of the use of the new lighting procedures on SA. For this objective, SART was the primary measure. Some additional relevant subjective results are also presented as supplementary information.

4.1 Planned Runway Incursions

In the four data collection scenarios, the confederate aircraft made a planned error that could have resulted in an accident with the B747-400 simulator if not detected by the B747-400 crews. Overall, of the 64 planned incursions in the data collection scenarios of the simulation, 34 resulted in the B747-400 crews holding short without crossing the hold-short line to avoid the incursion. Twenty-nine of the planned incursions resulted in actual RIs that were divided into three different descriptive categories: stopped short of the runway edge, crossed the runway and exited the runway, and collision. In the first, the B747-400 crews stopped past the hold short line, but did not enter the runway. In the second category, crews continued to cross but exited the runway in time to avoid a collision. In the third, the crews were in the runway and a collision occurred. One planned incursion never occurred because the confederate aircraft never departed. In the following paragraphs we will discuss how the incursions differed depending on the Standardization condition that crews were taxiing in.

Table 6 provides a summary of the results of the planned incursions. With the B747-400 crews cleared to cross the runway and the confederate aircraft taking off, the crews in the standard condition stopped before the hold short lines (thereby preventing incursions) 43% more often than the crews in the no standard condition. The probability of this outcome using the Fisher exact probability test was marginally significant at $p = .08$ (see Figure 2). Standard crews also crossed the runway 56% less often and were involved in 67% fewer collisions.

Table 6. Summary of Planned Incursions

Action of B747-400 Crews After Cleared to Cross	Standard Condition (# crews)	No Standard (# crews)
Held Short (at Hold Short Lines)	**20**	**14**
Stopped Short of Runway Edge	**6**	**6**
Crossed Runway and Exited Runway	**4**	**9**
Collision	**1**	**3**

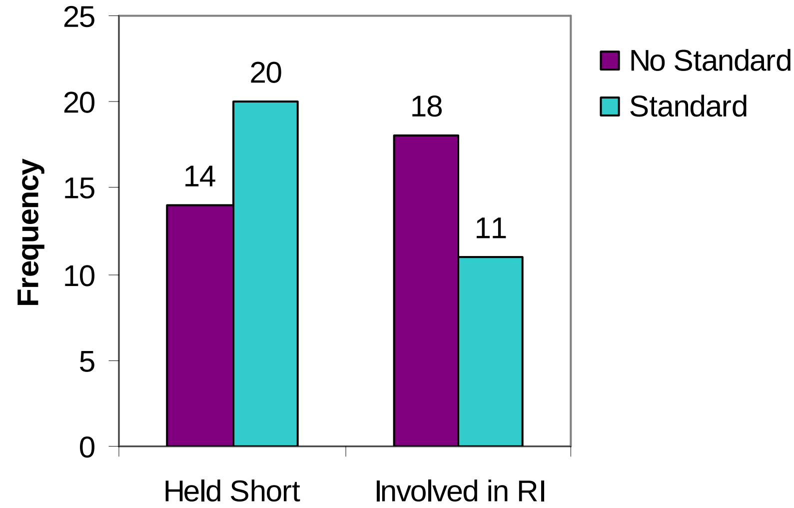


Figure 2. Overall number of runway incursions prevented by standardization condition.

An independent reviewer suggested that we conduct a focused analysis of the scenarios in the no standard condition in which the confederate aircraft had the landing lights on for Taxi in Position and Hold (TIPH). Examination of the data (see Table 7) showed that even in the no standard condition, the onset [turning on] of the landing lights for takeoff seemed to have had a positive effect. For the crews in the no standard condition, 2 of the 3 (67%) collisions and 7 of the 9 (78%) runway crossings occurred when the departing confederate aircraft had the landing light on for TIPH and remained on when takeoff clearance was received. Meanwhile, the majority of cases in which these crews stopped before entering the runway occurred when the landing light onset was coincident with takeoff, which accounts for 8 of the 14 (57%) times the crews held short, and 5 of the 6 (83%) times the crew stopped before crossing the runway edge. These results suggest that reserving the onset of the landing light for takeoff roll helps to prevent crossing aircraft from taxiing onto the runway. However, the results are confounded. The initial distance between the aircraft is greater in the lights-on for TIPH scenarios, making it potentially more difficult to determine whether the aircraft is moving. The viewing angles are also systematically different. There are other differences between these scenarios that are controlled in our design, but emerge when analyzed this way. The results were interesting nonetheless and may warrant future research.

Table 7. Planned Incursions for No Standard Condition

Action of B747-400 Crews After Cleared to Cross	No Standard (# crews)	No Standard (# crews when confederate had landing lights on for TIPH)
Held Short (at Hold Short Lines)	**14**	**6**
Stopped Short of Runway Edge	**6**	**1**
Crossed Runway and Exited Runway	**9**	**7**
Collision	**3**	**2**

4.1.1 Crews Holding Short

In the introduction of section 4.1, we compared the no standard and standard condition. We found that of the 64 incursions that were planned in the data collection scenarios, 34 resulted in the B747-400 crews holding short (without crossing the hold-short line). The scenarios are different in some respects (see Appendixes B and C); therefore, we will mention the general frequencies and show the results divided by scenario in subsequent analyses.

We counted the number of crews that held short for each scenario. The frequencies are displayed in Figure 3. The figure shows that crews taxiing in the standard condition held short more frequently than crews with no standard for 3 of the 4 scenarios. When we determined the probability of the outcomes using the Fisher exact probability test, the respective probabilities (p) for ORD 5, ORD 6, SFO 1, and SFO 3 were .23, .07, .30, and .23. Thus, we could not reject the null hypothesis at an alpha level of .05.

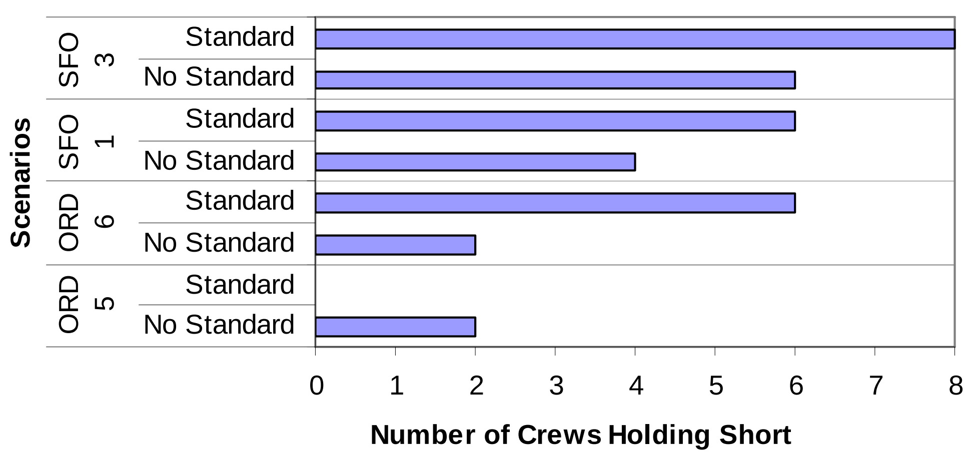


Figure 3. Number of crews holding short by standardization condition and scenario.

4.1.2 Runway Incursions

We counted the number of incursions that resulted in each type of scenario. Of the 64 incursions that were planned in the data collection scenarios, 29 resulted in actual incursions and the rest did not meet the criteria for a RI. The incursions were divided into three different descriptive categories: stopped short of runway edge, crossed the runway and exited the runway, and collision. Figure 4 shows the number of incursions in each category. Overall there were 11 RIs in the standard condition and 18 (or 63% more) in the no standard condition. The figure shows that in 3 of 4 scenarios, the no standard crews were involved in more incursions. Furthermore, the no standard crews were also involved in 3 collisions compared to 1 collision in the standard group. The probabilities of these outcomes did not reach statistical significance.

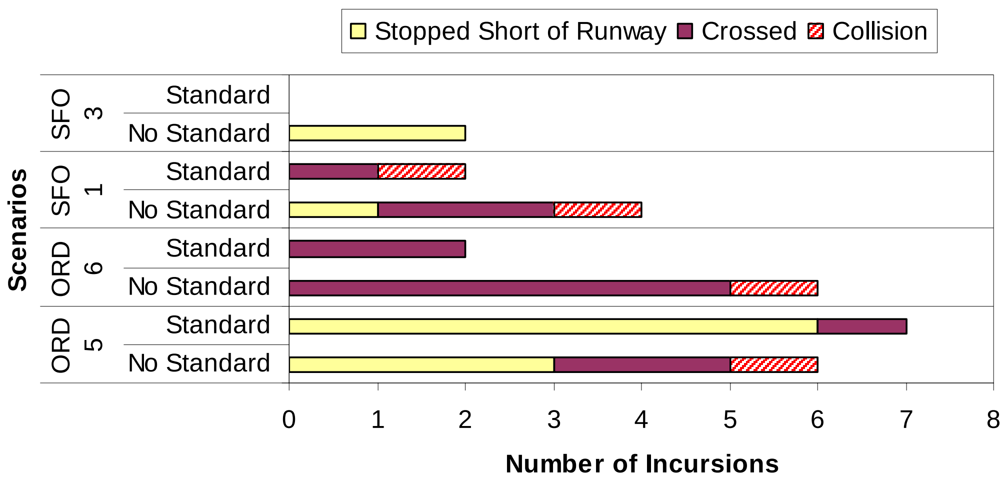


Figure 4. Number of crews involved in incursions by type of incursion, standardization condition, and scenario.

4.1.3 Severity Ratings

Each incursion was assigned a severity rating by the FAA Office of Runway Safety and Operational Services (see section 2.8.1.4). No differences in severity ratings were found between no standard and standard groups when individual scenarios (e.g., ORD 5, ORD 6, etc.) were analyzed using the Kolmogorov-Smirnov two-sample test (see Figure 5 for the data).

To further investigate, the sum of ratings assigned across the four incursion scenarios was used for the analysis. We analyzed the severity ratings using the *robust rank-order test.* We used the test to determine whether the severity of the incursions experienced by the no standard group exceeded those experienced by the standard group. We assigned each rating a number. The lowest number, one, was assigned to all *A* ratings, two was assigned to all *B* ratings, and so on. The sum of these numbers across all four scenarios for each crew was then used for the analysis. The observed value of the statistic, $\grave{U} = 1.31$, did not exceed the critical value at $\alpha = .05$. Therefore, we could not reject the hypothesis that there was no difference between the severity of incursions experienced by the two groups.

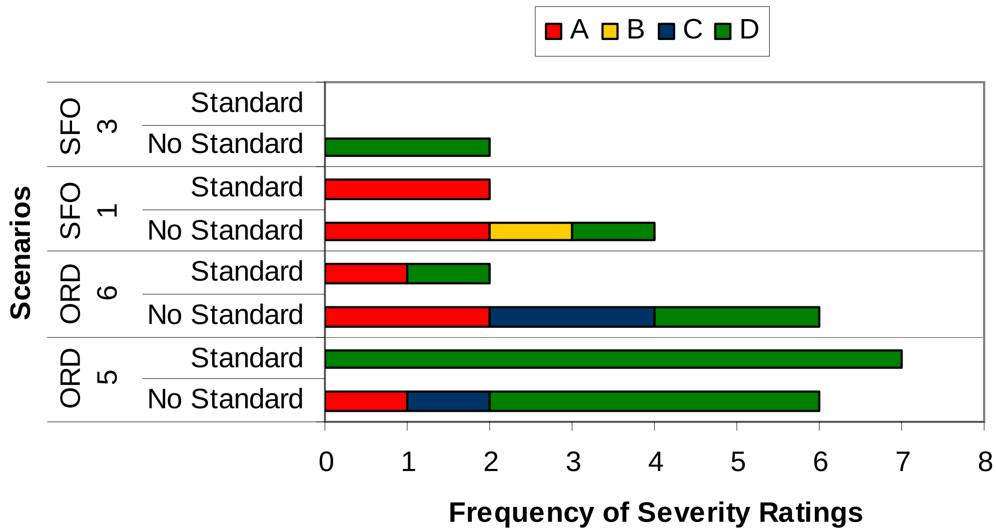


Figure 5. Number of crews involved in incursions by severity rating, standardization condition, and scenario.

4.2 Initial Response Times

Data for false alarm and correct detection of incursions for each crew were acquired via the Observer Post-Run form. Results from an analysis of this data indicated that crews were able to recognize conflicts with very high accuracy (*P*[hit] = .98 for standard and no standard) and no false alarms. Therefore, we were confident that we could examine the results from this section concerning initial response time data without being concerned about speed-accuracy trade-offs or response biases.

We conducted a task analysis to obtain a list of initial responses to the incurring aircraft. Possible responses included verbal reaction by either crewmember, an increase or decrease in power, pressing the brake[6], or a push-to-talk communication. We obtained verbalizations times from video recordings; all other data came from simulator data recordings.

Figure 6 represents average crew initial response times, in seconds, to an incursion event measured from the start of an incursion to the initial crew response. The start time was the actual movement of the incurring aircraft. Note that in the standard condition, the lights came on and the aircraft started to move simultaneously. We used the first and fastest of the crew reaction(s) to obtain the initial response time. The soundness of the data and adherence to statistical assumptions was then determined. We omitted crew five because the tapes with their verbal response times were damaged.

The data set was characterized by unequal samples and variances. Therefore, we chose nonparametric statistics to analyze the data. We used the Kolmogorov-Smirnov two-sample test to determine whether the two samples were drawn from the same population or populations with

[6] A program error resulted in the failure to collect braking times for 11 out of 16 crews. However, for the braking data that were collected, braking was never the fastest response.

the same distribution. The H_0 was that there was no difference in the reaction times between the standard and no standard conditions. The H_1 was that there was a difference in the reaction times between the environments. Results were in favor of rejecting H_0 ($Dm,n = .86$, $p < .05$). There was a statistically significant difference in reaction times between the two conditions indicating that the standard condition (i.e., SOP) reduced crew initial response times to RIs.

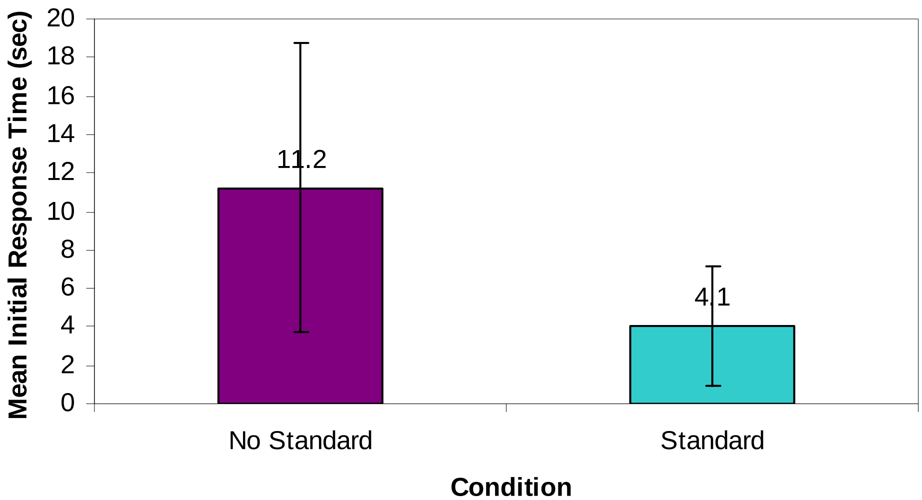


Figure 6. Mean (+/- standard deviation) Initial Response Times in seconds.

4.3 First Cue Data

We used a combination of initial response data, video recording analysis, and self-reported first cue detection (E/O Post Runs forms) to track each crewmember's first cue of an impending incursion (for scenarios with planned incursions), and to determine which crewmember detected the incursion first. Following each run, the E/O queried all crews and recorded which pilot noticed the potential incurring aircraft first, as well as the first cue for each of the pilots. (Note: as with all self-reported data, it is important to keep in mind that the pilot self-reports from the E/O logs inherently contain the potential for error.) We categorized all first cue data as *lights*, *movement*, or *other*. "Lights" refers to the landing lights of the departing aircraft that were turned on when the takeoff clearance was received and/or as the aircraft began takeoff roll. "Movement" refers to the noticeable movement of the conflicting aircraft. When "other" was indicated, it typically referred to pilots' reporting when the other crewmember called out that they noticed the impending incursion (i.e., their cue was the other pilot noticing first).

Each condition (standard and no standard) had a total of 32 runs with planned incursions; therefore, the expected number of responses for all pilot categories presented (Captain, First Officer, Crew) was n = 32. However, for each category some data was not available or it was inconsistent (i.e., pilot perceptions were not consistent with the objective data). Post simulation inspection indicated that the missing Captain/First Officer data were either due to an unobserved incursion or due to a failure to respond to the E/O's question. Upon inspection of the combined crew data, we found that several runs were not consistent across all data points. That is, pilot perceptions were not consistent with the objective data. We used video recordings and simulator data to validate which pilot recognized the potential incursions first. When data was inconsistent and could not be reconciled from the videotapes, the data was treated as missing.

4.3.1 First Cue Data - Captain vs. First Officer

These results indicate the first cue data of both the Captains and First Officers regardless of who detected the incursion first. For the standard condition, 4 responses from the Captains and 2 responses from the First Officers were missing, resulting in n=28 and n=30 respectively. Figure 7 depicts the frequency of the Captains' and First Officers' reported first cues of impending incursions in the standard condition. It is important to note that these crews received potential light cues (as prescribed in the SOP) in 100% of the runs. In general, lights were reported as the most commonly detected cue for both pilots, followed by other. Pilots reported movement as the first noticed cue the least. Between both crewmembers, first officers detected lights most often. Captains reported that their first cue for detecting incursions was lights 61% of the time, and other 32% of the time. Movement was the least detected first cue reported; it was reported only 7% of the time. First Officers detected lights cues in 83% of the scenarios, other cues 10% of the time, and movement was again the least reported cue at 7% of the time.

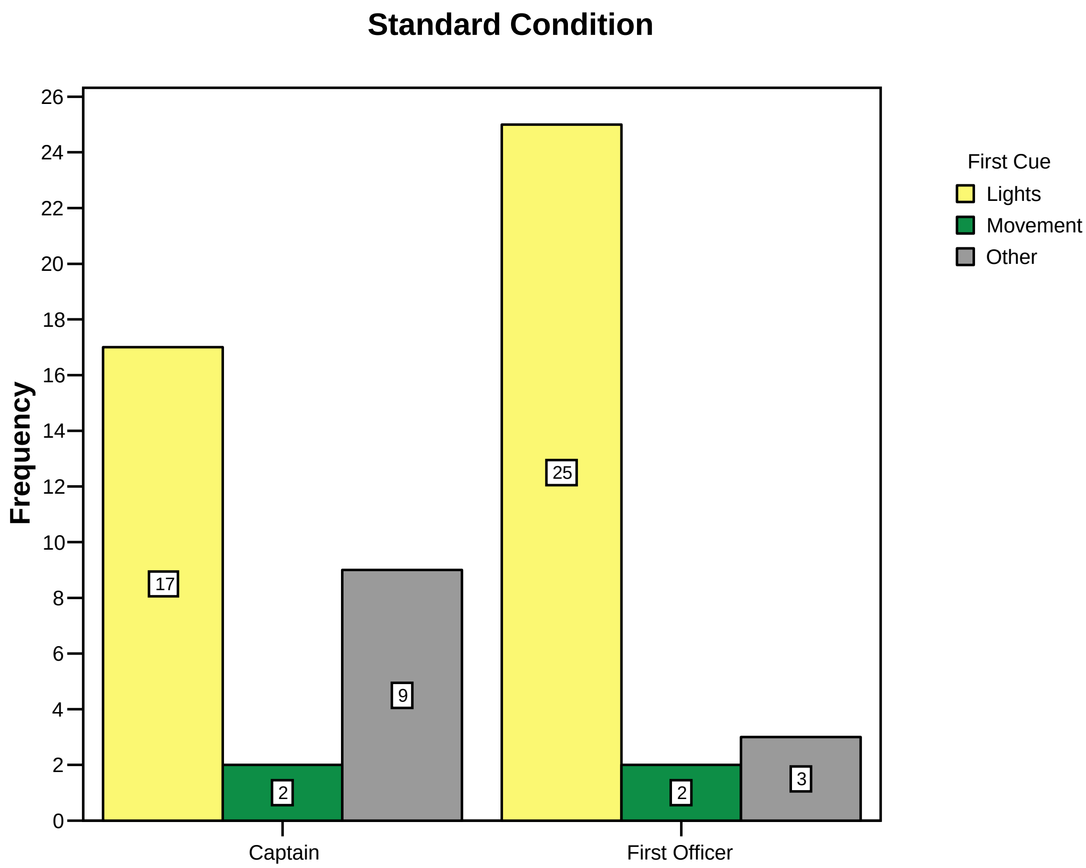


Figure 7. Standard Condition (overall): First cue reported by the Captain/First Officer that an incursion was occurring.

For the no standard condition, 1 response from the Captains and 2 responses from the First Officers were missing resulting in $n = 31$ and $n = 30$ respectively. Figure 8 depicts the frequency of the Captains' and First Officers' reported first cues of impending incursions in the no standard condition. It is important to note that these crews received potential light cues only 50% of the time in order to emulate real world conditions. In general, no individual cue stood out consistently. Captains reported first detecting movement and other cues 36% of the time each, while they reported lights as the first cue 29% of time (out of a possible 50%). First Officers detected movement cues first in 50% of the runs, lights cues 37% of the time (out of a possible 50%), and other only 13% of the time.

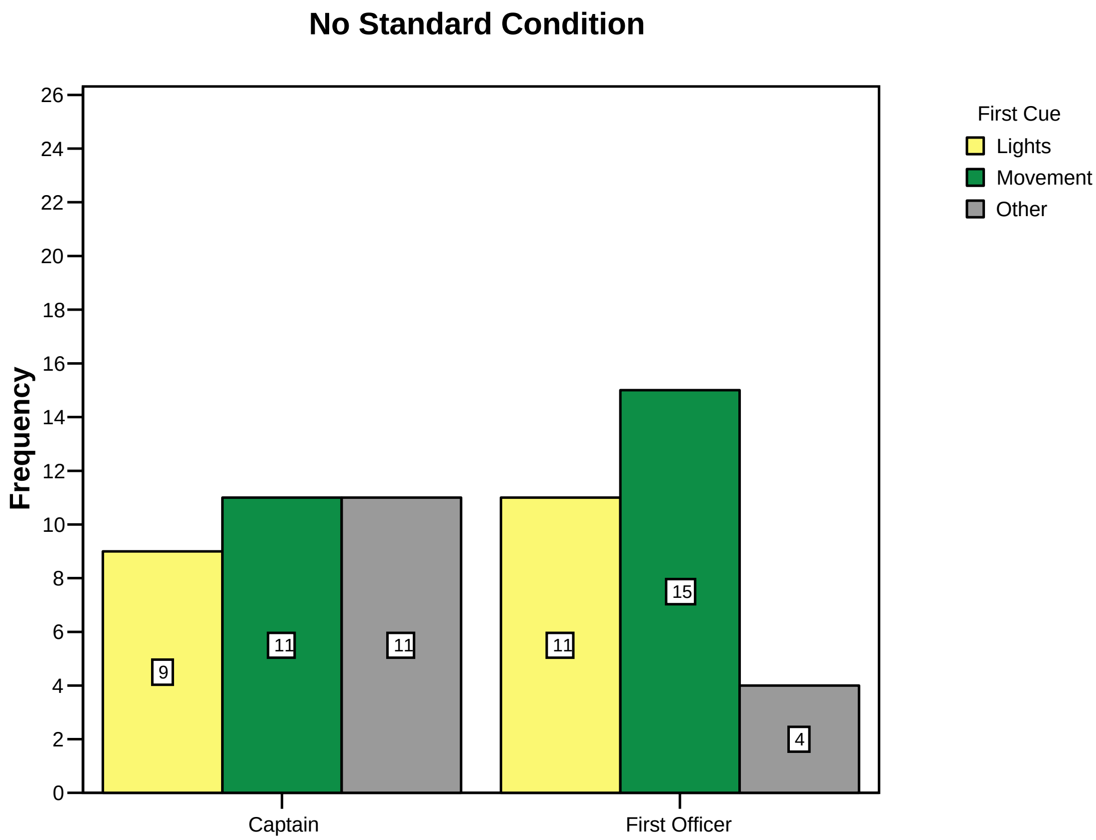


Figure 8. No Standard Condition (overall): First cue reported by the Captain/First Officer that an incursion was occurring.

We also compared crewmembers in each of the Standardization conditions to determine if they differenced in their reported first cue. We separated the data by scenario so that the frequency counts would represent a unique response. In our analysis, we counted the frequency of reports of "lights," movement, or other as a first cue. We collapsed the movement and other cue responses so that we would have two possible responses. However, the categories are separated in the figures below.

When we looked at the data by scenario and determined the probability of the outcomes using the Fisher exact probability test, the respective probabilities (p) for ORD 5, ORD 6, SFO 1, and SFO 3 were .66, .22, .08, and .41. Thus, we could not reject the null hypothesis at a $p < .05$. Regardless, the reader should note the pattern of results in Figure 9 indicating that Captains in the standard condition reported with greater frequency that their first cue of an impending incursion was "lights."

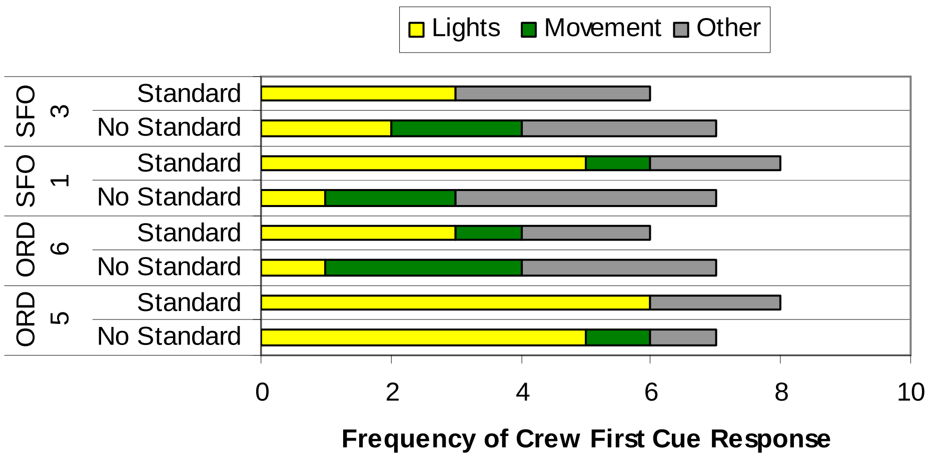


Figure 9. First cue reported by the Captain that an incursion was occurring shown by Standardization condition.

When we looked at the data by scenario and determined the probability of the outcomes using the Fisher exact probability test, the respective probabilities (*p*) for ORD 5, ORD 6, SFO 1, and SFO 3 were .50, .30, .03, and .01. Thus, results for two of the scenarios were statistically significant. The reader may note the pattern of results in Figure 10 where the First Officers in the standard condition reported more frequently that their first cue of an impending incursion was "lights."

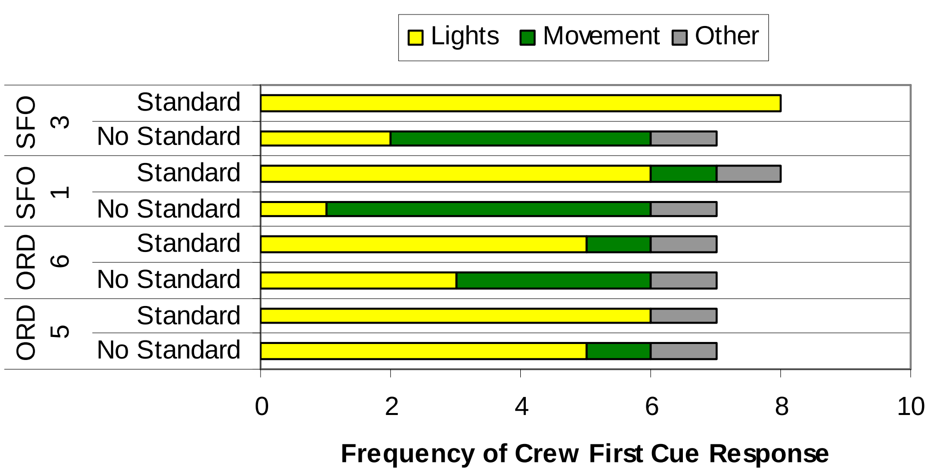


Figure 10. First cue reported by the First Officer that an incursion was occurring shown by Standardization condition.

4.3.2 First Cue Data - First to Detect

These results indicate the first cue data from the combined crew; therefore, data are reported only for the pilot of the crew who first detected the incursion. We computed the data using first cue self-report data (contained on the E/O Post Run form) in conjunction with objective data to accurately determine which pilot detected the incursion first. We omitted runs that contained data that contradicted videotaped recordings due to questionable reliability.

Figure 11 depicts the combined crews' first cue data for the standard condition. Of the 32 planned incursions, one data point was omitted because the crew never detected the incursion. Two other data points were omitted due to incomplete or inconsistent self-report data (n = 29). Again, it is important to note that these crews received potential light cues (as prescribed in the SOP) in 100% of the runs. Crews reported lights as their first cue 93% of time, while movement cued the crew only 7% of the time. As depicted in Figure 11, movement was the first cue in only 2 of the 29 runs analyzed in this data. First Officers in the standard group identified the incursion first 72% of the time, while Captains identified the incursion first the remaining 28% of the time. Furthermore, the crewmember stationed on the same side of the aircraft as the planned incursion was more likely to detect the occurrence of an incursion than the pilot on the opposite side. The opposite side crewmember detected the incursion 38% of the time, while the same side crewmember detected the potential incursion 62% of the time.

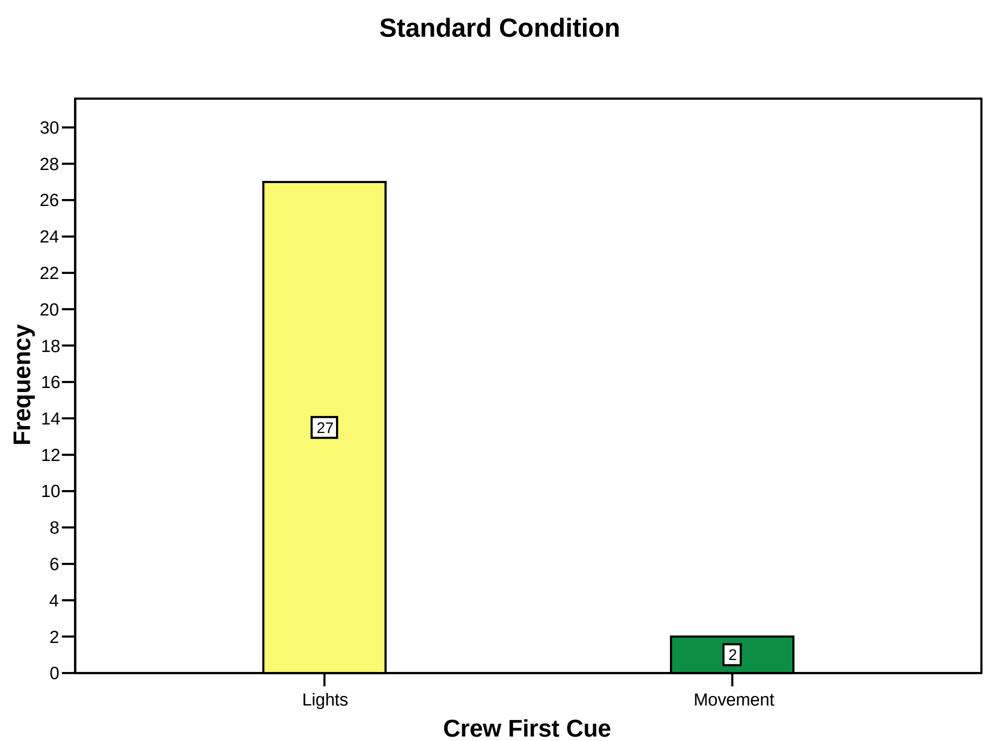


Figure 11. Standard Condition (overall): First cue reported by the crew that an incursion was occurring.

Figure 12 depicts the crews' first cue data for the no standard condition. Four runs from the 5th crew were not included in the summary because the video recording was damaged and no other reliable first cue data could be determined; three additional data points were omitted due to inconsistent self report data (n = 25). Again, it is important to note that these crews received potential light cues only 50% of the time in order to emulate real world conditions. For the no standard condition, movement was the primary cue for first detection of a potential incursion. The first pilot to detect incursions used movement as a primary cue in 60% of the runs; lights were detected first 40% of the time (out of a possible 50%). Contrary to the standard condition, Captains and First Officers were almost equally as likely to detect the impending incursion in the no standard condition. First Officers in the no standard group identified the incursion first 48% of the time, while Captains identified the incursion first 52% of the time. Same side crewmembers detected incursions slightly more often than opposite side crewmembers, 56% vs. 44%, respectively.

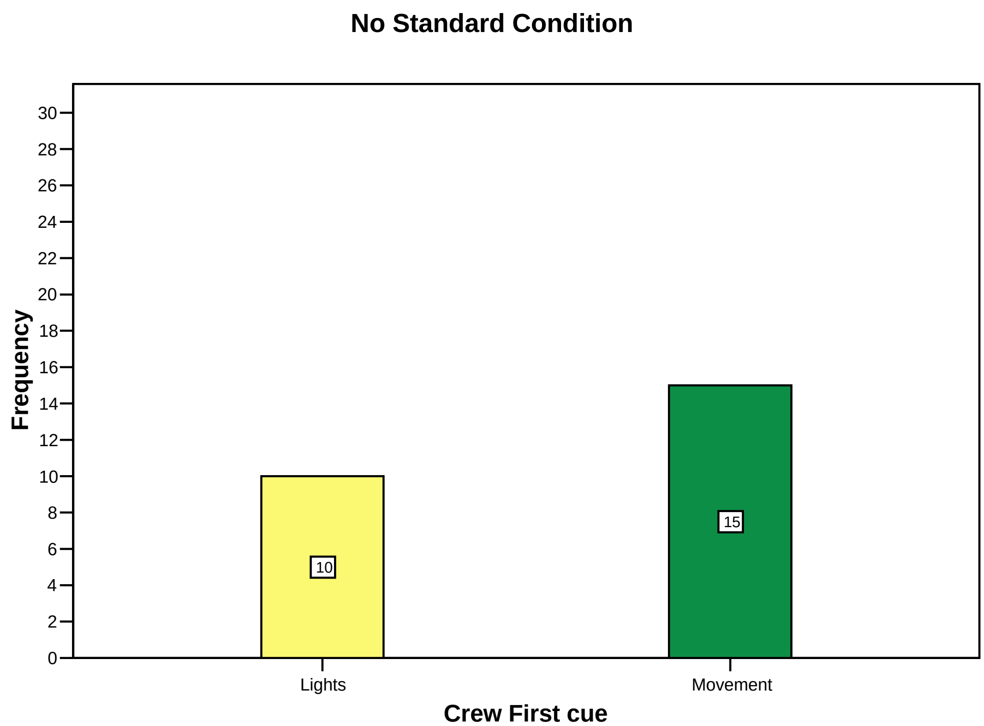


Figure 12. No Standard Condition (overall): First cue reported by the crew that an incursion was occurring.

When we looked at the data by scenario and determined the probability of the outcomes using the Fisher exact probability test, the respective probabilities (*p*) for ORD 5, ORD 6, SFO 1, and SFO 3 were .54, .02, .03, and .02. Thus, results for three of the scenarios were statistically significant. Crews in the standard condition reported more frequently that their first cue of an impending incursion was “lights” (see Figure 13).

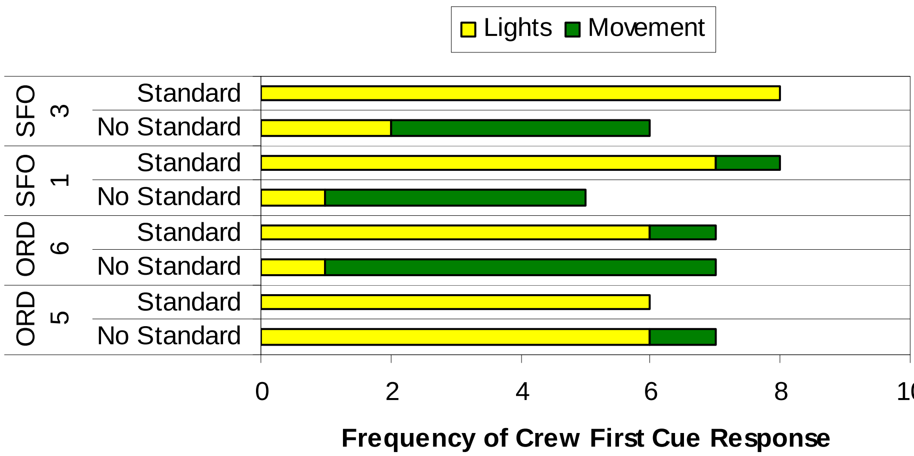


Figure 13. Standard Condition (by scenario): First cue reported by the crew that an incursion was occurring.

4.4 Situation Awareness

We used the three-dimensional version of the SART (Taylor, 1989) to collect the data used in this analysis. We asked crewmembers to select a rating from 1 to 7 (1 = Low, 7 = High) on the SART for each of the following items:

- Demand of attention: How demanding was the scenario on your attention?
- Supply of attention resources: How much attention did you have available to devote to the scenario?
- Understanding of the situation: How well did you understand the situation as it was in this scenario?
- Situation Awareness: Based on this scenario as it was simulated – how good was your ability to perceive elements in the environment (for example, your present situation, traffic movement, ATC interaction, etc.), to comprehend their meaning and project their status?

During interviews at the end of the day, several First Officers indicated that they were providing SA ratings for the task they were performing during the run. Because the First Officers were performing checklist related tasks, we analyzed the Captains' and the First Officers' SART ratings separately.

4.4.1 Captain Results

We used the median (*Md*) of the ratings provided across trials as the measure of central tendency in the analyses. The H_0 in the analyses was that the SART ratings given by Captains taxiing in the no standard conditions were the same as the ratings given by Captains taxiing in the standard conditions. The H_1 is that the SART ratings given by the Captains taxiing in the no standard conditions will be different from those given by the Captains taxiing in the standard conditions. The direction of the alternative hypothesis is given for each analysis.

Results are presented graphically in Figure 14 and Figure 15 and discussed by dimension in the following sections. We analyzed the data using the Wilcoxon-Mann-Whitney test (hereafter, Wilcoxon Test). We chose the Wilcoxon Test because this study uses two independent samples, which are small in n, and the measurement is on an ordinal scale. The significance level was set at $\alpha = .05$, $m = 8$ (no standard), $n = 8$ (standard).

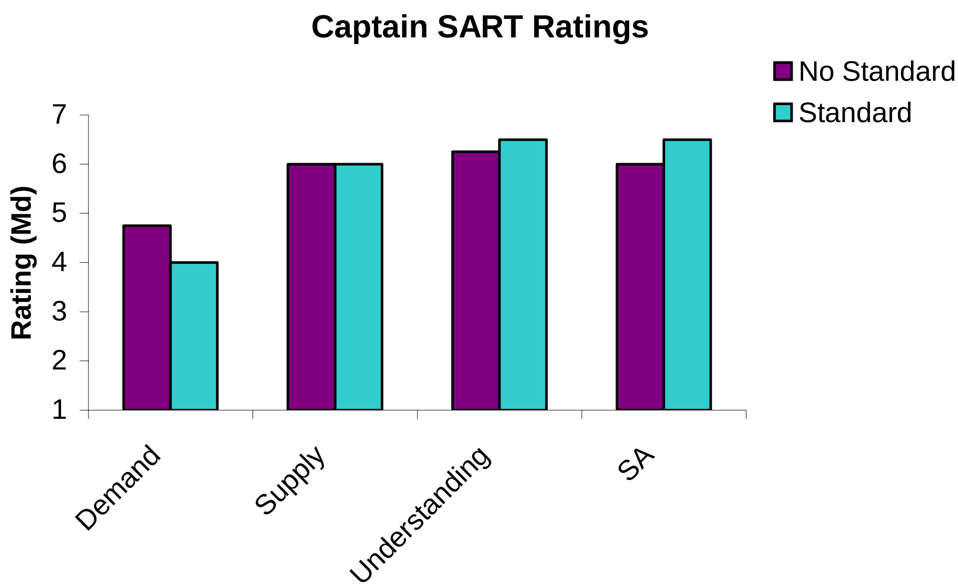


Figure 14. Captain SART ratings.

4.4.1.1 Demand of Attention

The hypothesis was that Captains operating in the standardized condition would have a reliable cue and, thus, there would be a lower demand for their attention. The comparison was in terms of what their ratings were for this dimension on the SART. A lower rating would indicate a lower demand for attention. The difference between the Demand of Attention ratings was not statistically significant between the groups, $P[W_{NS} \geq 70] = .4392$. Therefore, the data did not give evidence that justified rejecting H_0 at the .05 level of significance. These data do not support the hypothesis that the standardized condition provided a reliable cue that resulted in a lower demand of attention.

4.4.1.2 Supply of Attention Resources

The simulation team did not design this study to add any tasks that would take attention away from normal taxi operations. Had there been a difference between groups, the ratings for Supply of Attention resources may have indicated a bias in either group toward selecting higher or lower responses. The ratings appear to be equal in this dimension (as they were expected to be so).

4.4.1.3 Understanding of the Situation

The hypothesis was that Captains operating in the standardized condition would have more information about the intent or actions of the other aircraft, therefore, understand the situation more. The comparison was in terms of what their ratings were for this dimension on the SART.

Higher ratings would indicate a higher understanding of the situation. The difference between Understanding of the Situation ratings was not statistically significant between the groups, $P[W_S \geq 72] = .3605$. Therefore, the data did not give evidence that justified rejecting H_0 at the .05 level of significance. These data do not statistically support the hypothesis that the standardized condition has more information about the intent or actions of the other aircraft that result in higher understanding of the situation.

4.4.1.4 Situation Awareness

The hypothesis was that Captains operating in the standardized condition would have higher SA because of the additional information provided, and therefore, report higher ratings. The comparison was in terms of the ratings for SA on the SART. A higher rating would indicate a higher level of SA. The difference in the SA ratings was not statistically significant between the groups, $P[W_S \geq 80] = .1172$. Therefore, the data did not give evidence that justified rejecting H_0 at the .05 level of significance. These data do not support the hypothesis that Captains operating in the standardized condition have higher SA because of the information provided.

In summary, Captain SART ratings supplied no statistical evidence to support the hypothesis that the standard condition (i.e., the SOP) increased pilot SA. However, visual inspection of the different dimensions shows a pattern of ratings that suggests the SOP may have impacted SA in a positive way. In general, for Captains, there appears to be a slightly lower demand of attention in the standard condition, an equal supply of attention between conditions, a slightly higher understanding of the situation in the standard condition, and a slightly higher overall SA in the standard condition.

4.4.2 First Officer Results

First Officer SA was also compared between the no standard and standard group. The hypotheses for each dimension of the First Officers were the same as the Captains': SA would be improved with the standard procedure. For example, in the standard group, there would be a lower demand of attention. We tested the hypotheses by comparing the two groups' ratings. First Officers provided SA ratings for supply, demand, understanding, and overall SA.

A two-tailed Kolmogorov-Smirnov two-sample test was selected. Distributions of the ratings given by the two groups appeared to be different (in more than just central tendency) when examined visually. Therefore, the Kolmogorov-Smirnov, which is sensitive to any differences in the distributions of the samples, was desirable. With an $\alpha = .05$, $m = 7$, and $n = 8$, the largest discrepancies between the two cumulative distributions of demand of attention ($D_{m,n} = .25$), supply ($D_{m,n} = .59$), understanding ($D_{m,n} = .16$), and overall SA ($D_{m,n} = .18$) were not enough to exceed their respective critical values. Therefore, the null hypothesis could not be rejected.

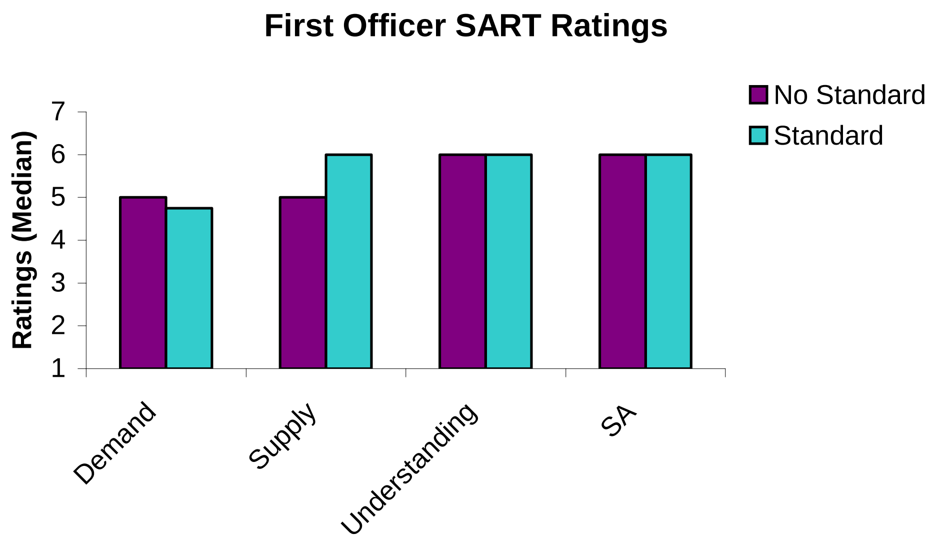


Figure 15. First Officer SART ratings.

In summary, First Officer SART ratings supplied no statistical evidence to support the hypothesis that the standard condition (i.e., the SOP) increased pilot SA. Visual inspection of the data indicates that for First Officers there appears to be a slightly lower demand of attention in the standard condition, a higher supply of attention in the standard condition, and an equal understanding of the situation and overall SA between conditions.

4.5 Workload

In addition to the SA questions used on the three-dimensional version SART (Taylor, 1989), one workload question was included. We asked crewmembers to select a rating from 1 to 7 (1 = Low, 7 = High) on the SART for the following question:

- Workload: How high was your workload because of the number of tasks you had to perform in the amount of time available?

The hypothesis was that the subjective workload reported by crewmembers in the standard condition would be more than that of those in the no standard condition. The comparison was in terms of what their ratings were for this dimension on the SART. The difference in the workload ratings was not statistically significant between the groups for Captains or First Officers. These data do not statistically support the hypothesis that crewmembers operating in the standardized condition have higher workload. There was no evidence (statistical or otherwise) that indicated the SOP affected workload.

4.6 Background Questionnaires

A total of 32 pilots acted as either Captain or First Officer in the AL^{2}ERTS study. All pilots were asked to complete a Background Questionnaire to provide researchers with information about their experience, range of skill, and other attributes. The results indicated that pilot participants varied in terms of demographics and experience. Of the 32 participants, 44% were active Title 14 of the Code of Federal Regulations Part 121 pilots; the remaining 56% were either furloughed or retired (for an average of 8 months). Participant ages ranged from 37 to 61, and all were male. The average total flying experience ranged from 9 to 43 years, with a mean of 32 years. Total years of B747 experience ranged from 1 to 31 years (with an average of 6 years 5 months). The active Part 121 pilots averaged 424 hours in the past 12 months. Sixteen percent of the participants did not have current certification to operate a B747. Of the 16 crews, 12 were from UAL and conformed to UAL procedures. The remaining four crews were mixed and generally adopted the cockpit configuration and lighting policy of the Captain's company (one Cathay Pacific crew, two Northwest Airlines (NWA) crews, and one UAL crew). The pilot conforming to their counterpart's procedures was sufficiently trained to mitigate any possible effects. According to observer logs, all participants remained motivated to participate throughout the study.

4.7 Post-Simulation Questionnaires

The Post-Simulation Questionnaires yielded many interesting results. All of the participants provided positive feedback overall. They all reported that they enjoyed the simulation and believed the standardized lighting procedures would be beneficial to real world operations.

Standard group participants were given more detailed questionnaires than the no standard group to draw out feedback about their experience with the new SOPs.

4.7.1 Safety

One hundred percent of the standard group participants indicated that they believed that the SOPs increased safety to some degree and felt confident using the new procedures. Participants' median response was 7 (1=Decreased Safety, 7=Increased Safety), or Increased Safety, when asked what effect, if any, the standardized landing light procedures had on runway traffic safety. Furthermore, none of the respondents indicated a neutral response or felt that safety was degraded in any way. These scores are consistent with debriefing comments and questionnaire feedback provided by the participants. When asked to expand upon the safety aspect of using standardized lighting procedures, 100% of the participants said that another aircraft's intent was more clearly indicated with a universal lighting system. They also indicated that they were able to react more quickly to potential situations than if they had to ascertain that there was a conflict by using the movement cue alone.

4.7.2 Confidence in Standardized Procedures

Participants reportedly felt confident using the SOPs. One hundred percent of the participants responded 5 or higher, with a median of 7, when asked if they were confident utilizing the standardized lighting procedures (1=Not at All, 7=A Great Deal). Scores strongly suggest that no participants had reservations relying on the lighting procedures as a messaging system in the simulation. These scores were also consistent with debriefing comments. Participants often commented that with good training and consistent use of the lighting procedures, the SOPs would provide a tremendous benefit overall.

4.7.3 Situation Awareness

Interestingly, the median response for the standard and no standard conditions in relation to the overall SA question was 6 (1=Low, 7=High). No scores were below 4 ("neutral") for either environment. Eighty-seven percent of those using SOPs reported "high" SA (6 or 7), compared to 81% for those that used their current (no standard) procedures, suggesting that poor SA was not an issue for either condition. However, during debriefs standard group participants consistently related that their SA was higher than that of the real world because they believed that the consistent messaging system enabled them to read other aircraft's intentions more clearly and they were able to more confidently predict other aircraft's movement. Figure 16 depicts frequency of response data for perceived levels of overall SA for both standard and no standard environments.

Q. Rate your perceived level of overall situation awareness during runs.

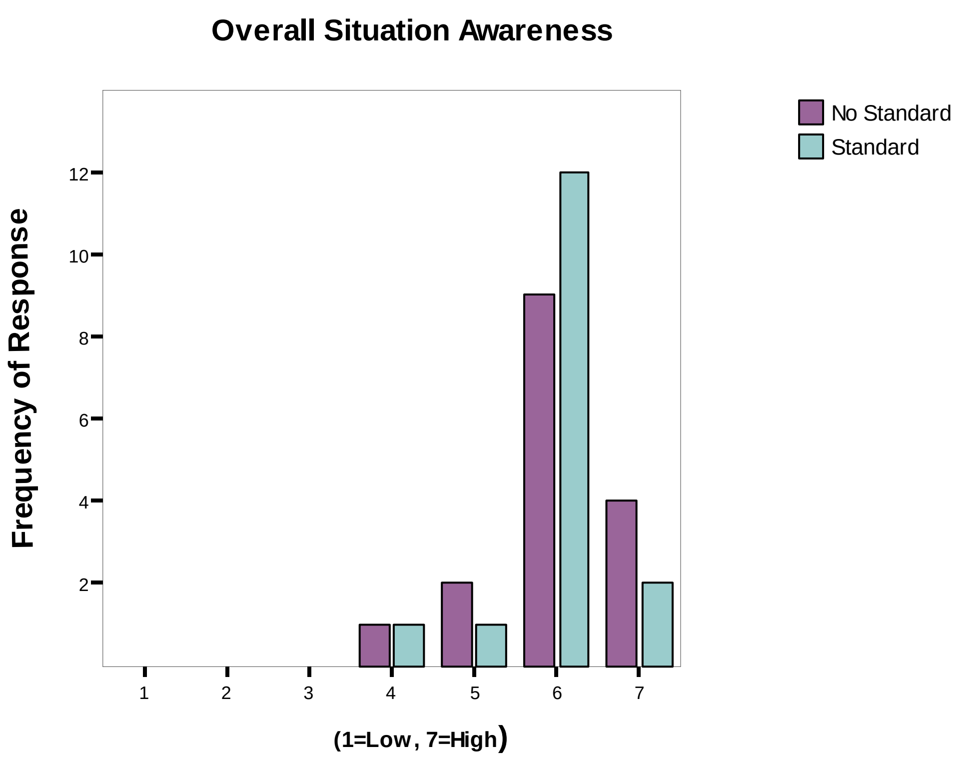


Figure 16. Overall Situation Awareness ratings.

4.7.4 Workload

Questionnaire analysis indicated that reported workload increases were largely contributed to factors other than the standardized lighting procedures. Participants who did report that the SOPs contributed to workload believed that with time and practice workload would decrease.

The crews were asked to rate their perceived level of overall workload during the runs (1= Extremely Low, 7= Extremely High). The standard group median of 4.5 revealed a very slight increase in perceived workload compared to the no standard median of 4, suggesting a neutral stance.

Participants that scored workload as a 4 or above were asked to explain. Some participants viewed a score of 4 as normal everyday operations; others believed that crossing multiple runways in a short period of time increased workload; several participants blamed unfamiliarity with ORD airfield and heightened vigilance for the increased score. Two standard group participants said, "[it] took some attention to remember to set the lights correctly."

However, when asked specifically how following the standardized landing light procedures affected standard crews' workload (1=Decreased Workload, 7=Increased Workload) the participant median indicated a *neutral* response of 4, indicating that the SOPs were not specifically contributing to workload. Four participants rated workload to be a 5, slightly above neutral; only one individual felt the workload was substantial enough to warrant a rating of 6. When asked to explain increases in standard group workload, the most common response was that multiple runway crossings in a short period of time resulted in a lot of switch manipulation/ flipping. However, when the seven participants who believed there was an increase in workload while using the standard procedures were asked if they believed that the increase in workload would decrease with time and practice, the median was 6 on a scale where 1= Not at All and 7=A Great Deal. In addition, 100% of the respondents believed workload would decrease to some degree; 93% believed it would decrease a good to Great Deal.

4.7.5 Realism and Training Summary

Nearly all participants felt that the simulation was realistic and the training was adequate to competently navigate the scenarios. The median response was 6 (1=Extremely Unrealistic, 7=Extremely Realistic) for realism of hardware, software, traffic runs, airport environments, and overall realism. Furthermore, 84% of the participants rated the "overall realism" of the simulation 5 or higher. The median rating for training was also 6 (1=Extremely Poor, 7=Extremely Good), with 97% of the participants rating the adequacy of training 5 or higher. Negative comments pertaining to realism generally reflected known limitations that we briefed to each participant each morning before the start of simulation. These comments related to the frequencies not being broken up between Tower and Ground, background chatter and traffic levels not being as substantial as those seen in real world operations at SFO and ORD, and color intensity and graphics in the simulation not representing objects as sharp they would appear in the real world.

4.8 Crew Debriefing Commentary

As previously mentioned, the focus of this study was to provide objective data as evidence to support or disprove the value of implementing a standard use of exterior aircraft lighting, particularly the use of aircraft landing lights. However, some subjective comments are included in the results as supplementary information. As such, we drew several observations from the debriefing sessions of the crews that experienced the standard conditions. In general, pilot reactions to the SOPs were very positive.

- All crews commented during the debrief sessions that they experienced increased SA when using the standard lighting procedures. They contributed the increase to clear communication of other aircrafts' intent.
- All crews agreed the overt signals communicated by the standard lighting procedures would provide a great benefit in the real world as long as the procedures were used correctly and consistently.
- All crews thought it was a good idea to implement the standard lighting procedures across all airlines and aircraft.

We also extracted pilot feedback from post-simulation questionnaires. In particular, pilots that flew in the standard conditions were asked how the standardized landing light procedures affected runway traffic safety during the simulation, if at all. All replies were positive and included responses such as the following: "You were able to have better situation awareness and increased reaction time", "Once the light system is understood, detection of movement is enhanced. Also, flight crew's intentions are indicated earlier," "Lights visible long before movement," "Provided early indication of intentions of other aircraft to either cross a runway or commence takeoff roll," and "Gave a good overt indication of what to expect from other aircraft."

This feedback from the professional pilots that had the opportunity to experience the affect of the new SOP in a high fidelity simulation environment clearly indicates favorable views toward the procedures.

5. Discussion

Sixteen crews composed of a Captain and First Officer participated in our study. They were instructed to taxi, depart, or land in 16 scenarios. Half of the crews operated in a baseline condition that had no standard procedures for using landing lights to indicate that aircraft are cleared to depart (no standard condition); the other half operated in an environment with standard procedures (standard condition). Both conditions included four scenarios in which a scripted confederate aircraft committed an error or followed erroneous instructions that resulted in a RI with the potential to result in an accident if not detected by the subject crew. We compared crews in the no standard condition and the standard condition in terms of their response to these scripted RIs. We will discuss the evidence that suggests that the standardized use of landing lights prevented the occurrence and the reduced severity of incursions, increased situation awareness, provided a direct message, and was accepted by the crews.

5.1 Prevention of Incursions

The first stated goal of this study was to investigate whether standardizing the use of aircraft landing lights to indicate that aircraft were cleared to depart prevented or reduced the severity of RIs or accidents. In the four data collection scenarios with scripted incursions, the data showed that crews taxiing in an environment with a standard use of landing lights held-short more frequently (thereby preventing more incursions) than those with no standard. Decomposition of the incursions showed that crews operating in an environment with no standard crossed the runway with greater frequency and were involved in more collisions.

The severity ratings analyses showed similar results. Figure 5 shows that the number of incursions was greater for crews operating without a standard use of landing lights and were also more severe in general. In the scenario where the standard group exceeded the no standard group in the number of incursions, all of the severity ratings were D for the standard group; the no standard group experienced A and C severity level incursions. Although we were unable to state that the differences discussed in this section were statistically significant, the pattern of results was consistently positive for an environment with a standardized use of landing lights.

5.2 Situation Awareness

Researchers generally believe that increases in SA will result in some type of operationally relevant improvement. SA itself is a construct that must be measured indirectly. In this study, we took subjective and objective measurements of SA and used them as evidence to indicate whether standardizing lighting procedures resulted in more or less SA. We considered all of the measures in tandem with the final desired result being a decrease in the occurrence or severity of RIs. When examined together, the evidence suggests that standard lighting procedures increased pilot SA. We discuss the patterns of SART ratings, response initiation time, and RIs as the evidence in the following paragraphs.

First, the pattern of responses on the SART for Captains suggested an increase in SA. There appears to be a slightly lower demand of attention in the standard condition, an equal supply of attention between conditions, a slightly higher understanding of the situation in the standard condition, and a slightly higher overall SA in the standard condition. Although the differences between standard and no standard conditions were not statistically significant, the results follow a pattern that we expected from an increase in SA. Furthermore, one of the limitations stated for this simulation was that the crews' awareness was already heightened because of the setting. To have enough power to show a statistical difference in SA under these circumstances, the design would have likely required an impractical increase in the sample size. Responses from First Officers were inconclusive.

Second, response initiation times suggested an increase in SA. We selected response initiation time as the accompanying measure to the SART. We paired this objective measure with the SART to make conclusions about SA because we feel that such a pairing is superior to relying on self-reports alone. Crews initiated a response about seven seconds faster in the standard condition; the difference was statistically significant. The faster response initiation time supports a higher SA in the crews operating in the standard environment.

When looking at the First Officer data, there is evidence for the standard lighting message; however, there were some inconsistencies in the data. For example, several First Officers reported that the SA ratings they provided reflected tasks that they performed inside the cab as their Captain taxied the 747. That would explain the difference in the pattern of responses between the two crewmembers and the lack of differences in some of the SART rating dimensions that we expected to be present. However, First Officers in the standard condition detected conflicts more often than their Captains while those in the no standard condition detected conflicts at about the same rate as their Captains. One would expect to see this reflected in the SART ratings, but First Officers in the standard condition only reported a significantly higher supply of attention. That was an unexpected difference in the SART ratings with no obvious reason.

5.3 Lights as a Direct Message Benefit

For the standard condition, when lights were attached with meaning, lights were the primary cue of an impending incursion while movement was the primary cue for the no standard condition. What is more interesting is the fact that lights were identified as a primary cue more frequently in the no standard runs as well, despite the fact that lights are used differently across airlines. Lights resulting in a primary cue for the no standard condition, in addition to common detection of lights in the standard condition, strongly suggest that lights are noticeable and send a strong signal of intent when concrete meaning is attached to them. It is well known that movement of an aircraft is difficult to detect. The supplemental data from this study (i.e., pilot feedback, audio and video tapes, E/O observations) also supports the observation that the onset of landing lights is a more noticeable signal that a departing aircraft is moving.

It is interesting to note that although potential light cues were only presented 50% of the time in the no standard condition, both Captains and First Officers still notably reported them as first cues. However, because lights were used inconsistently by different airlines in these scenarios, they could not be interpreted as a clear signal of intent. In other words, these crews could not reliably count on lights alone as an indication of an impending incursion.

Reported first cue results indicate that crewmembers in the standard conditions could rely (and usually did) on the lights as a primary cue instead of movement. This is evident because even though lights and movement cues were actually provided simultaneously in the standard condition, the data shows that pilots flying in this environment reported movement as the least common first cue. When both were consistently provided, lights were more likely to be detected earlier as the first cue because they were visible, implied clear meaning, and because it was difficult for pilots to discern initial movement of aircraft thousands of feet away. When potential light cues were only provided 50% of the time and without the understanding that this cue had a specific meaning (as they were in the no standard condition that intended to emulate current operations), the results were noticeably different. Crewmembers in the no standard conditions could not consistently or exclusively rely on lights because the companies used varying procedures and because there is no implicit message indicated by their use. Therefore, the data implies that crews operating in a standard environment (vs. no standard) had more time to react to incursions because "lights" cues were provided consistently, they had clear meaning of intent, and they were more easily detectable as a first cue before "movement".

5.4 Crewmember Roles in Detecting Incursions

We could not identify any clear trends for the pilot who first detected incursions or for same/opposite side identification of the impending incursions. The lack of consistent trends in these data suggest that both pilot position and aircraft side play only a small part in the detection of RIs. While First Officers in the standard condition were more likely to identify the potential incursion than the Captains (First Officers = 72% and Captains = 28%), Captains and First Officers in the no standard conditions were almost equally as likely to detect the impending incursion (First Officers = 48%, Captains = 52%). It is possible that task division differed in the standard condition allowing the First Officers to thoroughly scan the runways while Captains tended to other duties. However, it should be noted that upon inspection of the videotapes, differences between crewmember detection times were often off by only a fraction of a second. In the standard and no standard conditions, the crewmember stationed on the same side of the aircraft as the planned incursion was more likely to detect the occurrence of an incursion than the pilot on the opposite side, (62% vs. 38%, and 56% vs. 44% respectively). These results suggest that both crewmembers are extremely crucial to the detection of potential incursions.

5.5 User Acceptance

Some beneficial results from this study were knowledge of the pilots' acceptance of the procedure, that they felt it contributed to safety, and that it conveyed the message effectively. One hundred percent of the standard group participants indicated that they believed that the SOPs "Increased Safety" to some degree and felt confident using the new procedures. When asked to expand upon the safety aspect of using standardized lighting procedures, 100% of the participants said that another aircraft's intent was more clearly indicated with a universal lighting system and they were able to react more quickly to potential situations than if they had to ascertain that there was a conflict by using movement alone. The scores also strongly suggested that no participants had reservations relying on the lighting procedures as a messaging system in the simulation.

6. Conclusions

Subjective and performance data were analyzed to explore the procedural use of aircraft landing lights as a direct message to other pilots indicating that aircraft were cleared to depart. In general, the pattern of results supports the standardized use of the landing lights. Conclusions for this study are largely supported by the combination of observed patterns and trends in the data and subjective information gathered from the participants. In summary, the data shows that crews in the standard condition held-short more frequently, generally experienced less severe RIs when those occurred, initiated a response to RIs significantly faster, used the landing lights effectively as a first cue, and unanimously felt that safety was increased because of the standardized procedures.

The data showed with marginal significance that crews taxiing in an environment with a standard use of landing lights held-short more frequently (thereby preventing more incursions) than those with no standard. Crews with no standard crossed the runway with greater frequency and were involved in more collisions. As observed in the descriptive data, crews generally experienced incursions that were more severe when operating without a standard use of landing lights. The effect of having and not having a standardized use of the landing lights showed itself through repeated safe behaviors and unsafe behaviors respectively.

Landing lights provided a faster cue that there was potential for a collision than movement. When examining an aircraft on an intersecting runway, the primary concern of the crew was whether that aircraft was moving. If movement were easy to perceive, it would be sufficient as a cue that there is a potential for a collision. However, movement of the other aircraft is difficult to perceive because it is at such a great distance and its speed increases exponentially as it takes off (Regan, 1997). The standard procedure, however, provides an easy message: landing lights *on* means the aircraft may be moving and landing lights *off* means that it is not likely moving. Given that landing lights are more easily detected than movement (as simulated), they provide a message faster than movement (see section 4.3.1).

Initial response time results suggest that an optimal level of SA occurs sooner. The 3D SART rating trend of responses showed a slight increase in SA for Captains, although there may not have been enough power to detect the difference statistically. However, the objective supplementary measure to SA, initial response times to impending incursions involving a departing aircraft, were significantly faster for crews taxing in the standard condition. Given accurate knowledge of events in the environment (namely, an aircraft departing), a faster response means greater safety.

Feedback from the professional pilots who had the opportunity to experience the effects of the new procedure in this high fidelity simulation environment clearly indicated favorable views toward standardizing the use of landing lights. They felt it contributed to safety, increased their awareness of the intentions of other aircraft, and that it conveyed the intended message effectively.

In this simulation, we demonstrated the benefits of this procedure in an ideal environment where the complexity was relatively low and the lights were always visible. However, we need to conduct further simulations to determine the effects of various other factors, including consistency of the message, message conspicuity, and effects of the message on other human system elements. We need to determine the effects of providing an inconsistent message: for example, what is the effect on safety if there is not universal use of the procedure? We also need to determine under what conditions the lights are not visible (e.g., bad weather, acute angles, great distances) so that we can develop alternatives to maintain the same level of safety at intersections. Since conspicuity while holding in position is a prevalent concern for pilots, the use of pulse lighting (while in position and holding) used in combination with lights on steadily (upon departure clearance) can be evaluated as to whether it will provide an opportunity to improve perceived conspicuity while conveying the intent to begin takeoff roll upon receipt of a departure clearance. Further research on pulse lighting could also gather information on acceptability of these lights and their use from the pilot community. The use of aircraft landing lights to convey intent may provide valuable cues to ATC as well as to other pilots. We may conduct further simulations to determine if the standard use of landing lights has an additional positive effect on safety by extending the message of intent to this third party. Now that we have shown the benefits of the procedure under ideal conditions, we must find the conditions that are the exceptions and determine whether we can make any improvements.

REFERENCES

Federal Aviation Administration Advisory Circular 120-74A. (2003 September 26). *Parts 91, 121, 125, And 135 Flightcrew Procedures During Taxi Operations.* Washington, DC

Gawron, V. J. (2000). *Human performance measures handbook*. New Jersey: Lawrence Earlbaum Associates.

Federal Aviation Administration. (2002) *Runway incursion joint safety implementation team: Results and analysis.* Washington, DC

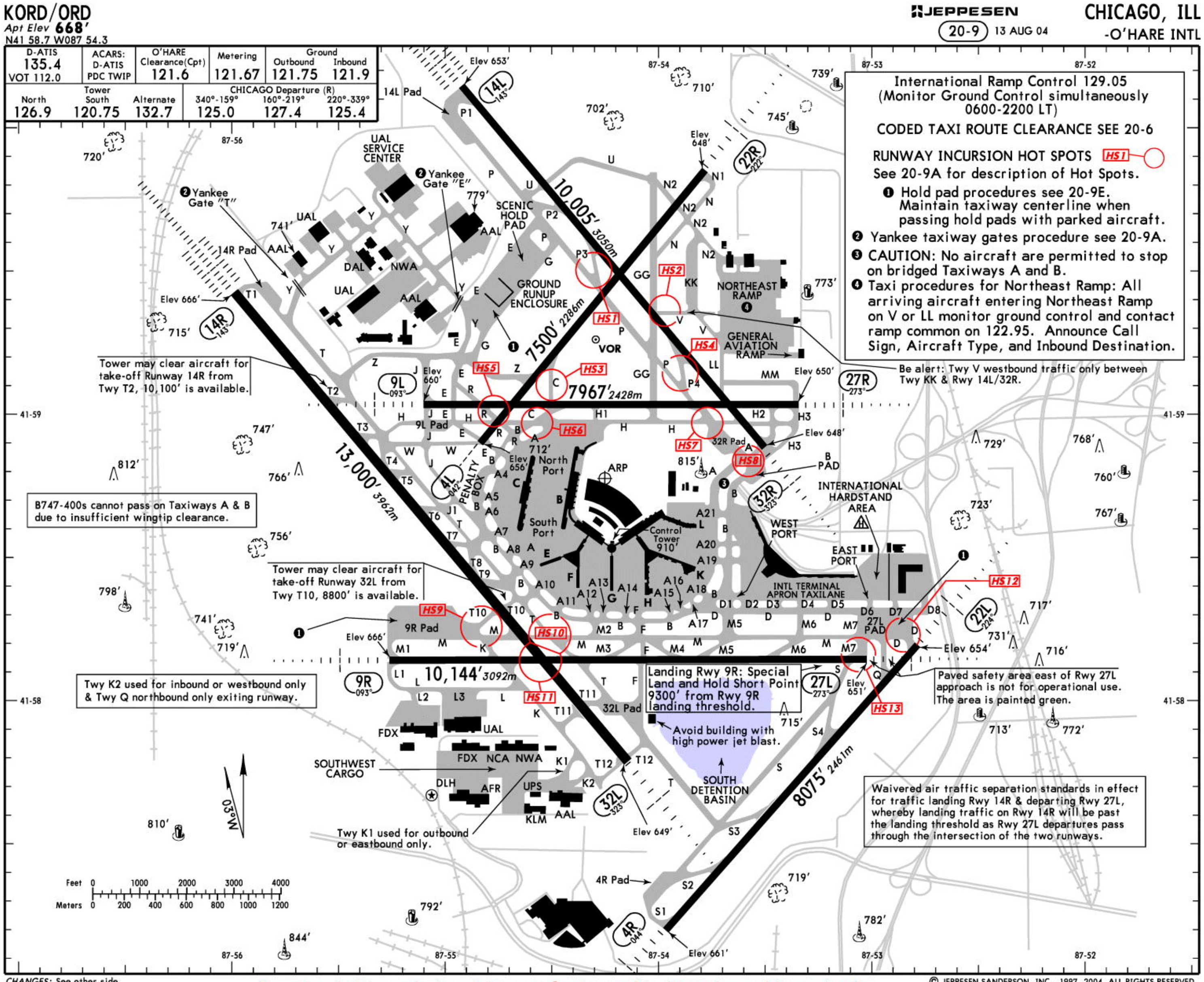
Regan, D. (1997). Perceptual motor skills and human motion analysis. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (pp. 174-218; 2nd ed). New York: John Wiley & Sons, Inc.

Sullivan, B. T., & Soukup P. A., (1996 July). *The NASA 747-400 flight simulator. National resource for aviation safety research* (AIAA-96-3517). Paper presented at the American Institute of Aeronautics and Astronautics, San Diego, CA.

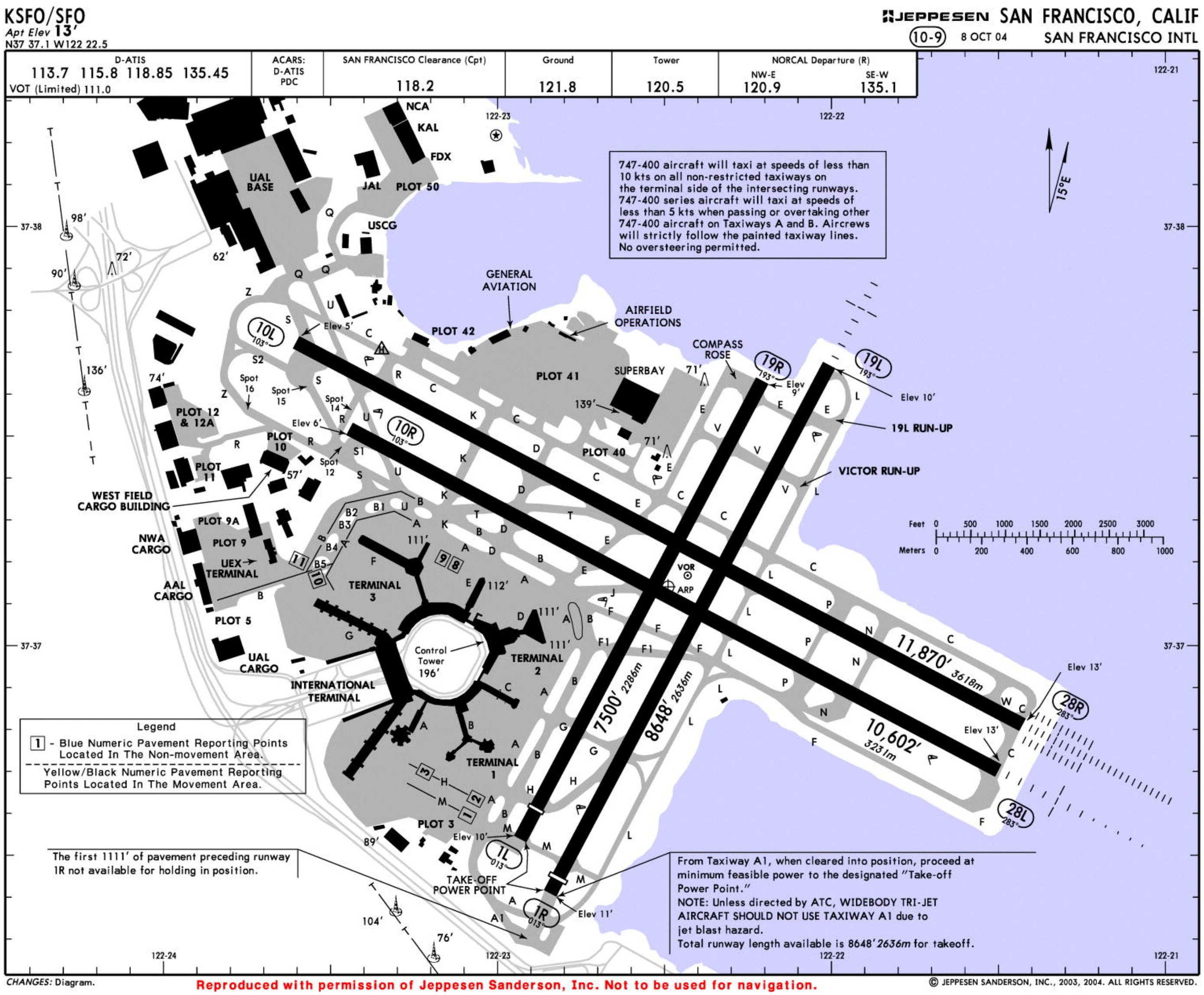
Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew system design. *Situational Awareness in Aerospace Operations*, AGARD-CP-478, 3-1 – 3-37.

APPENDIX A –SELECTED AIRPORTS – ORD & SFO

(The Jeppesen® airport diagrams of ORD and SFO found in this appendix are for simulation purposes only. They are not intended for commercial use or navigation purposes.)



KORD/ORD
Apt Elev 668'
N41 58.7 W087 54.3
JEPPESEN
20-9 13 AUG 04
CHICAGO, ILL
-O'HARE INTL
D-ATIS 135.4
VOT 112.0
ACARS: D-ATIS PDC TWIP
O'HARE Clearance(Cpt) 121.6
Metering 121.67
Ground Outbound 121.75 Inbound 121.9
Tower North 126.9 South 120.75 Alternate 132.7
CHICAGO Departure (R) 340°-159° 125.0 160°-219° 127.4 220°-339° 125.4
International Ramp Control 129.05
(Monitor Ground Control simultaneously 0600-2200 LT)
CODED TAXI ROUTE CLEARANCE SEE 20-6
RUNWAY INCURSION HOT SPOTS HS1
See 20-9A for description of Hot Spots.
❶ Hold pad procedures see 20-9E. Maintain taxiway centerline when passing hold pads with parked aircraft.
❷ Yankee taxiway gates procedure see 20-9A.
❸ CAUTION: No aircraft are permitted to stop on bridged Taxiways A and B.
❹ Taxi procedures for Northeast Ramp: All arriving aircraft entering Northeast Ramp on V or LL monitor ground control and contact ramp common on 122.95. Announce Call Sign, Aircraft Type, and Inbound Destination.
Be alert: Twy V westbound traffic only between Twy KK & Rwy 14L/32R.
Tower may clear aircraft for take-off Runway 14R from Twy T2, 10,100' is available.
B747-400s cannot pass on Taxiways A & B due to insufficient wingtip clearance.
Tower may clear aircraft for take-off Runway 32L from Twy T10, 8800' is available.
Twy K2 used for inbound or westbound only & Twy Q northbound only exiting runway.
Landing Rwy 9R: Special Land and Hold Short Point 9300' from Rwy 9R landing threshold.
Avoid building with high power jet blast.
Paved safety area east of Rwy 27L approach is not for operational use. The area is painted green.
Waivered air traffic separation standards in effect for traffic landing Rwy 14R & departing Rwy 27L, whereby landing traffic on Rwy 14R will be past the landing threshold as Rwy 27L departures pass through the intersection of the two runways.
Twy K1 used for outbound or eastbound only.
UAL SERVICE CENTER
SCENIC HOLD PAD
GROUND RUNUP ENCLOSURE
NORTHEAST RAMP
GENERAL AVIATION RAMP
INTERNATIONAL HARDSTAND AREA
INTL TERMINAL APRON TAXILANE
SOUTH DETENTION BASIN
SOUTHWEST CARGO
PENALTY BOX
North Port
South Port
WEST PORT
EAST PORT
Control Tower 910'
VOR
ARP
10,005' 3050m
7500' 2286m
7967' 2428m
13,000' 3962m
10,144' 3092m
8075' 2461m
14L Pad
14R Pad
9L Pad
9R Pad
32R Pad
32L Pad
4R Pad
27L PAD
B PAD
Yankee Gate "T"
Yankee Gate "E"
Feet 0 1000 2000 3000 4000
Meters 0 200 400 600 800 1000 1200
CHANGES: See other side.
© JEPPESEN SANDERSON, INC., 1997, 2004. ALL RIGHTS RESERVED.



KSFO/SFO
Apt Elev 13'
N37 37.1 W122 22.5
JEPPESEN SAN FRANCISCO, CALIF
10-9 8 OCT 04 SAN FRANCISCO INTL
D-ATIS 113.7 115.8 118.85 135.45
VOT (Limited) 111.0
ACARS: D-ATIS PDC
SAN FRANCISCO Clearance (Cpt) 118.2
Ground 121.8
Tower 120.5
NORCAL Departure (R) NW-E 120.9 SE-W 135.1
747-400 aircraft will taxi at speeds of less than 10 kts on all non-restricted taxiways on the terminal side of the intersecting runways. 747-400 series aircraft will taxi at speeds of less than 5 kts when passing or overtaking other 747-400 aircraft on Taxiways A and B. Aircrews will strictly follow the painted taxiway lines. No oversteering permitted.
GENERAL AVIATION
AIRFIELD OPERATIONS
COMPASS ROSE
19L RUN-UP
VICTOR RUN-UP
UAL BASE
USCG
PLOT 42
PLOT 41
SUPERBAY
PLOT 40
PLOT 12 & 12A
PLOT 10
PLOT 11
WEST FIELD CARGO BUILDING
PLOT 9A
PLOT 9
NWA CARGO
UEX TERMINAL
AAL CARGO
PLOT 5
UAL CARGO
TERMINAL 3
Control Tower 196'
INTERNATIONAL TERMINAL
TERMINAL 2
TERMINAL 1
PLOT 3
TAKE-OFF POWER POINT
11,870' 3618m
10,602' 3231m
7500' 2286m
8648' 2636m
Legend
1 - Blue Numeric Pavement Reporting Points Located In The Non-movement Area.
Yellow/Black Numeric Pavement Reporting Points Located In The Movement Area.
The first 1111' of pavement preceding runway 1R not available for holding in position.
From Taxiway A1, when cleared into position, proceed at minimum feasible power to the designated "Take-off Power Point."
NOTE: Unless directed by ATC, WIDEBODY TRI-JET AIRCRAFT SHOULD NOT USE TAXIWAY A1 due to jet blast hazard.
Total runway length available is 8648' 2636m for takeoff.
CHANGES: Diagram.
Reproduced with permission of Jeppesen Sanderson, Inc. Not to be used for navigation.
© JEPPESEN SANDERSON, INC., 2003, 2004. ALL RIGHTS RESERVED.

APPENDIX B – SUMMARY OF SCENARIO CHARACTERISTICS

Summary of Scenario Characteristics

Scenario	Airport	Lights	NASA aircraft	Confederate aircraft	~Angle of intersection	Planned Incursion	Incursion side	~Visual distance
12	ORD	No Standard	taxi	departure	90°	Yes	Left	2500’
13	ORD	No Standard	taxi	departure	135°	Yes	Left	6500’
14	SFO	No Standard	taxi	departure	135°	Yes	Right	7000’
16	SFO	No Standard	taxi	departure	90°	Yes	Right	4000’
28	ORD	Standard	taxi	departure	90°	Yes	Left	2500’
29	ORD	Standard	taxi	departure	135°	Yes	Left	6500’
30	SFO	Standard	taxi	departure	135°	Yes	Right	7000’
32	SFO	Standard	taxi	departure	90°	Yes	Right	4000’

APPENDIX C – DETAILED SCENARIO EVENTS

Detailed Scenario Events

Label Key	
ORD - Chicago-O'Hare Int'l Airport	SFO - San Francisco Int'l Airport
L – Standard Light Procedures	NL – No Standard Light Procedures
I – Incursion	NI – No Incursion

Scenario	Label	Description of scenario events
12	ORD5-NL-I	Same as scenario 28 except all aircraft will use lights as prescribed in the Company Lighting Policy (Appendix G) for "no standard" scenarios.
13	ORD6-NL-I	Same as scenario 29 except UAL173 turns on landing lights as aircraft enters departure runway for takeoff. All other aircraft will use lights as prescribed in the Company Lighting Policy (Appendix G) for "no standard" scenarios.
14	SFO1-NL-I	Same as scenario 30 except UAL657 turns on landing lights as aircraft enters departure runway for takeoff. All other aircraft will use lights as prescribed in the Company Lighting Policy (Appendix G) for "no standard" scenarios.
16	SFO3-NL-I	Same as scenario 32 except all aircraft will use lights as prescribed in the Company Lighting Policy (Appendix G) for "no standard" scenarios.
28	ORD5-L-I	NASA747 IP at the preset taxi position runway 32L (Pushed back from C Concourse, north of A5, heading 204). The pilot has been briefed that he is standing by for Ground Control to contact him. NASA747 is instructed to taxi to runway 32L via taxiway A7, T, T10, and K, hold short of 32L. As NASA747 enters taxiway T, UAL166 departs runway 27L, followed by AAL267. As NASA747 approaches the hold line for 32L on T10, AAL671 departs runway 32L. COA772 (incurring aircraft) taxis into position on runway 32L. A clearance is issued to COA732 for takeoff clearance on runway 27L. NASA747 is issued a clearance to cross runway 32L. An incursion occurs when COA772, believing the departure clearance was for him, begins takeoff roll on 32L as NASA747 begins to cross runway 32L on T10. The incurring aircraft's landing lights come on as the aircraft begins takeoff roll on 32L. All other aircraft will use lights as prescribed in the Company Lighting Policy (Appendix G) for "standard" scenarios.
29	ORD6-L-I	NASA747 IP - Pushed back from the International Terminal abeam taxiway A19, heading 322. The pilot has been briefed that he is standing by for Ground Control to contact him. NASA747 is instructed to "Taxi to runway 14L, straight ahead to join "B" then turn left on "P", hold short of taxiway H. As NASA747 starts to taxi the scenario starts with AAL274 and NWA476 on approach to 09L. After AAL274 lands and clears the runway, COA 375 departs on 09L. After NWA476 lands on 09L UAL173 taxies into position for takeoff on 09L. NWA476 clears the runway at "P" and when clearing "P" onto "H", NASA747 is cleared to cross runway 09L. A Takeoff clearance is issued to UAL163 to depart on 09R. The incursion occurs when UAL173 starts its takeoff roll on 09L as NASA747 is crossing the same runway. The incurring aircraft's landing lights appear as aircraft begins takeoff roll. All other aircraft will use lights as prescribed in the Company Lighting Policy (Appendix G) for "standard" scenarios.

Scenario	Label	Description of scenario events
30	SFO1-L-I	NASA 747 IP – Pushed back from gate 86, heading 075. GC clears NASA747 to taxi to runway 28R via taxiways A, E, and C, hold short of taxiway E. As NASA747 taxies, UAL152 is cleared to land on 28L. UAL556 is number two to land 28L. As NASA747 turns onto taxiway E and UAL152 crosses 1L on landing roll, AAL253 is cleared for takeoff on runway 01L. UAL556 is cleared to land on 28L. As UAL556 crosses the 28L threshold, Local Control clears COA354 to taxi into position and hold 28R and UAL657 (incurring aircraft) to taxi into position and hold runway 28L When UAL556 passes taxiway E, Local Control clears NASA747 to cross 28L hold short of 28R for departing traffic. Local Control then clears COA354 for takeoff on 28R. Incursion occurs when UAL657 commences takeoff roll on 28L when COA354 was cleared for takeoff and NASA747 enters runway 28L while crossing. The incurring aircraft's landing lights appear as aircraft begins takeoff roll. All other aircraft will use lights as prescribed in the Company Lighting Policy (Appendix G) for "standard" scenarios.
32	SFO3-L-I	NASA747 IP – Pushed back from gate 86, heading 075. Ground Control clears NASA747 to taxi to runway 28R via A, F, hold short of runway 1L. Ground Control advises NASA747 that traffic is holding in position on runway 1R and clears NASA747 to cross runway 1L and runway 1R. NWA462 (incurring aircraft) takeoff position runway 1R. NWA462 begins takeoff roll (turns on landing lights) without takeoff clearance. Incursion occurs when NASA 747 enters runway 1R and NWA462 is commencing takeoff roll (landing lights go on as takeoff roll commences). All other aircraft will use lights as prescribed in the Company Lighting Policy (Appendix G) for "standard" scenarios.

APPENDIX D – E/O OVER THE SHOULDER FORM

E/O OVER-THE-SHOULDER FOR

Participant Codes:
Run Number:
Date: _____/_____/ 2003

Captain	First Officer
Incursion □ Detected incursion o ___first ___ second o ___verbal indication of detection o ___motor indication of detection o ___only confirmed after it was pointed out *Motivation during scenario* □ Positive (attentive and cooperative) □ Negative (alert, but uncooperative) □ Apathetic (just trying to get it over with)	*Incursion* □ Detected incursion o ___first ___second o ___verbal indication of detection o ___motor indication of detection o ___only confirmed after it was pointed out *Motivation during scenario* □ Positive (attentive and cooperative) □ Negative (alert, but uncooperative) □ Apathetic (just trying to get it over with)

Navigation Errors____

Other Events____

SIMULATOR PERFORMANCE

Out-the-Window display	Please describe
□ *Flawless*	
□ Flicker	
□ Distortion	
□ Target clearly out of place	
□ Other ______________________	

Controls	Please describe
□ *Flawless*	
□ No feedback	
□ No response	
□ Lag in response	
□ Unexpected response	
□ Other ______________________	

NASA747 Behavior	Please describe
□ *Flawless*	
□ Unrealistic performance	
□ Behaves as if heavier than specified	
□ Behaves as if lighter than specified	
□ Other ______________________	

APPENDIX E – E/O POST RUN FORM

E/O POST-RUN FORM

Participant Codes:
Run Number:
Date: _____/_____/ 2003

Observer, do not assume that crewmembers were aware that an incursion occurred if they do not mention an incursion in their description.

1. Did the B747 crew use proper light procedures? ☐ YES ☐ NO
2. **OBSERVER ASK**: DID YOU EXPERIENCE ANY UNFORSEEN EVENTS?

☐ No Event

☐ Event 1 Description____

☐ Event 2 Description____

3. **OBSERVER ASK:** AFTER YOU NOTICED THE PROBLEM, HOW LONG DID YOU WAIT TO REACT?

Event 1 Reaction____

Event 2 Reaction____

OBSERVER, PLEASE DRAW OUT WHAT CUED THEM.

4. The Captain indicated that an incursion occurred: ☐ YES ☐ NO
   a. If *YES,* describe the event____

____

   b. What cued the Captain (indicate 1st, 2nd, etc.)? __ Lights __Movement __Other _______

5. The First Officer indicated than an incursion occurred: ☐ YES ☐ NO
   a. If *YES*, describe the event____

____

   b. What cued the FO (indicate 1st, 2nd, etc.)? __ Lights __Movement __Other ___________

6. I (the observer) believe that an incursion occurred ☐ YES ☐ NO
   a. If *YES*, describe the event____

PEASE WRITE ANY ADDITIONAL COMMENTS ON BACK

APPENDIX F – EXCERPT FROM ADVISORY CIRCULAR

10. USE OF EXTERIOR AIRCRAFT LIGHTS TO MAKE AIRCRAFT MORE CONSPICUOUS.

a. General.

(1) Exterior aircraft lights may be used to make an aircraft operating on the airport surface more conspicuous. Pilots may use various combinations of exterior lights to convey their location and intent to other pilots. Certain exterior lights may also be used in various combinations to signal whether the aircraft is on a taxiway or on a runway, in position on the runway but holding for takeoff clearance, crossing an active runway, or moving down the runway for takeoff.

(2) Because adherence to the guidelines in this AC are voluntary and aircraft equipment varies, pilots are cautioned not to rely solely on the status of an aircraft's lights to determine the intentions of the pilot(s) of the other aircraft. Additionally, pilots must remember to comply with operating limitations on the aircraft's lighting systems.

b. Exterior Lights. To the extent possible and consistent with aircraft equipage, operating limitations, and pilot procedures, pilots should illuminate exterior lights as follows:

(1) Engines Running. Turn on the **rotating beacon** whenever an engine is running.

(2) Taxiing. Prior to commencing taxi, turn on **navigation, position, anti-collision, and logo lights, if available.** To signal intent to other pilots, consider turning on the taxi light when the aircraft is moving or intending to move on the ground, and turning it off when stopped, yielding, or as a consideration to other pilots or ground personnel. Strobe lights should not be illuminated during taxi if they will adversely affect the vision of other pilots or ground personnel.

(3) Crossing a Runway. All exterior lights should be illuminated when crossing a runway.

> **CAUTION: Pilots should consider any adverse effects to safety that illuminating the forward facing lights will have on the vision of other pilots or ground personnel during runway crossings.**

(4) Entering the Departure Runway for Takeoff. When entering a runway after being cleared for takeoff, or when taxiing into position and hold, pilots should make their aircraft more conspicuous to aircraft on final behind them and to ATC by turning on lights (except landing lights) that highlight the aircraft's silhouette. Strobe lights should not be illuminated if they will adversely affect the vision of other pilots.

> **NOTE: The SOP of turning on landing lights when takeoff clearance is received is a signal to other pilots, ATC, and ground personnel that the aircraft is moving down the runway for takeoff.**

(5) Takeoff. Landing lights should be turned on when takeoff clearance is received, or when commencing takeoff roll at an airport without an operating control tower.

11. SUMMARY.

a. Taxi operations require constant vigilance on the part of pilots. Pilots need to be continually aware of the movement and location of other aircraft and ground vehicles. Taxi operations require the same planning, coordination, and proper execution as other phases of flight operations. Sterile cockpit discipline is always appropriate while taxiing, even under normal weather conditions.

b. During low-visibility taxi operations, additional vigilance is absolutely essential. Pilots must pay particularly close attention to instructions from ATC and must insist on correct readback and hearback. Additionally, pilots should pay close attention to readback and hearback between ATC and other aircraft. Any ambiguity or uncertainty should be promptly resolved by clarification with ATC. When clear of an active runway, pilots should be prepared to stop in position to resolve any questions about position on the airport or clearance from ATC.

c. Safe aircraft operations can be accomplished and incidents eliminated if pilots are properly trained and correctly accomplish standard taxi operating procedures and practices.

APPENDIX G – COMPANY LIGHTING POLICIES

NOTE: Taxi, Logo, Runway turnoff, and Wing lights are not available in the CVSRF scenarios. Beacon (anti-collision) lights will be on whenever an engine is running for all companies.

☼ = lights on	**No Standard Procedures (per Current Policies)**			**Standard Procedures (per AC 120-74A)**		
	Navigation	Strobe	Landing	Navigation	Strobe	Landing
American						
Taxiing	☼			☼		
Crossing a runway	☼			☼	☼	☼
Entering the departure runway for Taxi Into Position & Hold (TIPH)	☼	☼		☼	☼	
Takeoff clearance received and/or commencing takeoff roll	☼	☼	☼	☼	☼	☼
Continental						
Taxiing	☼			☼		
Crossing a runway	☼	☼		☼	☼	☼
Entering the departure runway for TIPH	☼	☼		☼	☼	
Takeoff clearance received and/or commencing takeoff roll	☼	☼	☼	☼	☼	☼
Northwest						
Taxiing	☼			☼		
Crossing a runway	☼			☼	☼	☼
Entering the departure runway for TIPH	☼	☼		☼	☼	
Takeoff clearance received and/or commencing takeoff roll	☼	☼	☼	☼	☼	☼
United						
Taxiing	☼			☼		
Crossing a runway	☼	☼	☼	☼	☼	☼
Entering the departure runway for TIPH	☼	☼	☼	☼	☼	
Takeoff clearance received and/or commencing takeoff roll	☼	☼	☼	☼	☼	☼

APPENDIX H – LIGHTING ASSESSMENT

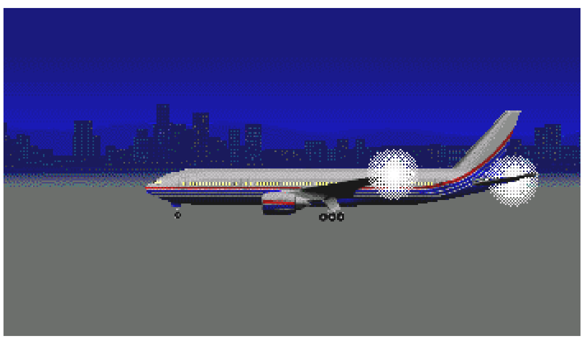
LIGHTING USAGE ASSESSMENT

Date _____/_____/ 2003
Participant Code ____________
Check One
□ Captain
□ First Officer

Please indicate when to use your **landing lights** throughout the scenarios.

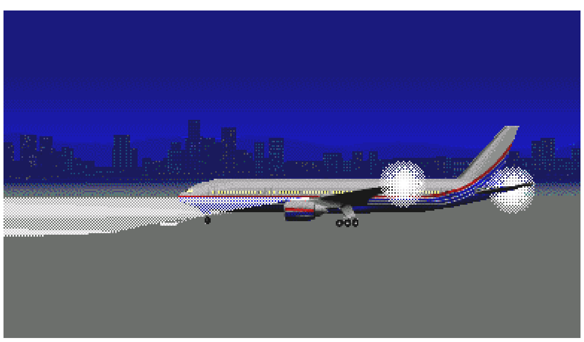
	(X = lights on)
Taxiing	
Crossing a runway	
Entering the departure runway for Taxi into Position & Hold (TIPH)	
Takeoff clearance received and/or commencing takeoff roll	

Using the aircraft's lighting as a guide, please indicate the aircraft's **intent** by placing a check mark next to all that apply.



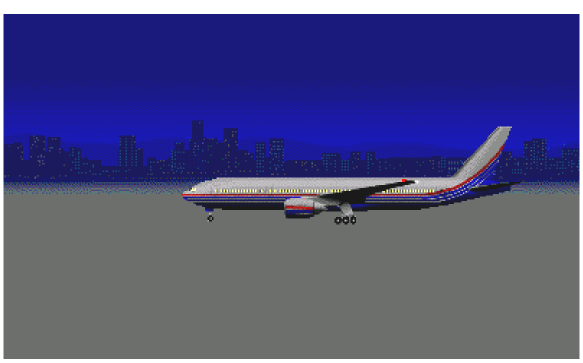
[Logo, anti-collision, navigation, and strobe lights on. No landing lights.]

- □ Taxiing
- □ Crossing a runway
- □ Entering the departure runway for TIPH
- □ Takeoff clearance received and/or commencing takeoff roll



[Logo, anti-collision, navigation, strobe, and landing lights on.]

- □ Taxiing
- □ Crossing a runway
- □ Entering the departure runway for TIPH
- □ Takeoff clearance received and/or commencing takeoff roll



[Logo, anti-collision, navigation, taxi, and strobe lights on. No landing lights.]

- □ Taxiing
- □ Crossing a runway
- □ Entering the departure runway for TIPH
- □ Takeoff clearance received and/or commencing takeoff roll

APPENDIX I – CONSENT

Ames Research Center

CATEGORY II – HUMAN RESEARCH MINIMAL RISK CONSENT

To the Research Participant: Please read this consent form and the attached protocol and/or subject instructions carefully. Make sure all your questions have been answered to your satisfaction before signing.

A. **I agree to participate** in the ____ research experiment as described in the attached protocol or subject instructions. I understand that I am employed by ___________________________ who can be contacted at ______________________________ .

B. **I understand that my participation could cause me minimal risk*, inconvenience, or discomfort.** The purpose and procedures have been explained to me and I understand the risks and discomforts as described in the attached research protocol.

C. **To my knowledge,** I have no medical conditions, including pregnancy, that will prevent my participation in this study. I understand that if my medical status should change while I am a participant in the research experiment there may be unforeseeable risks to me (or the embryo or fetus if applicable). I agree to notify the Principal Investigator (PI) or medical monitor of any known changes in my condition for safety purposes.

D. **My consent to participate has been freely given.** I may withdraw my consent, and thereby withdraw from the study at any time without penalty or loss of benefits to which I am entitled. I understand that the PI may request my withdrawal or the study may be terminated for any reason. I agree to follow procedures for orderly and safe termination.

E. **I am not releasing NASA from liability for any injury** arising as a result of my participation in this study.

F. **I hereby agree that all records collected by NASA in the course of this study are available** to the research study investigators, support staff, and any duly authorized research review committee. I grant NASA permission to reproduce and publish all records, notes, or data collected from my participation, provided there will be no association of my name with the collected data and that confidentiality is maintained, unless specifically waived by me.

G. **I have had an opportunity to ask questions and have received satisfactory answers** to all my questions. I understand that the PI for the study is the person responsible for this activity and that any questions regarding the research will be addressed to him/her during the course of the study. I have read the above agreement, the attached protocol and/or subject instructions prior to signing this form and I understand the contents.

* ***Minimal Risk*** means that the probability and magnitude of harm or discomfort anticipated in the research are not greater, in and of themselves, than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.

Signature of Research Participant Date	Signature of Principal Investigator Date
Printed/Typed Name of Research Participant	Printed/Typed Name of Principal Investigator
Address	Telephone Number of Principal Investigator
City, State, Zip Code	**Subject Signature:** Authorization for Videotaping
Telephone Number of Test Subject	**Subject Signature:** Authorization for Release of Information to Non-NASA Source(s)

ARC 475(Mar 01)E Previous editions are obsolete

APPENDIX J – BACKGROUND QUESTIONNAIRE

BACKGROUND QUESTIONNAIRE

The following information is requested for reporting data relevant to the simulation.

Your personal information will be kept completely confidential and will not be included in any of the reports or documents that will be produced as a result of this study. When necessary, individuals will be identified as Pilot A, Pilot B, etc.

Participant Code: __________

Date: ____________________

1. Are you currently an active Part 121 pilot?

   Circle one: YES NO

2. If you are not currently an active pilot, how long have you been furloughed/ retired?

   Years: ____________ Months:_______________

3. Are you currently certified to operate a B-747 aircraft?

   Circle one: YES NO

4. Please list your current ratings:

   ____
   ____
   ____
   ____

5. How much total experience do you have operating a B-747 aircraft?

   Years: ____________ Months: _______________

6. Please estimate your total Part 121 hours for the past 12 months.

   __________ Hours

7. What is your total experience as a pilot (all aircraft types)?

   Years: ______________ Months: ______________

8. For which airline do you currently work?

   ____

9. If you are currently furloughed or retired, for which airline did you last work?

____

10. With whose airline procedures will you abide when operating the B747-400 simulator today?

____

11. What is your age? ___________

11. What is your sex? __________

12. Do you wear eye glasses (clear) during departures, arrivals or while taxiing on the airport?

Circle one: YES NO

13. Do you wear sunglasses during departures, arrivals or while taxiing on the airport?

Circle one: YES NO

14. Please rate your previous computer experience.

No Experience	① ② ③ ④ ⑤ ⑥ ⑦	Very Experienced

15. Please rate your previous simulator experience.

No Experience	① ② ③ ④ ⑤ ⑥ ⑦	Very Experienced

APPENDIX K – SITUATION AWARENESS RATING FORM

Participant Number ____________
Check One
- □ Captain
- □ First Officer

For observer use only			
Date / /03	**Time** :	**PM/AM**	
Run Number			

(Circle One)

Item		**Low**						**High**
1	**Demand of attention** How demanding was the scenario on your attention?	1	2	3	4	5	6	7
2	**Instability of situation** How likely was the scenario to change suddenly (e.g., from usual operations to unusual events/operation, usual communications to unusual communications, etc.)?	1	2	3	4	5	6	7
3	**Complexity of situation** How complicated was the scenario?	1	2	3	4	5	6	7
4	**Variability of situation** How variable were the factors in the scenario (e.g., the change in the number of things going-on that required your attention)?	1	2	3	4	5	6	7
5	**Supply of attention resources** How much attention did you have available to devote to the scenario?	1	2	3	4	5	6	7
6	**Arousal** How alert and ready for action did you feel throughout the scenario?	1	2	3	4	5	6	7
7	**Concentration of attention** How concentrated were you on the scenario?	1	2	3	4	5	6	7
8	**Division of attention** How divided was your attention among different things going-on in the scenario (e.g., between operations inside the cockpit, other aircraft, communications, etc.)?	1	2	3	4	5	6	7
9	**Spare mental capacity** How much attention did you have left over to deal with new events, should they happen?	1	2	3	4	5	6	7
10	**Understanding of the situation** How well did you understand your situation as it was in this scenario?	1	2	3	4	5	6	7
11	**Information quantity** How much information (e.g., from ATC, instruments, visual, etc.) about your situation were you able to obtain throughout the scenario?	1	2	3	4	5	6	7
12	**Information quality** How good or valuable was the information (e.g., from ATC, instruments, visual, etc.) that you obtained throughout the scenario?	1	2	3	4	5	6	7
13	**Familiarity** How knowledgeable and familiar were you with the operation (e.g., taxiways, runways, traffic, communications, etc.) in the scenario?	1	2	3	4	5	6	7
14	**Situation awareness** Based on this scenario as it was simulated - how good was your ability to perceive elements in the environment (for example, your present situation, traffic movement, ATC interaction, etc.), to comprehend their meaning, and to project their status?	1	2	3	4	5	6	7
15	**Workload** How high was your workload because of the number of tasks you had to perform in the amount of time available?	1	2	3	4	5	6	7

APPENDIX L – POST-SIMULATION QUESTIONNAIRES

Post-Simulation Questionnaire

Date : _______
Participant Code: L______
Run: _______

Position worked (Please circle): Captain First Officer

Instructions: All questions are specific to this simulation. Answers will not be generalized to other pilots, airlines, airports, etc. Please answer the following based upon your overall experience in the demonstration. As always, your identity will remain anonymous.

Simulation

1. Did you observe any unforeseen events during the simulation?

Circle One YES NO

2. If you answered YES to the above question, please list the signals or cues that typically made you aware that an event was impending?

____
____
____
____

Standard Landing Light Procedures

3. During the simulation, what **effect**, if any, did the standardized landing light procedures have **on runway traffic safety**?	Decreased Safety	① ② ③ ④ ⑤ ⑥ ⑦ None	Increased Safety

4. Explain how the standardized landing light procedures affected runway traffic safety during the simulation, if at all.

____
____
____
____
____

5. During the simulation, did you feel **confident utilizing standardized landing light procedures** as a runway "messaging system"?	Not at All	① ② ③ ④ ⑤ ⑥ ⑦	A Great Deal

6. Rate your **consistency in conforming to the new procedures**.	Extremely Poor	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Good

7. Please explain any inconsistencies.

8. Please indicate **when you had landing lights on** througho throughout the scenarios.

	(X = lights on)
Taxiing	
Crossing a runway	
Entering the departure runway for Taxi Into Position & Hold	
Takeoff clearance received and/or commencing takeoff roll	

9. Rate the other aircraft in the scenario's **consistency in conforming to the new procedures**.	Extremely Poor	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Good

10. Please indicate **when the other aircraft had their landing lights on** throughout the scenarios.

	(X = lights on)
Taxiing	
Crossing a runway	
Entering the departure runway for Taxi Into Position & Hold	
Takeoff clearance received and/or commencing takeoff roll	

Workload

The term ***workload*** (as used in questions 11 - 15) refers to both the cognitive and physical demands imposed by your tasks.

11. Rate your **perceived level of overall workload** during runs.	Extremely Low	① ② ③ ④ ⑤ ⑥ ⑦	Extremely High

12. If you rated your overall workload level as a 4 or higher in question 11, please explain what contributed to your workload.

____
____
____
____

13. Did following the **standardized landing light procedures affect your workload**?	Decreased Workload	① ② ③ ④ ⑤ ⑥ ⑦	Increased Workload

14. If you perceived an increase in workload in question 13, please explain how the standardized landing light procedures contributed to your workload.

____
____
____
____

If you reported an increase in workload in question 13,

15. Do you believe the increase in **workload would decrease with time and practice**?	Not At All	① ② ③ ④ ⑤ ⑥ ⑦	A Great Deal

Situation Awareness

The term ***situation awareness*** (as used in question 16) refers to how well you were able to perceive the elements in the environment (for example, your present situation, traffic movement, ATC interaction, etc.), to comprehend their meaning, and to project their status.

16. Rate your perceived level of **overall situation awareness** during the runs.	**Low**	① ② ③ ④ ⑤ ⑥ ⑦	**High**

Simulation Fidelity

17. Rate the **realism of the overall simulation experience** compared to actual pilot operations.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

18. Rate the **realism of the simulated hardware** compared to actual equipment.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

19. Rate the **realism of the simulated software** compared to actual functionality.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

20. Rate the **realism of the simulated traffic runs** compared to actual National Airspace System (NAS) traffic.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

21. Rate the **realism of the simulated airport** compared to the actual airport.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

22. Do you have any comments or suggestions for improvement about our simulation capability?

____
____
____
____
____

Training

23. Rate the **adequacy of the training** you received for the simulation.	Extremely Poor	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Good

24. If your response to question 25 was "Poor" (rated 1-3), please identify the training deficiencies and describe what should be done to improve the training (e.g., longer practice sessions, more robust scenarios, etc.).

____
____
____

Comments

25. Is there anything about the study that we should have asked or that you would like to comment about?

____
____
____
____
____
____
____
____

Post-Simulation Questionnaire

Date : _______

Participant Code: NL______

Run: _______

Position worked (Please circle):   Captain   First Officer

Instructions: All questions are specific to this simulation. Answers will not be generalized to other pilots, airlines, airports, etc. Please answer the following based upon your overall experience in the demonstration. As always, your identity will remain anonymous.

Simulation

1. Did you observe any unforeseen events during the simulation?

   Circle One   YES   NO

2. If you answered YES to the above question, please list the signals or cues that typically made you aware that an event was impending?

____
____
____
____
____

Workload

The term ***workload*** (as used in questions 3 and 4) refers to both the cognitive and physical demands imposed by your tasks.

3. Rate your **perceived level of overall workload** during runs.	Extremely Low	① ② ③ ④ ⑤ ⑥ ⑦	Extremely High

4. If you rated your overall workload level as a 4 or higher in question 3, please explain what contributed to your workload.

____
____
____
____

Situation Awareness

The term ***situation awareness*** (as used in question 5) refers to how well you were able to perceive the elements in the environment (for example, your present situation, traffic movement, ATC interaction, etc.), to comprehend their meaning, and to project their status.

5. Rate your perceived level of **overall situation awareness** during the runs.	**Low**	① ② ③ ④ ⑤ ⑥ ⑦	**High**

Simulation Fidelity

6. Rate the **realism of the overall simulation experience** compared to actual pilot operations.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

7. Rate the **realism of the simulated hardware** compared to actual equipment.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

8. Rate the **realism of the simulated software** compared to actual functionality.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

9. Rate the **realism of the simulated traffic runs** compared to actual NAS traffic.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

10. Rate the **realism of the simulated airport** compared to the actual airport.	Extremely Unrealistic	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Realistic

11. Do you have any comments or suggestions for improvement about our simulation capability?

____
____
____
____
____

Training

12. Rate the **adequacy of the training** you received for the simulation.	Extremely Poor	① ② ③ ④ ⑤ ⑥ ⑦	Extremely Good

13. If your response to question 12 was “Poor” (1-3), please identify the training deficiencies and describe what should be done to improve the training (e.g., longer practice sessions, more robust scenarios, etc.).

____
____
____
____
____

General

14. What strategies, if any, could be implemented to help reduce the potential for runway incursions?

____
____
____
____
____
____

Comments

15. Is there anything about the study that we should have asked or that you would like to comment about?

____
____
____
____
____
____
____
____
____
____
____

APPENDIX M – DAILY PILOT SCHEDULE

The daily pilot schedule was subject to minor revisions.

Time (PST)	Daily Schedule
08:00am :15	Pilots arrive at security
08:30am :45	Official start: Pilot Briefing, Consent & Background Questionnaire
9:00am :15	Familiarization, Safety Briefing & Training
:30	Run1
:45	Run2
10:00am	Run3
:15	Run4
:30	***Break***
:45	***Break***
11:00am	Run5
:15	Run6
:30	Run7
:45	Run8
12:00pm :15 :30 :45 01:00pm	***Lunch***
:15	Run9
:30	Run10
:45	Run11
02:00pm	Run12
:15	***Break***
:30	***Break***
:45	Run13
03:00pm	Run14
:15	Run15
:30	Run16
:45	***Break***
04:00pm :15 :30	Post Sim Questionnaire & Debriefing