Summer 2012

The Effect of Control and Display Lag on UAS Internal Pilot Manual Landing Performance

Marshall Everett Lloyd
Embry-Riddle Aeronautical University - Daytona Beach

THE EFFECT OF CONTROL AND DISPLAY LAG ON UNMANNED AIR SYSTEM INTERNAL PILOT MANUAL LANDING PERFORMANCE

by

Marshall Everett Lloyd

B.S., Royal Military College of Canada, 2006

A Graduate Thesis Submitted to the

Department of Human Factors and Systems

in Partial Fulfillment of the Requirement for the Degree of

Master of Science in Human Factors and Systems

Embry- Riddle Aeronautical University

Daytona Beach, Florida

Summer 2012

The Effect of Control and Display Lag on UAS Internal Pilot Manual Landing Performance

by

Marshall Everett Lloyd

This thesis was prepared under the direction of the candidate's thesis committee chair, Kelly Neville, Ph.D., Department of Human factors and Systems and has been approved by the members of the thesis committee. It was submitted to the Department of Human and Systems and has been accepted in partial fulfillment of the requirements for the degree of Science in Human Factors and Systems.

THESIS COMMITTEE:

Kelly Neville, Ph.D., Chair

Dahai Liu, Ph.D., Member

Richard Stansbury, Ph.D., Member

MS HFS Program Director

Department Chair, Department of Human Factors and Systems

7-5-2012

Associate Vice President for Academics

Acknowledgements

This author would like to express his sincere gratitude and respect to his thesis chair, Dr. Neville. Her expansive knowledge in her field provided guidance and insight that enriched and strengthened this study. Without her influence, this study would have been trivial and ineffectual. Dr. Neville has been an excellent mentor who fostered intellectual growth in directions the author never expected.

The author would like to express his great appreciativeness to his other committee members—Dr. Liu and Dr. Stansbury. Dr. Liu has been an outstanding instructor whose exceptional knowledge and aspiring standards were a constant source of inspiration that contributed to the author acquiring numerous invaluable skills. Dr. Stansbury graciously devoted time from his busy schedule to make a meaningful contribution to this study.

The author would also like to express his many thanks to Matthew Grasso who graciously devoted countless hours of outstanding work to singlehandedly develop custom software for this experiment. Beyond any doubt, his most generous and impressive efforts ensured the success of this study.

Additionally, the author would like to acknowledge the entire Human Factors and Systems department—faculty, staff, and students. They have made the past two years possible and valuable by providing their experience, knowledge, and unique personalities.

Finally, this author would like to express a special acknowledgement to his parents, Monte and Christine Lloyd. Any success on the part of the author is due to their selfless support and strength. They both continue to inspire to this day.

Table of Contents

List of Figures

List of Tables

Abstract

Author: Marshall Everett Lloyd

Title: The Effect of System Lag on Unmanned Air System Internal Pilot Manual Landing Performance

Institution: Embry-Riddle Aeronautical University

Year: 2012

An important characteristic of UASs is lag because it can become a considerable challenge to successful human-in-the loop control. As such, UASs are designed and configured to minimize system lag, though this can increase acquisition and operation costs considerably. In an effort to cut costs, an organization may choose to accept greater risk and deploy a UAS with high system lag. Before this risk can be responsibly accepted, it must be quantified.

While many studies have examined system lag, very few have been able to quantify the risk that various levels of lag pose to an internally piloted, manually landed UAS. This study attempted to do so by evaluating pilot landing performance in a simulator with 0 ms, 240 ms, and 1000 ms of additional lag. Various measures were used, including a novel coding technique.

Results indicated that 1000 ms of lag was unsafe by all measures. They also indicate that 240 ms of lag degrades performance, but participants were able to successfully land the simulated aircraft. This study showed the utility of using several measures to evaluate the effect of lag on landing performance and it helped demonstrate that while 1000 ms poses a high risk, 240 ms of lag may be a much more manageable risk.

Future research suggested by this research includes: investigating lag between 240 ms and 1000 ms, introducing different weather phenomena, developing system lag training techniques for operators, and investigating the effect of aides such as predictive displays and autopilot-assisted recovery.

Introduction

Unmanned air systems (UASs) include air vehicles that have no on-board pilots; instead, they are remotely controlled by operators on the ground. The entire UAS includes the pilot, the ground control station, the air vehicle, and it can also include other components such as communications satellites. The use of UASs has greatly expanded since the 1990s. Well-funded militaries such as those of the United States and Israel have led the way, but there is a great deal of potential for organizations with greater fiscal constraints such as smaller militaries or commercial operators to expand the use of UASs even further (e.g., Frost & Sullivan, 2007). These users may be tempted to procure UASs with less than ideal configurations in order to cut operating costs. Quantifying the risks associated with non-ideal UAS configurations is an important step in UAS expansion and one to which this research will attempt to contribute.

UAS Background

There are a plethora of variations to UAS configurations. This research will not focus on the physical air vehicle variations, but instead on a specific landing system configuration. At the moment, there is a wide variety of possible UAS landing system configurations. Firstly, some UAS landing systems require manual control for landing while others use varying degrees of automation. Of the manually controlled systems, there are two major types: external pilot (EP) and internal pilot (IP). EP landing is how radio-controlled airplanes are landed. IP landing systems have the pilot positioned in a ground control station (GCS) and the pilot does not directly see the UAS air vehicle. Instead the IP is provided with a display of the UAS air vehicle's forward looking camera. The IP performs the actual task of landing using stick and rudder techniques. Due to the similarity of tasks between IP UAS and manned aircraft operation (e.g., McCarley & Wickens, 2005), the GCS is often set up much like a manned aircraft simulator in that there are stick and rudder controls with a display of what is ahead of the air vehicle, only the display image is the image output of the UAS's forward looking camera instead of a simulated environment. The IP landing system is the UAS configuration that this research will be investigating.

While some have assumed that UASs will be highly automated (e.g., Hopcroft, Burchat, & Vince, 2006; Van Erp & Van Breda, 1999), there will likely be a role for manually landed IP UASs in the future for several reasons. First, the most prevalent contemporary medium altitude long endurance (MALE) UAS is the General Atomics MQ Predator series, which uses manual IP landing. It stands to reason that as this series is replaced by another system, some will be sold off to other nations, or even decommissioned and sold to industry. While the actual systems are likely to stick around for some time to come, it is conceivable that the landing systems would be upgraded with some form of automation. This may be true, but there are advantages to manual control such as decreased response time, decreased mission planning time, and the potential for greater flexibility in IP situational awareness when compared against automatic control systems (e.g., de Vries 2005; Parush, 2006). Furthermore, the capability to manually land a UAS may be kept as a back up to an automated landing system for use in case of an emergency. In any of these applications, careful consideration needs to be given to the automation versus manual trade-offs.

There are many additional human factors challenges to UAS human controller performance when compared to manned aircraft. They include limited field of view (FOV), display layout, reduction of sensor modalities, poor resolution, and system latency (e.g., Van Erp & Van Breda, 1999; Williams, 2004). Some, such as reduction of FOV or display layout, will vary greatly from one UAS to another since there are many different UAS display configurations currently in use. Other challenges, such as system latency, also known as lag, are common to nearly all UASs.

Lag

It is well established that system lag affects human control behavior (e.g., Conklin, 1957; Ferrel, 1965; Ricard, 1994) in human in-the-loop (HITL) systems. A HITL system is one where a human operator has control capabilities and receives continuous feedback (e.g., van Paassen & Mulder, 2006). In a HITL system, an increase in system lag will decrease the ability of the human operator to control the system. This is true for both display and control lag (e.g., Jennings, Reid, Craig, & Kruk, 2004). Early investigations into the effect of lag in HITL were based on traditional engineering control theory, which has its roots in mechanical engineering (e.g., Franklin, Powell, & Emani-Naeini, 1994). *Figure 1* depicts a simple model of a closed loop control system upon which much of control theory is based (e.g., Franklin et al., 1994). It is a closed loop control system because the results of control input and disturbances are captured by output sensors and fed back into the system for further action. An open loop control system is one where there is no opportunity to modify control of the system based on sensor feedback. An example of an open loop system is an unguided rocket: once the rocket is fired, no adjustments can be made to the system.

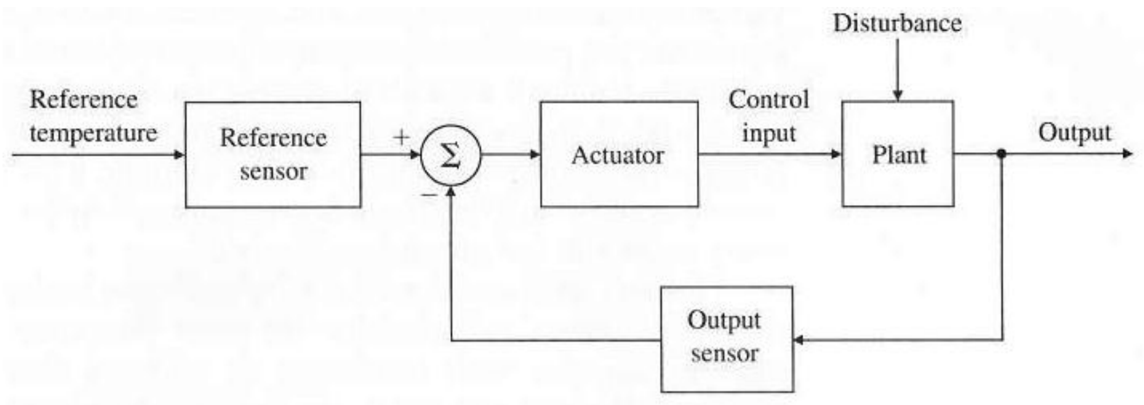

Figure 1. Closed-loop control system model. Adapted from Franklin et al. (1994).

Characteristics of closed loop control systems are increased accuracy when compared to open loop control systems, but with a tendency towards oscillation, or instability (e.g., DiStefano, Stubberud, & Williams, 1990). Instability can be best understood by first considering overshoot. Overshoot occurs when a controller overcompensates for a disturbance (e.g., Franklin et al., 1994; Whiteley, 1946). If a ship is heading north and needs to be heading east, a captain (i.e., the controller) will turn to starboard until the ship is heading east. If no other action is taken, once the ship is heading east, it will continue to turn and overshoot its desired heading. To correct for this overshoot, the captain turns to port in order to return to an easterly heading. If the desired heading is overshot again, then system is oscillating. If a system never stops oscillating while seeking a desired output, then it is considered unstable. There are many different definitions and criteria for control system stability (e.g., Franklin et al., 1994; DiStefano et al., 1990; Whiteley, 1946), but for the purposes of this discussion, the greater a system's oscillations and the longer the system oscillates in response to a disturbance, the less stable it is. In the example of the ship, the operator-induced oscillations are evidence of system instability.

The enormous accuracy advantage and the significant stability challenge of closed-loop control systems is why they are widely used and studied. Early investigations into lag incorporated the HITL controller into the closed-loop control systems model. The result was the development of the Crossover Model (COM; McRuer & Jex, 1967) and the Optimal Controller Model (OCM; Baron & Kleinman, 1969), both of which have seen wide spread use over the past forty years (e.g., Day, 1977; Mulder, van Passen, Flach, & Jagacinski, 2006; Wickens, 1992).

Crossover model. The Crossover Model (COM) was developed first (McRuer, & Jex, 1967) and has been widely used since. This model is built on the earlier quasi-linear models. Quasi-linear models were based on the observation that nonlinear aircraft-pilot systems could be approximated by equivalent linear systems (e.g., Jury & Pavlidis, 1963; Mulder et al., 2006). The system in question would still have a nonlinear remnant that could not be captured by its equivalent linear system, but it was negligible under certain conditions and did allow for fairly reliable system behavior prediction, provided those conditions were met (e.g., Mulder et al., 2006). The predictive power of quasi-linear models is limited by their need to be confined to environments that are similar to those in the experiments from which they were derived. This fact forces each quasi-linear model to be highly specialized and relevant to very specific conditions (McRuer & Jex, 1967).

As researchers strove to validate different quasi-linear models using empirical data (e.g., McRuer & Jex, 1967; McRuer, & Weir, 1969; Stapleford, McRuer, & Magdaleno, 1967), it was observed that HITL controllers would modify their behavior in response to the system. From these observations, and in an effort to overcome the specialization required for quasi-linear models, the COM was developed (McRuer & Jex, 1967).

At the heart of the COM is the HITL controller's systematic adjustment to the system. In order to describe the systematic adjustment, the human controller is treated as a second transfer function, *Hp(s),* in series with the machine controller, *Hc(s)*, as seen in Figure 2. U*(s)* is the original state of the system, *Y(s)* is the altered state of the system, and *Σ* represents the feedback of new information to the human controller (i.e., *U(s)* – *Y(s)* = change in system state observed by the human controller).

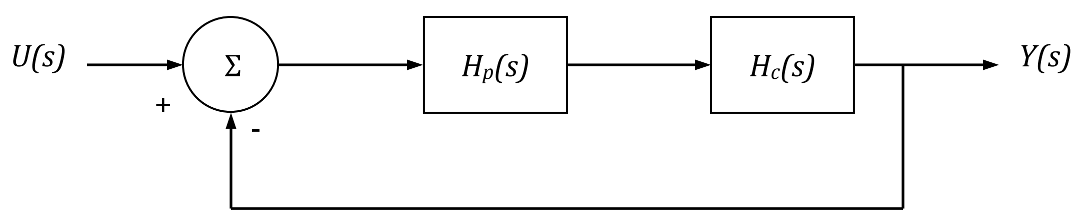

Figure 2. Crossover model transfer function.

The transfer function for Figure 2 is:

$$H(j\omega) = \frac{Y(j\omega)}{U(j\omega)} = \frac{H_p(j\omega)H_c(j\omega)}{1 + H_p(j\omega)H_c(j\omega)}. \quad (7)$$

When McRuer and Jex (1967) ran tests with different types of machine controllers, they found that the human operator would modify behavior to accommodate the different controllers. In other words, *Hp(s)* would adapt to changes in *Hc(s)*. The operators would adapt their gain and command frequency such that the amplitude ratio between input and output would approach 1. Additionally, the output frequency would approach the *crossover frequency*. What this means is that the operators would adopt a control input frequency and magnitude that would result in the output magnitude being the same as the input magnitude. This input-output amplitude ratio of 1 is what defines the crossover frequency. Furthermore, it was observed that, near the crossover frequency, the HITL controller showed similar characteristics across all types of machine controllers. That is to say the adaptation was systematic and universal. The common characteristic was the open-loop behavior of the system and is referred to as the crossover model:

$$Y_{OL}(j\omega) = H_p(j\omega)H_c(j\omega) \approx \frac{\omega_c}{j\omega}e^{-j\omega\tau_e} \tag{8}$$

where $Y_{OL}(j\omega)$ is the open-loop output, ω_c is the crossover frequency, and τ_e is the human controller's time delays (McRuer & Jex, 1967). Note that closed-loop feedback design often involves using the results of the open-loop output (Franklin et al., 1994), in other words, removing the feedback and analyzing system behavior. Knowing the machine controller's characteristics, the crossover frequency, and human time delays thus allows for an approximate prediction of the human controller's transfer function (McRuer & Jex, 1967). The above results were parameterized in a simple function used to describe pilot control behavior, which is known as the simplified precision model:

$$H_p(j\omega) = K_p\frac{1+\tau_L j\omega}{1+\tau_I j\omega}e^{-j\omega\tau_e} \tag{9}$$

where Kp is the pilot gain, τ_L is the lead time constant, and τ_I is the lag time constant. The two time constants are what the pilot modifies in order to adapt to the different machine controllers (Mulder et al., 2006). Equations (8) and (9) are the heart of the COM and they have been validated using empirical data in many experiments (e.g., Chan, Jhoun, & Childress, 1997; Feth, Groten, Peer, & Buss, 2009; Inaba & Matsuo, 1997; McRuer & Jex, 1967), though they lose accuracy as the conditions from the experiments are modified. For instance, the human controller can only accurately increase gain for small amounts of system lag, thus the COM does not hold true for systems with large amounts of lag (McRuer & Jex, 1967).

COM limitations. The COM was a significant advance in HITL control theory. It used empirical data to demonstrate that human controllers can be modeled for a simple tracking task and that the human controller adapted to various systems in a similar way, specifically by adapting the amplitude and frequency of controller inputs to the various types of machine controllers (Wickens, 1992). This allowed for additional predictive power that was not present in earlier quasi-linear models. The COM, however, cannot be considered a revolutionary step forward. The empirical data it was based on was collected from experienced pilots in a fixed base simulator while only using a single axis of control and display (McRuer & Jex, 1967). It is most relevant for simple, single-loop tracking tasks with minor disturbances. As these conditions are violated, the basic COM loses validity. Furthermore, the COM predicts changes in operator control input behavior, which suggests that operators will adapt to the systems they are attempting to control, but it does not provide insight into the success of operator input in controlling said system. Knowing that operators will adopt a certain behavior when an air vehicle is in level flight with minimal control lag and disturbances cannot help determine if an operator will be able to successfully control the air vehicle when asked to perform more complex maneuvers with significant lag and/or disturbances. In other words, COM is useful for describing behavior under docile conditions called for by its base assumptions, but is of limited use in determining success or failure in more extreme conditions.

Optimum control model. Like COM, the Optimum Control Model (OCM; Baron & Kleinman, 1969) was a further development of control theory to take into account a wider array of human limitations and feedback loops. While COM describes what a human operator is doing, OCM describes what the human operator should be doing (Mulder et al., 2006). The basic assumption is that operators who are properly trained and motivated will consistently behave optimally. Though it assumes optimal controller behavior, the model still takes into account the fact that the controller will be human. Incorporating time delays in perception and neuromuscular response, as well as perception noise and motor response noise helps capture these human limitations (Baron & Kleinman, 1969). Figure 3 is a visual representation of the model that has been adapted from Mulder et al. (2006).

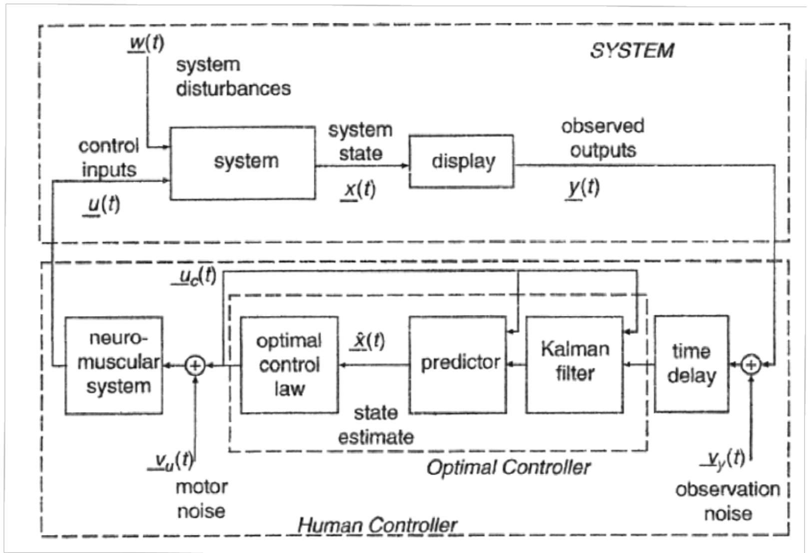

Figure 3. Optimal Control Model (adapted from Mulder et al., 2006).

Unlike COM, OCM remains in the time domain, with the system dynamics being described by the following equation:

$$\dot{x}(t) = Ax(t) + Bu(t) + Ew(t) \tag{10}$$

where the changing system state, $\dot{x}(t)$, is the sum of the control inputs, $x(t)$, outputs, $u(t)$, and disturbances, $w(t)$. Note that A, B, and E are constant matrices that describe characteristics of multiple loops in the system (Baron & Kleinman, 1969). It is assumed that the outputs of the system are depicted on a control panel with several displays, which can be described by Equation (11):

$$y(t) = Cx(t) + Du(t) + Hw(t) \tag{11}$$

with C, D, and H being constant matrices. The human aspect of the model is broken down into two parts: the optimal observer and the optimal regulator (Baron & Kleinman, 1969).

The optimal observer scans the appropriate displays for the relevant information, which is referred to as the Kalman filter. The Kalman filter can be thought of as the optimal observer's noise filtering process whereby the observer picks out useful information and filters out what is judged to be superfluous information. The output of the Kalman filter is the observer's understanding of the system's state (Mahoney, 1994). The understanding of the system's state allows the observer to predict system behavior and provide a signal for the controller, $y_{ob}(t)$. The output of the optimal observer is:

$$y_{ob}(t) = y_d(t-\tau) + v_y(t-\tau) \tag{12}$$

where τ is the perceptual time delay of the observer (Mulder et al., 2006). With the Kalman filter process, the observer will attempt to minimize the effect of display noise, $v_y(t)$. Note that there is no error introduced in the predictor process since it is assumed that the observer has complete understanding of the system's characteristics (Baron & Kleinman, 1969; Kleinman, Baron, & Levison, 1971).

The observer output, $y_{ob}(t)$, is then transformed into a control output, $u(t)$, by the optimal regulator's control signal, $u_c(t)$ (Mulder et al., 2006).

$$u(t) = u_c(t) + v_u(t) \tag{13}$$

with $v_u(t)$ being the motor noise, which can include neuromuscular noise and delay (Baron & Kleinman, 1969; Kleinman et al., 1971; Mulder et al., 2006).

In order to determine the optimal control behavior, OCM employs a *cost functional*. A cost functional is a mathematical expression of state and control that describes undesirable outcomes (Mulder et al., 2006). In other words, it places a penalty on unstable control components (Krstic, 2008). Solving for the OCM involves minimizing the cost functional, equation 14, of a system.

$$J(u) = E\left\{\sum_{i=1}^{m} q_i y_i^2 + \sum_{i=1}^{l} r_i u_{ci}^2 + \sum_{i=1}^{l} g_i \dot{u}_{ci}^2\right\} \quad (14)$$

where *Q, R*, and *G* are constant cost function matrix weightings related to their respective task functions, y_i, u_{ci}, and $\dot{u}_{ci}$ for the *m* number of inputs and *l* number of outputs (Baron & Kleinman, 1969; Mulder et al., 2006). Solving for the minimal cost functional will provide the OCM solution. The OCM solution minimizes the costs, which can also be thought of as minimizing variance or errors. This solution will indicate which of the feedback loops the optimal controller should focus on (Baron & Kleinman, 1969). For example, in the case of a pilot in a cockpit, the solution identifies the instruments that provide the best information for the optimal controller (i.e., the optimal pilot) to react to. It has been shown empirically that operators will in fact select the most effective feedback loop identified by the OCM solution (e.g., Dong, Hsiang, & Smith, 2009; Hess, 1981; Junker & Levison, 1977; Kleinman et al., 1971; Baron, 1973; Ganesh & Bajcsy, 2008). This way, the OCM solution not only identifies the optimal control feedback loops, but HITL controller behavior as well.

The OCM was of limited usefulness when it was first developed because it involved building and manipulating arrays of derivative equations, which can be very time consuming if done by hand. With the advance of low cost computational capabilities, it has become easier to employ in a variety of systems (e.g., Ganesh & Bajcsy, 2008, Dong et al., 2006). This has allowed for further validation with empirical data. Much of the empirical data suggest that under the right circumstances, the OCM assumptions of human operators behaving optimally are sound (e.g., Dong et al., 2009; Hess, 1981; Hsiang, Dong, & Karakostas, 2006; Junker & Levison, 1977; Kleinman et al., 1971; Baron, 1973; Ganesh & Bajcsy, 2008).

OCM limitations. Applying the OCM is challenging because it is highly parameterized. It requires that many of the system and disturbance characteristics (recall *A, B, D, D, E, H*, and *W*), task characteristics (*Q, R,* and *G*), and human characteristics (τ, v_y, and v_u) be known a priori. The more completely these characteristics are known, the more accurate the OCM solution becomes. Building an OCM for a particular system under particular conditions requires specific testing of the system in the condition of interest. This will establish the aforementioned characteristics, but it can be challenging. Another shortcoming of the OCM is the assumption that the HITL controller is able to construct a perfect mental model of the system. This is not always the case for complex and challenging systems, and the OCM is considerably less accurate at the extremes of performance, such as when the system has significant amounts of lag (Kleinman et al., 1971). Much like COM, OCM can be used to describe operator control input and suggests that operators modify their behavior to accommodate the system, but it does not provide insight into whether the operator will successfully control the system, especially when the systems are in less than ideal configurations.

Acceptable Lag

To reduce the lag-induced burden on in-the-loop human controllers, systems are designed with the goal of having no lag. While this is a worthy pursuit, it is virtually unobtainable. The reality is that all systems will have a certain amount of lag. For example, all aircraft have some, albeit small, magnitude of systemic lag. The

COM and OCM described above were able to demonstrate that human controllers can accommodate these small amounts of lag (Kleinman et al., 1971; McRuer & Jex, 1967). However, the larger amounts that can be expected in UAS operations (e.g., de Vries, 2005) violate the models' base assumptions.

Because zero system lag is difficult to attain, some amount of system lag must be acceptable, though too much results in loss of situational awareness, higher cognitive workload, and poor human controller performance (de Vries, 2005; Parush, 2006). As mentioned, much research has been done to examine the detrimental effect lag has on human performance (e.g., Conklin, 1957; Ferrel, 1965; Ricard, 1994). Early research focused on simple tasks such as tracking or inserting a peg into a space. Later research looked into the effect of lag in virtual environments. For instance, Watson, Walker, Ribarsky, and Spaulding (1998) looked at a simple task of grabbing and placing a virtual object, while Davis, Smyth and McDowell (2010) looked at a more complex task of maneuvering a virtual vehicle through a course.

Thus the challenge is to determine what level of lag is acceptable. There is no standard threshold level of lag. In some cases, it has been shown that lag as small as 16-80 ms has caused an effect on human controller performance (Ellis, Mania, Aldestein, & Hill, 2004; Patterson, Winterbottom, & Pierce, 2006), while in others, operator performance remained at an acceptable level with up to 700 ms of lag (Davis, Smyth, & McDowell, 2010). Some systems have been tested with levels of lag up to 3 s without a significant increase in operator error (Ferrel, 1965; Lane, Carignan, Sullivan, Akin, Hunt, & Cohen, 2002). What these articles show is that there are a plethora of factors that contribute to lag-induced human control degradation.

An important contributing factor to the effect of lag on HITL control performance is task complexity. The degree of freedom has been shown to affect simple manual tasks, while open and closed loop tasks have also been shown to contribute to degradation in varying degrees (Lane et al., 2002). Chang and So (1999) demonstrated that task difficulty was an important factor in determining the effect lag had on performance. As the difficulty of the task increased, the effect of lag would also increase. Watson et al. (1998) also observed this phenomenon when they researched varying levels of system responsiveness in a virtual environment. The time sensitivity of a task can also have an effect. Some tasks allow for the operator to "wait-and-see" the effects of each control input (e.g., Chang & So, 1999; Ferrel 1965; Lane et al., 2002; Watson et al., 1998), while other tasks that are more time sensitive and require constant inputs may not (e.g., Davis, et al. 2010; Ricard, 1995; Sullivan, Ware, & Plumlee, 2006).

Likewise, not all lag is equal. There has been much work done with different magnitudes of constant lag, but there is also an increasing interest in variable lag (e.g., Davis, Smyth, & McDowell 2010), which has been shown to degrade performance even further than constant lag. And while it is well established that increasing the magnitude of system lag decreases controller performance, this is not necessarily a linear relationship and it has been observed that there can be a threshold of lag magnitude beyond which performance drastically drops off (e.g., Lane et al., 2002).

For all the knowledge that has accumulated on system lag, there are very few generalizations that transcend all tasks and can be directly applied in a practical manner. The generalizations that have come out of the research are qualitative rules. For instance, the less variable the lag, the better the controller performance (e.g., Davis, Smyth, & McDowell 2010), or the smaller the magnitude of the lag the better (e.g., Conklin, 1957). These kinds of generalizations provide designers with goals, but not practical thresholds. Likewise, any quantitative results from the research to date have been task and environment specific. For instance, Lane et al. (2002) found that any more than 1.5 seconds of lag significantly reduced controller performance while piloting a zero buoyancy vehicle to perform a specific task under the conditions of the experiment. This is useful information for operating robots in zero gravity, but may not apply to the design of UASs.

Even when research on lag is confined to flight, there can be varying thresholds. System lag in an aircraft flight simulator was shown to produce significant flight path deviations at levels as low as 50 ms of delay (Ricard, 1995), while other research suggests that display or control lags below 144ms have no effect on pilot ratings of helicopter handling qualities (Jennings, Reid, Craig, & Kruk, 2004). This emphasizes the need to test specific configurations in order determine acceptable levels of lag. Furthermore, it emphasizes the importance of using multiple measures in order to determine the effects.

There has been some research into the effect of lag on UASs (e.g., Tadema, Theunissen, & Koeners, 2007; Thurling, 2000), however these have suffered from weak experimental designs. Thurling (2000) used only six pilot participants, had noticeable carryover effects between tests, and altered the experimental procedure midway through testing. This experiment also only tested aircraft in level flight while performing minor maneuvers. Tadema et al. (2007) also used only six pilot participants and they were helicopter pilots conducting fixed wing landings. Participants also flew only once or twice in each condition (number of flights per participant was not controlled) and no statistical analysis was conducted on the results.

Furthermore, the Cooper-Harper Rating (CHR; Cooper, & Harper, 1969) scores were used as part of the results, but they contradict the objective data in Tadema et al. (2007). The CHR is a subjective pilot evaluation of aircraft handling qualities. In the Tadema et al. (2007) results, they report that pilots rated UASs as “uncontrollable” using the CHR system when latency was 325-500 ms, but that the pilots never actually lost control of the aircraft and were in fact able to land safely every time. This suggests that the CHR measure of “uncontrollable” does not correspond to an uncontrollable aircraft and may not properly convey the risk of operating such an aircraft. Other measures used by Tadema et al. (2007) were air vehicle position at landing threshold, root mean square (RMS) operator control stick deflections, and RMS error of the air vehicle position. These three measures seem to provide better insight into operator performance, though it is difficult to determine from Tadema et al. (2007) because no statistical analyzes were conducted.

Contextual Control Model. As mentioned above, the COM and OCM suggest that HITL controllers will modify their behavior to accommodate system lag, but these models cannot predict a lag threshold and lose validity with larger amounts of lag. In order to determine lag thresholds of a specific HITL system, one must test the system. However, interpretation of the results can be challenging and even contradictory, as was observed by Tadema et al. (2007) when pilots rated UASs with lag as uncontrollable yet never lost control.

For the evaluation and interpretation of HITL UAS performance, it is useful to view the human and machine as a joint cognitive system (JCS); this departs from the strategy of deconstructing the system into a collection of sub processes with distinct inputs and outputs (Hollnagel & Woods, 1983). This way, the quality of performance of the entire joint system is evaluated instead of the individual sub processes in isolation (Hollnagel, 1997; Klein et al., 2003). Both COM and OCM equate the human operator to an electro-mechanical sub component and in so doing, narrow the scope of analysis, as is done when the COM simply describes operator crossover frequency (McRuer & Jex, 1967). The JCS approach of acknowledging sub processes while keeping a broad view of the performance of the entire system (Hollnagel, 1997; Klein et al., 2003) is useful when evaluating the operational performance of a whole system.

The Contextual Control Model (COCOM) describes JCS performance as a mix of feedback and feedforward control actions (Hollnagel & Woods, 2005). JCS control is dictated by the dynamic interplay between situation understanding, actions, and information feedback. Control actions (reactive or proactive) occur based on the understanding of the situation and on feedback, which are in turn affected by the actions taken (Hollnagel & Woods, 2005). An important aspect of COCOM is the description of control modes. While degrees of control vary along a continuum

(Worm, 1999), the Contextual Control Model (COCOM) identifies four different modes of control: *scrambled*, *opportunistic*, *tactical*, and *strategic* (Fujita & Hollnagel, 2004; Hollnagel & Woods, 2005). Control modes are described as:

- *scrambled*, which consists of random trial and error with little to no planning or thinking;
- *opportunistic*, which involves only limited planning due to lack of understanding, limited time, and/or poor information feedback, which results in inefficient actions and wasted attempts;
- *tactical*, which has useful feedback and takes some delayed effects into account with actions following known rules, though the selection of actions can still be ad hoc; and
- *strategic*, which has high quality information feedback that facilitates consideration of high level goals and understanding of dependencies between actions and multiple goals (Hollnagel & Woods, 2005)

Hollnagel and Woods' (2005) description of the control mode characteristics is outlined in Table 1.

Table 1

Control Modes. Adapted from Hollnagel & Woods (2005)

Control Mode	Number of goals	Subjectively available time	Evaluation of outcome	Selection of action
Strategic	Several	Abundant	Elaborate	Based on models/predictions
Tactical	Several (limited)	Adequate	Detailed	Based on plans/experience
Opportunistic	One or two (competing)	Just adequate	Concrete	Based on habits/association
Scrambled	One	Inadequate	Rudimentary	Random

The aforementioned control modes provide guidelines that help classify JCS performance along the continuum of control. These control modes have in fact been used in Human Reliability Analysis (HRA) in order to estimate the probability of failure (e.g., Fujita & Hollnagel, 2005). Knowing the probability of failure provides insight into the level of risk to the JCS associated with each control mode.

If a JCS is operating in a certain control mode during normal conditions, it is reasonable to assume that altering the condition in such a way that they become novel and challenging will affect the control mode. The operator will have a poorer understanding of the system's behavior in the new conditions, which will negatively affect the selection of actions. A poorer selection of actions will result in less desirable outcomes and will challenge the operator's ability to manage multiple goals. As the operator manages goals one at a time based on simple associations (or worse, through random action), the subjective amount of time is reduced to just adequate or insufficient. The result of challenging abnormal conditions is that the control mode is degraded.

Applying COCOM. There has been little to no research into the relationship between system lag and COCOM control mode, so it is difficult to determine the exact effects. It is reasonable to expect that an experienced pilot will exhibit tactical or strategic control when flying an aircraft without any additional lag. The pilot will choose his/her actions based on their mental model of how the JCS works and on his/her experience. This will place the event horizon well into the future. They will be able to manage many goals (e.g., altitude, airspeed, heading, etc.) and will have either an adequate or abundant amount of time in order to make decisions. As increasing amounts of lag are introduced, the pilot will be required to readjust his/her model of the JCS and will have less experience to draw upon when selecting a course of action. Actions will instead be chosen based on simple associations (e.g., aircraft is too high, point nose down immediately) and the pilot will be able to manage fewer and fewer goals simultaneously. With less available time to choose actions, the event horizon will come closer and closer to the present. It is reasonable to expect that increasing levels of lag will result in lower control modes.

Safe versus unsafe flight. The definitions for tactical and strategic control modes can be used to describe an air vehicle system with lag that is being safely controlled. The operator demonstrates an understanding of the system and can ably compensate for lag. In contrast to tactical and strategic control, the definitions for scrambled and opportunistic control modes describe an air vehicle that is being controlled in a manner that is unsafe. The operator does not have an understanding of the system, is only reacting to immediate concerns, and at best, operating the system extremely inefficiently. This classification of safe control and unsafe control modes will assist with data analysis and interpretation in this study.

Research Objectives

The current state of knowledge leaves those who are interested in the effects of lag on a specific HITL JCS in a situation whereby specific tasks need to be investigated under specific lag conditions in order to better understand the actual risks involved. The goal of this research is to investigate a specific task in a human controlled system: the manual landing of a UAS by an IP. The landing task was chosen because it has historically been the phase of manually controlled UAS flight to be most prone to operator error (e.g., Parush 2006; Williams, 2004), thus making it the phase that is most likely to be affected by lag with catastrophic results. The lags to be investigated are the amount that could be the result of signal transmission to and from satellites in beyond line of sight (BLOS) operations.

Lag in BLOS UAS operations. The United States Air Force (USAF) and Army control overseas UAS missions with operators located on the US mainland. Control signals are transmitted BLOS to UASs via geostationary satellites with a constant presence over a specified area of the earth (e.g., de Vries, 2005). This results in an increased system lag during the mission. In order to eliminate this component of system lag during the high-risk phases of flight, control of a UAS is passed over to operators that are positioned at overseas air bases. These operators then control the take offs and landings in a line of sight configuration (LOS). The LOS configuration still experiences lag, but not nearly as much as the BLOS configuration. If the US-based operators were to attempt to land on an airfield that is overseas, or anywhere BLOS, the fixed amount of time that it takes for the signals to travel to and from the satellite would greatly increase system lag.

The USAF and Army have the resources to reduce the system lag risk in BLOS operations by deploying additional pilots to control UASs in LOS configurations when necessary, but not all organizations that might use UASs will have the resources to support additional pilots. An organization may choose to accept the increased risk due to system lag associated with BLOS operations in order to save resources (i.e., one pilot in one location as opposed to two pilots in two locations). It is also conceivable that a UAS configured for LOS landing may be forced to land BLOS due to technical difficulties or weather. In this case, the UAS operator may want to attempt a landing, especially if the alternative is to safely ditch the aircraft, but the airport manager may not want to accept the risk of a failed UAS landing. Without the value of a pilot's life on board, the airport manager may be less compelled to accept the risk of a UAS crash shutting down operations. A third example would be a UAS with a malfunctioning automatic landing system that is attempting to land BLOS. Much like the second example, there are risks involved with attempting a BLOS manual landing. Before these kinds of risk can be responsibly accepted, they must be quantified. Assessing pilot performance of LOS and BLOS manual landings is an important step in quantifying this risk. Pilot performance could be evaluated by not only recording glide slope deviations, but also by examining the JCS control mode during the decent.

Based on a geostationary communications satellite at an altitude of 3.58×10^7m and the speed of light (3.00×10^8 m/s), it takes a signal 239ms to travel up to a satellite and back to earth. This is approximately the amount of fixed lag that is introduced to a UAS when it operates just BLOS (the additional distance the signal travels to get to the UAS as opposed to back to the GCS is negligible), and will likely be noticeable to operators (e.g., Chen, Haas, & Barnes, 2007; Moss, Muth, Tyrrell, & Stephens, 2010). Total system lag is the sum of the BLOS lag, hardware, and software lag. Investigating 239ms of lag is assuming the best-case scenario, i.e., technological advances allow for instantaneous computation. A more realistic scenario is one where various aspects of the datalink, such as encryption, compression, synchronization, and weather distortions, contribute additional delay. De Vries (2005) estimates a realistic datalink delay to be 1000 ms, though in certain configurations it can be much higher. Note that BLOS lag affects both control and display signals.

Research approach. In an attempt to quantify the increased risk of conducting IP manually controlled UAS BLOS landings, participant conducted approaches in LOS and BLOS conditions while their performance was assessed. These conditions were no lag, 240 ms of lag, and 1000 ms of lag; these levels will be referred to as the *no lag*, *low lag*, and *high lag* conditions respectively. Trials for all lag conditions were varied across wind conditions, which consisted of *ideal wind* conditions and an *adverse wind* condition (a crosswind gust). The wind manipulation was introduced in order to better simulate potential adverse environmental conditions that occur in real operations and it provided participants with the opportunity to overcome a system disturbance.

Measuring performance. Several techniques were used to measure participant performance. In a similar study, Tadema et al. (2007) used aircraft deviations from the target glideslope expressed as root means square error (RMSE). This study uses a similar measure, which will be referred to as *RMSE*. Another measure used by Tadema et al. (2007) was aircraft position at the landing threshold: if the position was within certain criterion, then the attempt was considered a landing success. This uses a similar measure, referred to as *landing success*, though with additional criterion.

In addition to RMSE and landing success, two more measures were used: *landing control* and *entire approach control*. Two coders independently evaluated each approach in order to determine which control mode best described the JCS's performance during the final 2000 ft of the approach (i.e., *landing control*), and during the entire approach (i.e., *entire approach control*). A detailed description of entire approach control, and landing control can be found in Appendix A. Recall that tactical and strategic control modes are considered safe control modes. Once entire approach control coding was complete, the distances of flight path segments associated with a safe control mode (AWSCM) were summed. This gave each approach a *distance AWSCM* score.

Hypotheses

RMSE. It is hypothesized that when lag is introduced to the system, it will increase RMSE when compared to no lag for both the ideal and adverse wind conditions. Between the two levels of lag, it is hypothesized that the larger magnitude of lag (1000 ms) will result in greater RMSE than the smaller magnitude of lag (240 ms). It is hypothesized that introducing the adverse wind condition will significantly increase RMSE in all lag conditions.

Landing success. It is hypothesized that when lag is introduced to the system, it will decrease the landing success rate when compared to no lag for both the ideal and adverse wind conditions. Between the two levels of lag, it is hypothesized that the larger magnitude of lag (1000 ms) will result in fewer landing successes than the smaller magnitude of lag (240 ms). It is hypothesized that introducing the adverse wind condition will decrease the landing success rate in all lag conditions.

Landing control. It is hypothesized that when lag is introduced to the system, it will shift pilot control to a lower control mode for both the ideal and adverse wind conditions. Between the two levels of lag, it is hypothesized that the larger magnitude of lag (1000 ms) will more often result in a lower landing control mode code than the smaller magnitude of lag (240 ms). It is hypothesized that introducing the adverse wind will shift pilot control to a lower control mode in all lag conditions.

Entire approach control. It is hypothesized that when lag is introduced to the system, it will decrease the distance associated with a safe control mode (AWSCM) when compared to no lag for both the ideal and adverse wind conditions. Between the two levels of lag, it is hypothesized that the larger magnitude of lag (1000 ms) will result in less distance AWSCM when compared to the smaller magnitude of lag (240 ms). It is hypothesized that introducing the adverse wind condition will decrease the distance AWSCM in all lag conditions.

Methods

Participants

This research involved 13 participants. Excluding two outliers over 50 years old, the average age was 23.73 years (*SD* = 3.93). The participants were a random pool of volunteer general aviation pilots who, excluding two outliers with over 10,000 hrs, had an average 699.00 hrs of flying experience (*SD* = 626.54). Participants were recruited using flyers distributed to Embry-Riddle instructor pilots and posted around the Embry-Riddle campus. There have been instances when lag has been shown to trigger simulator sickness (e.g., Casali & Wierwille, 1980; Draper, Viire, Furness, & Gawron, 2001; Patterson et al., 2006), so for safety, participants completed a simulator sickness questionnaire (SSQ; Kennedy, Lane, Berbaum, & Lilienthal, 1993) shortly after their first exposure to the lag condition. If they exhibited signs of simulator sickness, they were withdrawn from the study.

Participants took part in three sessions, with each session lasting 30 - 45 minutes. Sessions were separated by 1-7 days. Participants were given $10 at the start of each session as compensation for their time. The best performance was awarded $100 and the second best was awarded $20. The two participants with the best performance was the one who had one of the five lowest total RMSE scores and top two greatest number of successful landings. Participants signed an informed consent form (see Appendix B) and were allowed to end their involvement in the experiment at any time. The informed consent form notified the participants about the nature of the study, the risks, their rights, and the prize money.

Materials

Apparatus setup. The flight simulation software was X-Plane9 (Craighead, Murphy, Burke, & Goldiez 2007) with Saitek yoke, rudder, flap, and throttle controls. As seen in Figure 4, the participant's station was to the left of the experimenter. The simulated aircraft was a Cessna 172N because it is a common aircraft that participants were familiar with through their flight training. A simulated UAS such as the MQ-1 would have been ideal, but there was no readily available pool of MQ-1 pilots to draw from for this experiment. The assumption is that participants in this study will be as familiar with the Cessna 172N as a UAS pilot would be with a UAS. Flight data were recorded at 4 Hz using an X-Plane plugin that was developed for this experiment. Network impairment was simulated using NetDisturb (ZTI, 2012).

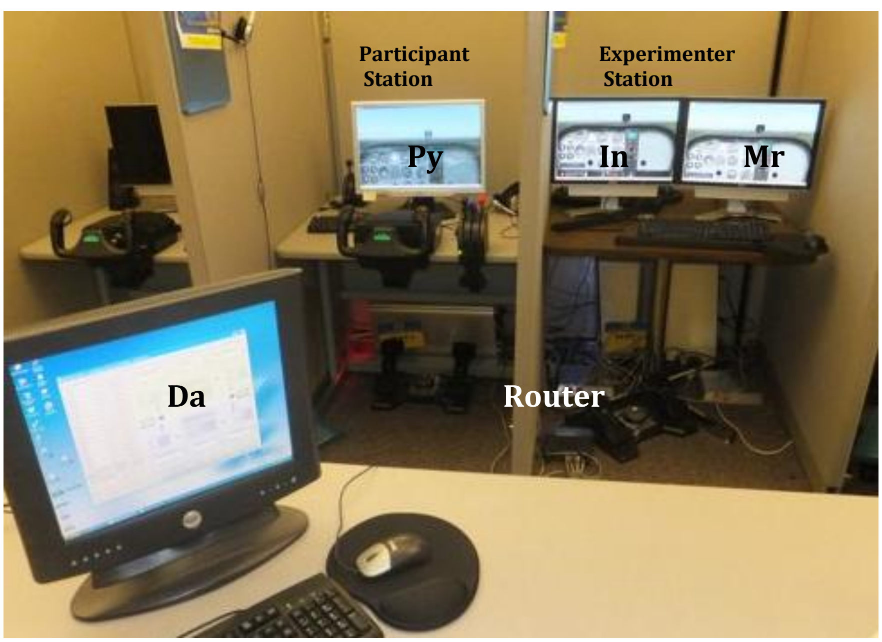

Figure 4. Study apparatus setup.

Hardware. The apparatus consisted of a network of four computers connected via a router. Table 2 lists the equipment used. All computers had 1 Gb/s network interface cards and 60 Hz monitors. Three computers were used to simulate components of the UAS network and are referred to as *Master* (*Mr*), *Intermediary* (*Iy*), and *Participant* (*Pt*). The Mr and Pt computers were connected directly to the router. The fourth computer, referred to as *Delay* (Da), was used to run NetDisturb, the network simulation software. The Da computer had two network interface cards, one that was connected to the router and the second that was connected to Iy. In this way, Iy was connected to router via Da, which ran NetDisturb and could thus impede the connection.

Table 2

List of Apparatus Equipment

Component	Hardware
Mr	Dell XPS630
Iy	Dell XPS710
Pt	Dell XPS710
Da	Dell Precision
Router	Linksys WRT54GL
Yoke	Saitek Proflight Yoke System
Rudder	Saitek Proflight Rudder Pedals
Throttle	Saitek Throttle Quadrant
Flaps	Saitek Throttle Quadrant

Network relations. Figure 5 shows the network configuration alongside a BLOS UAS network. The X-Plane simulation was run on Mr and the cockpit display was exported to Iy, which in turn was exported to Pt. The participant used aircraft controls that were connected to Mr and controlled the simulation on Mr; however, the participants were not able to see the Mr monitor. Instead, they used the Pt monitor, which was displaying data that had been exported from Mr to Iy (via Da), and then from Iy to Pt (via Da as well). The Mr monitor displayed what would happen to the air vehicle if there was no control latency, while the Iy monitor displayed what would happen if there was only control latency. The Pt monitor displayed the result of both display and control latency. The simulation scenarios were loaded, controlled, and supervised from Mr.

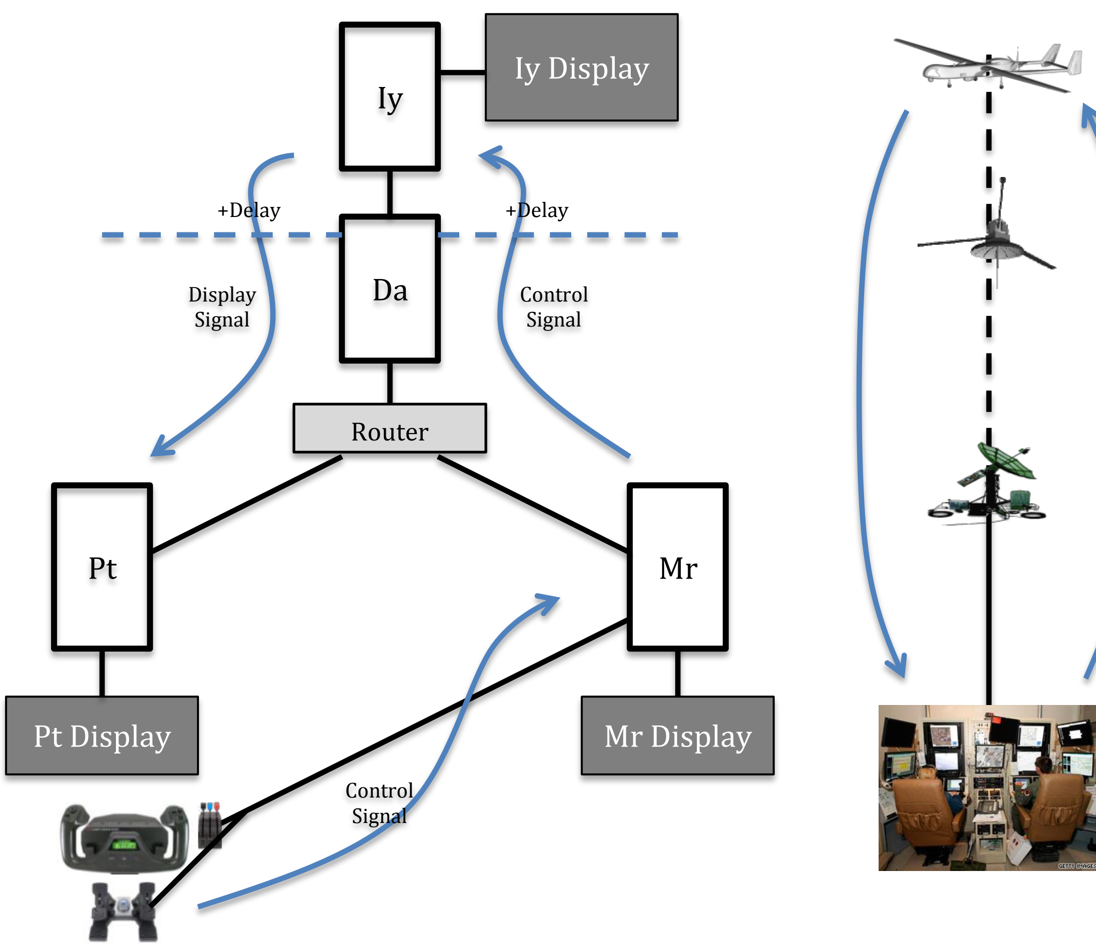

Figure 5. Apparatus network connections (left) compared to the intended simulated BLOS UAS network (right). Note that Da is positioned to be able to impair the network in the same way that a satellite connection does.

Plugin. An X-Plane plugin was developed specifically for this experiment. The plugin and associated files are included as supplementary material and are the intellectual property of Matthew Grasso. The plugin set the experimental conditions, allowed for easy transition between trials, and recorded flight data into an excel file. Detailed instructions on plugin use can be found in Appendix C.

Apparatus Testing and Validation

Determining PAPI slopes. Participants used the precision approach path indicator (PAPI) to guide their approaches and the vertical component of RMSE was measured as deviations from the PAPI glideslopes. The PAPI is designed to guide the pilot along an ideal 3° glideslope and it does so by indicating when the aircraft is too high or too low. In this way it in fact indicates a range of acceptable glideslopes. As such, it was necessary to determine the PAPI upper and lower glideslope boundaries. This was done by positioning the aircraft at three different distances from the touchdown point, 2,438 ft, 11,064 ft, and 19,501 ft as shown in Figure 6, and then changing the aircraft's altitude until the upper and lower ideal PAPI boundaries were observed. Figure 7 is a screen shot of the aircraft below a lower boundary. The points at which the PAPI indication changed from ideal to non-ideal were then recorded, as shown in Figure 8, and used to calculate slope with respect to the PAPI position. The slopes between points were calculated as well.

Table 3 shows a summary of all the PAPI angles that were observed. Note that the altitude reported in this table is the vertical distance between the aircraft and the PAPI, not the aircraft's altitude above sea level.

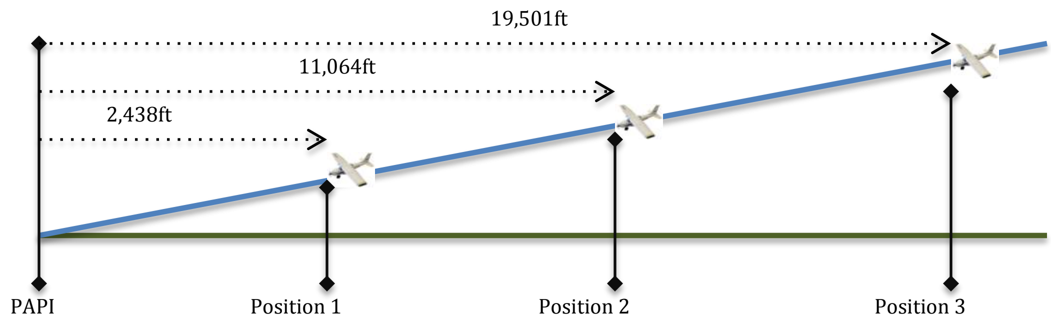

Figure 6. PAPI upper and lower thresholds were observed from three different positions.

Figure 7. A screenshot of the aircraft below the lower PAPI boundary, as indicated by one white and three red lights.

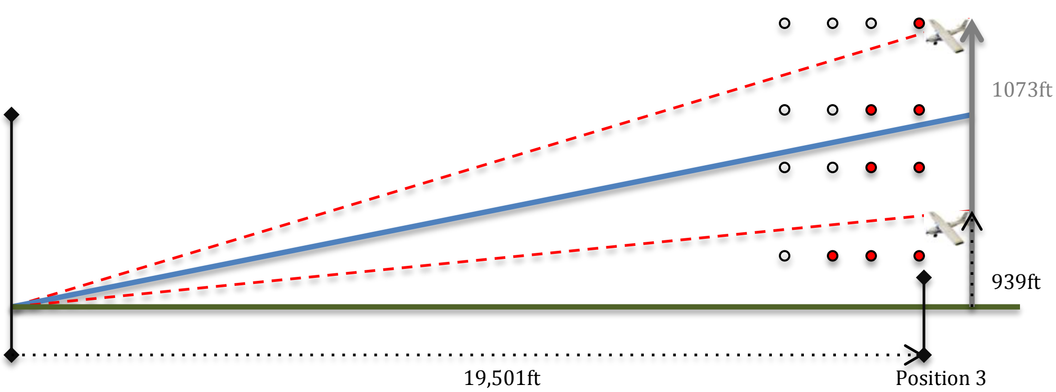

Figure 8. An example of PAPI indication changes as thresholds are crossed at different altitudes.

Table 3

PAPI Glideslope Boundary Observations

		Aircraft Position Comparisons				
Boundary	Measure	PAPI-1	1 - 2	PAPI-2	2 – 3*	PAPI-3*
Upper	Alt. (ft)	136	485	621	452	1,073
	Dist. (ft)	2,438	8,626	11,064	7987	19,051
	Slope	3.20°	3.22°	3.22°	3.24°	3.23°
Lower	Alt. (ft)	118	425	543	396	939
	Dist. (ft)	2,438	8,626	11,064	7,987	19,051
	Slope	2.78°	2.82°	2.81°	2.84°	2.83°

**The experimental procedure has the aircraft start at 11,064 ft from the runway and as such, the observed slopes from position 3 were not considered, though notice that they are similar to the other observed slopes.*

There are inconsistencies in the observed slope. This is partially due to the fact that X-Plane allows the aircraft's altitude to be modified by one foot at a time. As such, it was only possible to observe the PAPI threshold to the nearest foot of altitude. This can make a noticeable difference when dealing with small angles. For example, if the altitude difference for "PAPI – 1" is one foot lower, then the slope becomes 3.18°. The imprecision of the altitude manipulations led to inaccurate and inconsistent slope observations. To accommodate the imprecisions, the strictest observed slope less 0.1° was selected as the PAPI boundary slope. For example, of the upper slopes, 3.20° is the most stringent, but adding 0.1° makes it more conservative. Likewise, 0.1° was subtracted from the most stringent of the lower bounds. The result was that 3.21° and 2.81° were used for the upper and lower PAPI boundaries, respectively.

Testing the plugin. The plugin allows for upper and lower glideslope limits to be set. It then calculates the vertical deviation of the aircraft with respect to the relevant boundary. In order to validate the plugin, the aircraft's instruments were read while it was moved around to various points. These readings were then compared to the Excel file generated by the plugin. The average difference between the two measurements was M = 1.21 ft (SD = 0.33) for 430 data points.

The plugin also sets the runway centerline as the default ideal horizontal alignment. It then determines the horizontal deviation of the aircraft with respect to the runway centerline. This function was tested by comparing aircraft GPS coordinates with the plugin's horizontal deviation output. The perpendicular distance between the runway centerline axis and the aircraft's GPS coordinates was calculated and compared to the horizontal deviation recorded by the plugin. The difference between the two measures was calculated for 729 data points and was M = 3.17 ft (SD = 1.78), which was considered satisfactory. The plugin also calculated the total deviation by summing the horizontal and vertical vectors, as shown in Figure 9. This calculated result was verified by computing the hypotenuse of the horizontal and vertical deviations.

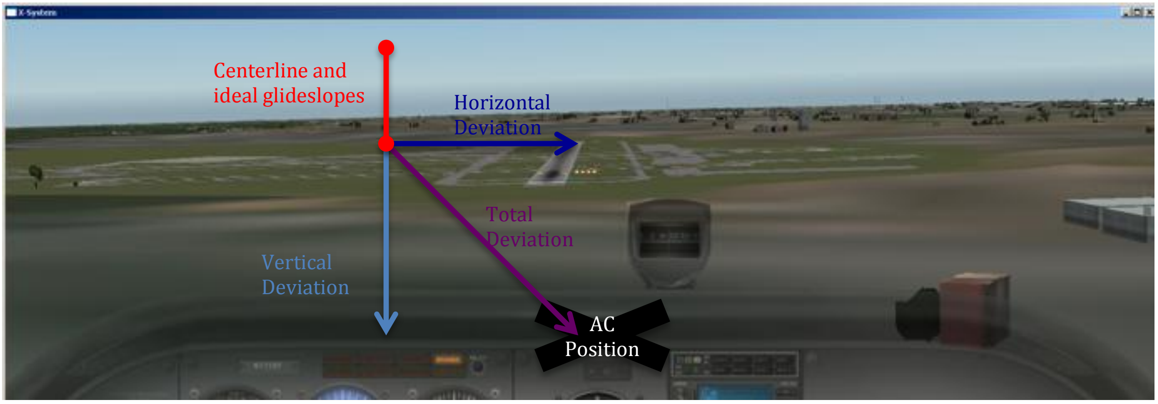

Figure 9. A screenshot of the aircraft deviating from the ideal glideslope.

Testing delay. Lag was introduced using the NetDisturb software. The accuracy of NetDisturb was tested using the Microsoft Windows built-in ping utility. The ping utility was used to run a ping test, which is a common tool used to test network communication success and latency. The ping test was run several times at all three lag settings and always reported latency within 1 ms of the NetDisturb setting.

The system's baseline lag was tested using direct observation to further validate the apparatus. A high-speed video camera was used to test and measure baseline system delay with NetDisturb set at zero delay. It was a Fujifilm FinePix F505 digital camera capable of recording video at 320 fps, which is equivalent to 3.125 ms per frame. The camera would be used to record the apparatus, and then Quicktime would be used to advance the recording frame by frame to observe changes over fixed time frames. This procedure is explained in further detail below.

To validate the test method, the video camera recorded a stopwatch and then Quicktime was used to playback the video frame by frame. By advancing the recording stopwatch recording by one frame, it was possible to observe the stopwatch advance by approximately 3 ms. This was approximate because the stopwatch screen latency did not permit accurate observations of less than 1 ms.

With evidence to support the testing method, the baseline lag of the apparatus was tested. The test involved recording the three monitors while maneuvers were performed. The recording was then replayed frame by frame in Quicktime, during which the experimenter looked for marker events. A marker event is a distinct event that occurs during the flight and is displayed on the monitor. In this case, the marker event was a change in the aircraft's orientation such that the pedestal mounted compass moved above or below the horizon. Notice that the compass is above the horizon in Figure 10, and then moves below the horizon in *Figure 11*.

Figure 10. The compass is above the horizon in this frame.

Figure 11. The attitude of the aircraft has change such that the compass is no longer visible again the blue sky.

When a marker event was observed on the Master monitor (see the right side in Figure 10 and Figure 11), the recording was then advanced frame by frame until the event occurred on the Participant monitor (see the left side in Figure 10 and Figure 11). The number of frames between events was recorded and converted into time in order to determine the baseline lag during the marker event. Figure 12 shows a series of screen captures as the recording was advanced by two frames. Notice as the outline of the pedestal mounted compass breaks horizon. In this case, the delay between the primary monitor and participant monitor was observed to be at most one frame, or 3 ms.

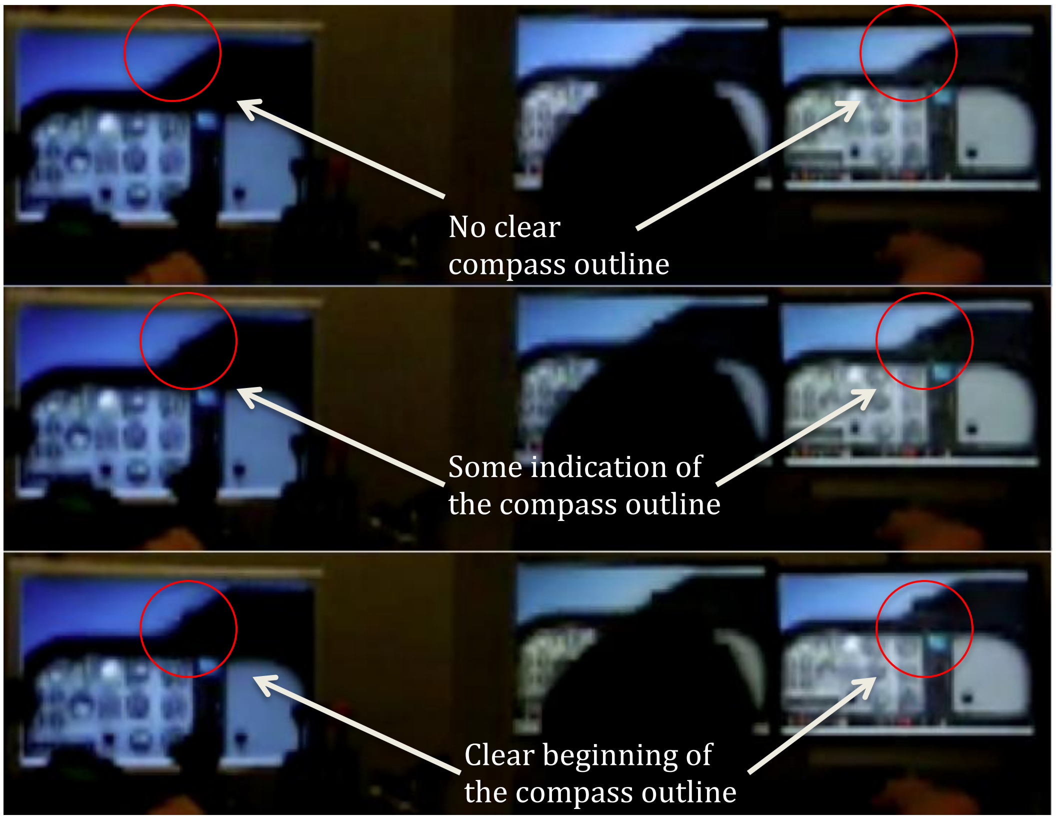

Figure 12. The recording is being advanced frame by frame to observe a marker event.

Figure 12 does not show a delay, however there were instances when marker events were separated by as many as 12 frames. Figure 13 shows one such event. In this case, the aircraft was in a shallow dive and the compass disappeared on the primary monitor 12 frames (36 ms) prior to disappearing on the participant's monitor.

Figure 13. Observing delay between marker events.

Twelve marker events were observed and recorded. The low resolution of the digital recording meant that at times it was difficult to determine exactly when the compass disappeared or reappeared above the horizon. When it was unclear, the larger frame count was used. When it appeared that no frames separated marker events, it was recorded as one frame because zero lag is highly unlikely. Thus the values in Table 4 are conservative and represent the largest observed lag.

Table 4

Observed Baseline System Lag

	Marker Event Separation	
Event #	Number of Frames	Time (ms)
1	1	3
2	1	3
3	13	41
4	4	13
5	2	6
6	1	3
7	9	28
8	1	3
9	12	38
10	1	3
11	1	3
12	7	22
M(*SD*)	4(5)	14(15)
Median(*MAD*)*	1(0.5)	3(1.5)
*Mode**	1	3

**Median and mode are reported since the distribution does not* *appear to be normal.*

Sources of this baseline lag include internal computing processes required to run windows and X-Plane, network communications, and monitor refresh rates. While more observations would be needed for a more complete understanding of the baseline lag distribution, it seems to be skewed. It appears that the majority of the time, the baseline system lag is near zero, but that it will periodically spike to higher levels. There are many components of the apparatus that operate on frequencies. For example, X-Plane transmits display data at 60 Hz, the network cards transmit 1 Gb/s, the router transmits 100 Mb/s, and the monitors refresh at 60 Hz. The periodic spikes in lag could be the result of the various frequencies lining up or data bottlenecks occurring at the router.

Based on the observations made during the lag test, the baseline system lag is negligible. Although lag may momentarily spike, these spikes are much lower than the lag manipulations (240 ms and 1000 ms). Furthermore, the observed spikes are at levels that are similar to or lower than baseline lags that have been reported in other comparable studies. Thurling (2000) reported a UAS baseline lag of 372 ms and Tadema et. al (2007) reported a baseline lag of 50 ms. Other studies of lag in aviation have shown that 50 ms (Ricard, 1995) and 144 ms (Jennings et al., 2004) had no effect on pilot performance. Apparatus testing detected some baseline system lag, though it was observed at levels that are considered acceptably low.

Dependent Variables

RMSE. Deviations from the target PAPI glideslope range and the runway centerline were recorded by the X-Plane plugin. The hypotenuse of the lateral and vertical deviation was the aircraft glideslope deviation at any given moment. This deviation was recorded four times every second. Each deviation was squared, and then the mean of the squared deviations for a single approach was calculated. Finally, the square root of the mean squares was taken in order to give the RMSE. The RMSE was used to measure performance because it is more sensitive than the mean error and it gives greater weight to large deviations.

Landing success. Two hypothetical landing gates were established at 1000 ft and 500 ft before the target touchdown location. Both gates were centered on the 3° glideslope. The first gate was 100 ft tall and wide, while the second gate was 55 ft high and 80 ft wide. Any approach that passed outside of either gate was considered a failed approach. Figure 13 shows an example of the landing gates. Note that this only shows the vertical flight profile (i.e., altitude), but the horizontal flight profile (i.e., alignment with the runway centerline) was also considered. A final criterion for landing success was landing speed. If airspeed at the 1000 ft gate was below 55KIAS or above 70KIAS (Cessna Aircraft Company, 1977), it was considered a failed approach. All of the above criteria need to be satisfied in order for the landing to be considered a success.

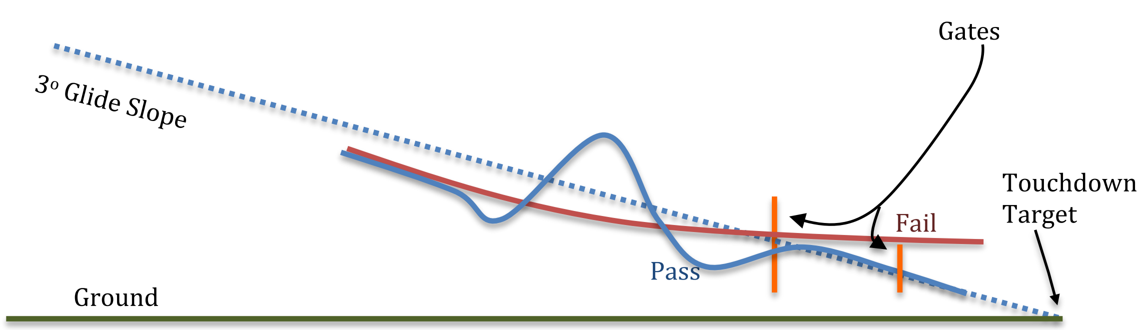

Figure 14. Landing success is determined at gates.

Control mode coding. Two coders independently evaluated each approach in order to determine which control mode best described the JCS's performance. The same two coders were used for all approaches. Coder training was conducted according to the coder training plan, which can be found in Appendix A. It consisted of an explanation of general JCS control mode theory, how it pertains to this experiment, and how coding was to be conducted for this experiment. Coders practiced coding on a small set of sample data, which was used to discuss and address differences in coding between the coders. The coder training plan includes detailed descriptions of the various control modes and can be found in Appendix A.

Coders were provided with graphs that superimposed actual aircraft flight paths over the target flight path. Recall that the target flight path is the glideslope as indicated by the PAPI, which has an upper and lower boundary, as well as horizontal alignment with the runway centerline. The coders evaluated the actual flight paths based on the definitions in the coder training plan and assigned control mode codes to the approaches. The training plan describes how aircraft position, trajectory, and trajectory changes are evidence used to determine a control mode. Furthermore, the training plan provides detailed guidelines for the coding activity, which helps maintain reliability and reduces subjectivity of the coding data.

Landing control. The landing control measure was a single control mode for each trial. Coders assigned a single code to the last 2000 ft of each approach. Coders were only allowed to select a single control mode code that best described the behavior exhibited by the JCS. Figure 15 shows an example of a coder's evaluation of the final 2000 ft of an approach, which was marked as opportunistic. Notice the considerable overshoot of the inefficient flight path, which is consistent with the description of opportunistic control in Appendix A.

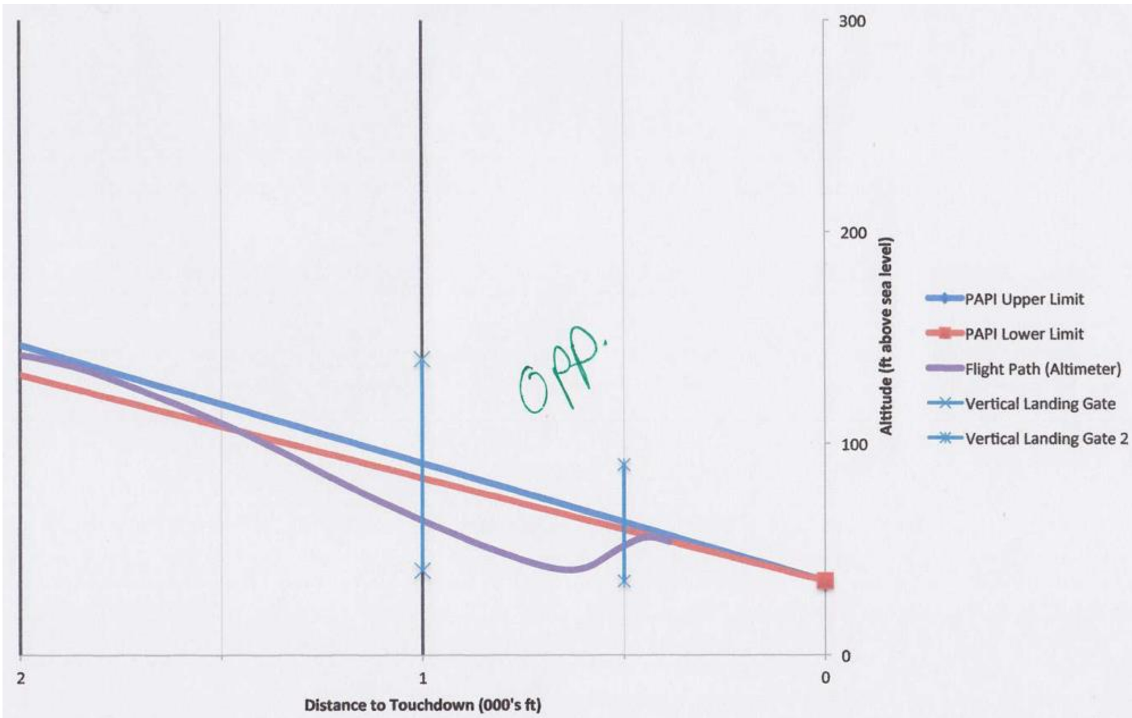

Figure 15. An example of the final 2000 ft of a flight path graph with a coder's hand written evaluation of landing control mode.

Entire approach control. This exercise was repeated with graphs of the aircraft flight path during the entire approach. Coders assigned codes to the entire approach. As shown in Figure 16, there can be evidence of multiple control modes over the course of an approach. Coders assigned control modes to sections of the approach as they saw fit, with each section spanning a certain distance of the approach. The entire approach control measure was the sum of distances associated with a safe control mode (AWSCM). Figure 16 shows 5,500 ft of the approach associated with scrambled control and the rest being associated with opportunistic control. Because there are no sections coded as tactical or strategic control, the measure for this approach was recorded as 0 ft AWSCM.

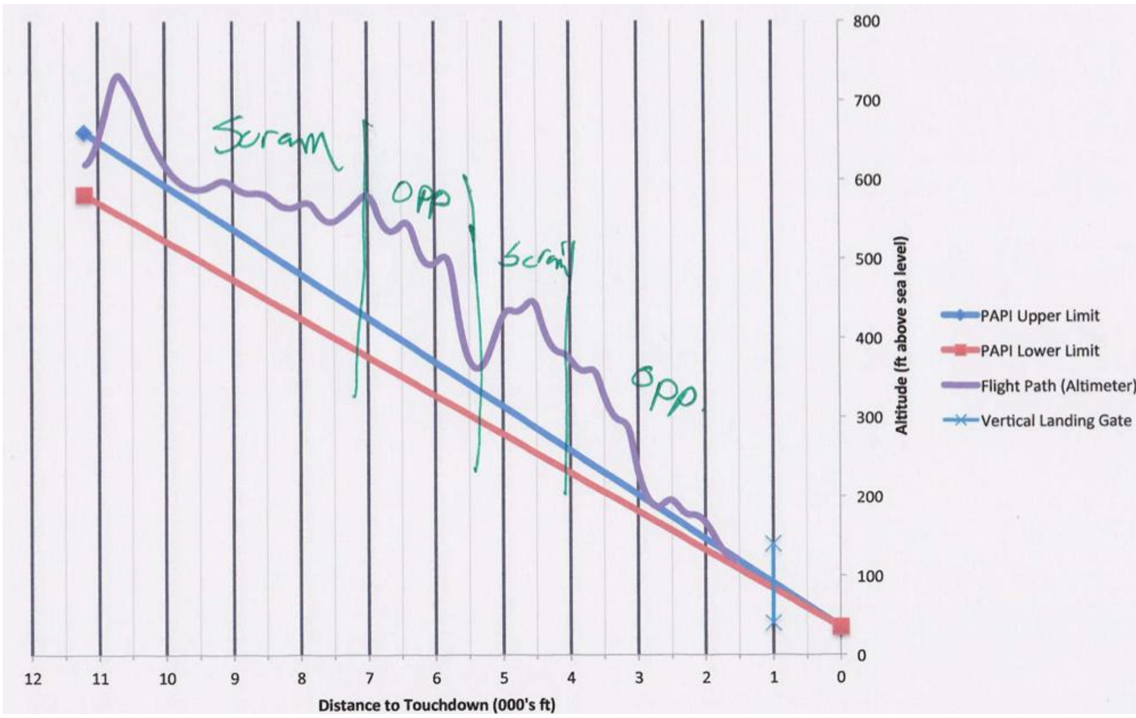

Figure 16. An example of an entire approach graph with a coder's hand written evaluation of control mode.

Analysis of measures. RMSE is ratio data and thus its associated hypotheses can be tested with an analysis of variance (ANOVA) and post hoc tests as appropriate. Conversely, landing success, landing control, and entire approach control are ordinal data that require non-parametric analyses. To test the hypotheses for these three measures, and to avoid losing statistical power through the use of an omni-bus test, five a priori hypotheses were established. The same five a priori hypotheses were for all three non-parametric measures: landing success, landing control, and entire approach control.

The first three a priori null-hypotheses were to test if there was an effect on performance due to wind at each of the three levels of lag, where u_{ij} is performance in lag condition *i* (1 = none, 2 = 240 ms, 3 = 1000 ms) with *j* wind (1 = ideal, 2 = adverse).

$$H_{o1}: \mu_{11} = \mu_{12} \tag{16}$$

$$H_{o2}: \mu_{21} = \mu_{22} \tag{17}$$

$$H_{o3}: \mu_{31} = \mu_{32} \tag{18}$$

The next two null hypotheses were to test if there was an effect on performance due to lag. These focus on where one would expect the smallest differences in performance between levels of lag. For example, it would be expected that the smallest difference in performance would occur between the most challenging no lag condition (with adverse wind) compared to the least challenging low lag condition (with ideal wind). In other words, a weak lag manipulation with an adverse wind manipulation compared to strong lag manipulation without the wind manipulation.

$$H_{o4}: \mu_{12} = \mu_{21} \tag{19}$$

$$H_{o5}: \mu_{22} = \mu_{31} \tag{20}$$

It could be argued that one should also look at where one expects the largest differences in performance, such as between no lag and high lag, or between no lag ideal wind and low lag adverse wind. However each additional pairwise comparison further reduces the statistical power and the ability to detect differences. Thus looking at many differences would reduce the ability to see small differences. It was decided to place greater importance on finding small differences than all differences.

Intercoder disagreements. Two coders were used to assign COCOM control modes to approaches for both the entire approach and landing control. Disagreements between the two coders were resolved by taking the average score. For entire approach data, this meant averaging the distance AWSCM.

For landing success, recall that scrambled, opportunistic, tactical, and strategic are control modes along a continuum. They were thus given numerical values of 1, 2, 3, and 4 respectively. Averaging ordinal data is generally an unacceptable practice since the difference between data points is unknown. However in this case it is possible because a non-parametric ranks test is only concerned with ranking data points from highest to lowest. Thus, an average between 2 and 3 is marked as 2.5, and the non-parametric rank test simply treats it as a number that is greater than 2 and less than 3. When the ranks test is used to analyze the data, 2.5 is ranked after 2 and before 3, just as any number between 2 and 3 would.

It was assumed that reliability of the final approach codes was reflective of reliability of the entire approach control codes. Intercoder reliability of all landing control was used to measure intercoder reliability, though using a sample of the data is an acceptable practice. Intercoder reliability was investigated using Cohen's kappa (Cohen 1960; Lombard, Snyder-Duch, & Braken, 2010; Wood, 2007).

Design

A 3 x 2 repeated measures factorial design was utilized. The independent variables were lag and wind. There were three levels of lag: no lag, 240 ms (low lag), and 1000 ms (high lag). There were two levels of wind: the first was a gentle head wind of 5 kts (ideal wind condition), and the second was a gentle headwind of 5 kts with a 20 kt crosswind gust 1.6 NM from the runway that lasted for 2 seconds (adverse wind condition). Each participant was exposed to all six conditions.

Participant involvement was spread over three sessions that took place 1-7 (M = 3.97, SD = 2.11) days apart. Each session consisted of 16 trials at a particular level of lag. The presentation order of sessions, and thus lag, was counterbalanced with at least two participants for each lag presentation order. Of the 16 trials in each session, the first eight were for participants to warm up and these data were not included in the analysis. For the final eight trials, half were randomly selected to have the adverse wind condition. In other words, a session provided data on four adverse and four ideal wind condition trials at a certain level of lag, and then this was repeated two more times at the two remaining levels of lag. This resulted in a total of 24 trials per participant (four for each of the six conditions).

Procedure

Participants read and signed an informed consent form (see Appendix B). Simulator sickness was explained to each participant and each was asked if they felt any symptoms immediately prior to starting the experiment. Participants were instructed to seek and maintain a 3° glide slope based on the PAPI and to stay aligned with the center of the runway (see instructions script in Appendix D). They were instructed to maintain the glideslope for the entire approach and to aim for the target touchdown markings on the runway, which are 1000 ft past the start of the runway. They were also instructed that the target landing speed for touchdown was 60-65 knots indicated airspeed (KIAS; the airspeed as indicated by aircraft instrumentation). Participants were given a preflight weather briefing, which described the ideal headwind and the possibility of crosswind gusts, though the possibility of crosswind gusts was not discussed in detail. Lastly, participants were informed that the first few trials would be warm ups, but that they would not know when data collection was to begin.

Each trial began with the aircraft 11,064 ft from the runway, at an altitude of 618 ft above sea level and traveling at 75 KIAS in level flight towards the runway. The participants conducted the approach and the simulation was terminated 400 ft short of the touchdown markers. The SSQ was administered after two approaches and again at the end of the 16 trial sessions. A detailed, step-by-step procedure explaining how to use X-Plane and the plugin can be found in Appendix C.

Participants were asked a set of questions by the researcher after completing each session of trials at a particular level of lag. The questions were:

- How do you think you performed during this part of the experiment?;
- Were these approaches difficult?; and
- Did you change how you fly the aircraft to compensate for the system settings?

Participants were prompted to explain their responses.

Results

Glideslope Deviations

Glide slope deviations were the only ratio data recorded in this experiment. As such, the analysis of variance (ANOVA) of glideslope deviations was the only analysis that could provide insight into interactions between the two independent variables and it was considered the primary analysis for this experiment. The nonparametric analyses of other measures acted as additional, and nonetheless important, support. One participant's RMSE data were discarded because they had such difficulty with the high lag condition that they would crash within seconds of being handed control. As a result, they were not able to stay airborne long enough to accumulate a meaningful RMSE. No participants reported simulator sickness symptoms. All other participant data were included in the analyses.

The RMSE descriptive statistics are in Table 5 and Figure 17. Mauchly's test indicated that sphericity was violated ($p < 0.05$). The Greenhouse-Geisser sphericity (Greenhouse & Geisser, 1959) adjustment was used to accommodate the violation of sphericity. A two-way repeated measures ANOVA with an alpha level of 0.05 was conducted to asses the effect of lag and wind on RMSE.

Table 5

RMSE Descriptive Statistics

		Ideal wind	Adverse wind	Main effect
Lag	*n*	*M (SD)*	*M(SD)*	*M(SD)*
No	12	43.20 (17.81)	44.29 (19.87)	43.74 (18.47)
Low	12	55.14 (19.29)	63.42 (19.28)	59.28 (19.33)
High	12	123.12 (56.50)	143.84 (55.71)	133.48 (55.89)
Main Effect	12	73.82 (49.94)	83.85 (55.95)	

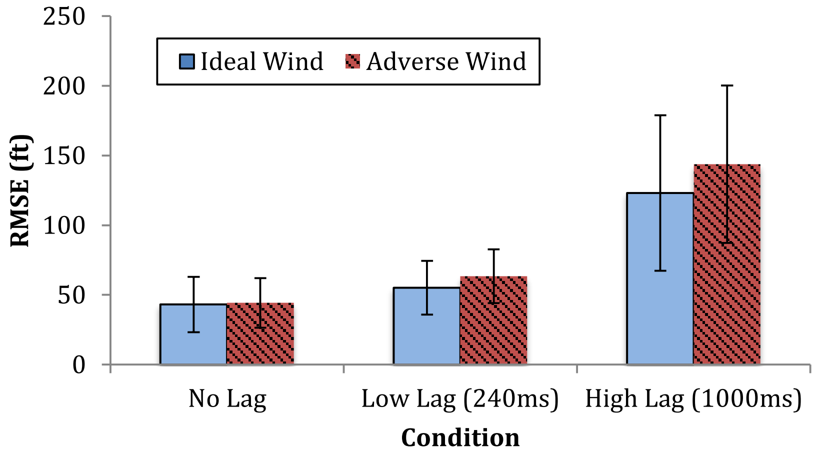

Figure 17. RMSE means and standard deviations across lag and wind conditions.

The analysis revealed a main effect of lag on RMSE ($F(2, 22) = 28.801$, $p < 0.001$, $\eta_p^2 = 0.724$). The partial eta squared indicates that 72.4% of variance is caused by the lag manipulation. A post hoc power analysis experiment (Onwuegbuzie & Leech, 2004) indicated that if the sample represented the population, there is a 99.9% chance that a difference in RMSE due to lag would have been detected in this. The mean RMSEs in the three lag conditions are shown in Figure 18.

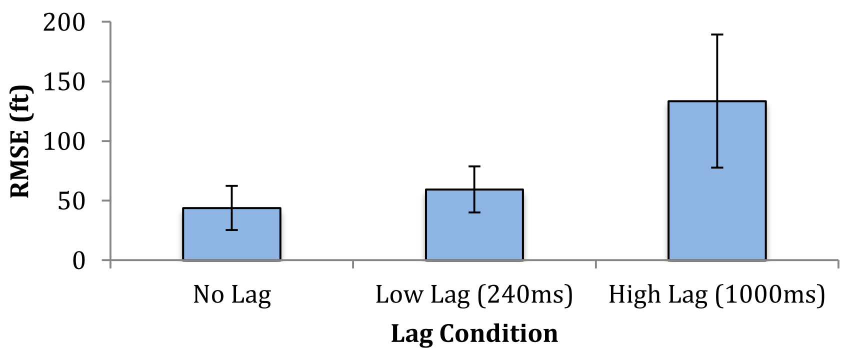

Figure 18. RMSE main lag effects and standard deviations.

The analysis also revealed a main effect of wind on RMSE ($F(1, 11) = 15.186$, $p = 0.002$, $\eta_p^2 = 0.580$). The partial eta squared indicates that 58.0% of variance is caused by the wind manipulation. A post hoc power analysis (Onwuegbuzie and

Leech, 2004) indicated that if the sample represented the population, there is a 94.3% chance that a difference in RMSE due to wind would have been detected in this experiment. The mean RMSEs in the two wind conditions are shown in Figure *19*.

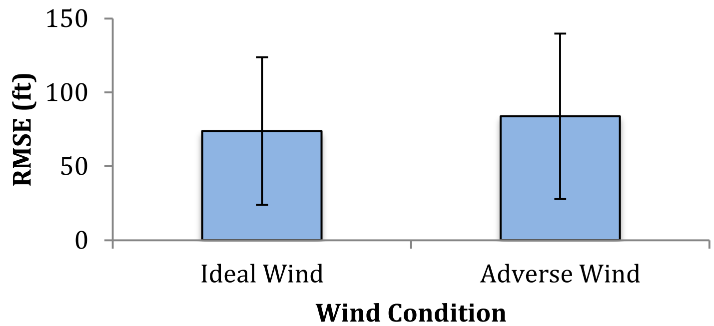

Figure 19. RMSE main wind effects and standard deviations.

No interaction was observed. A post hoc power analysis (Onwuegbuzie and Leech, 2004) indicated that if the sample represented the population, there is a 43.9% chance that a significant difference in RMSE due to a lag and wind interaction would have been detected in this experiment.

A post hoc analysis was conducted in order to determine where the differences in RMSE due to lag manipulations were. The Bonferroni adjustment was used in order to control for the family wise error inflation that accompanies multiple pairwise comparisons. The post hoc test revealed a difference between No Lag and Low Lag ($p = 0.005$), No Lag and High Lag ($p < 0.001$), and Low Lag and High Lag ($p = 0.001$).

Learning. The first eight trials of each session were for the participants to warm up and the data were not used in the analyses. The analyses were based on the data from the remaining eight trials (recall there were four trials for both the adverse and ideal wind conditions). Learning was investigated by conducting a repeated measures ANOVA of the trials collapsed across lag and wind conditions and the means can be found in Figure 20. Mauchly's test indicated that sphericity was violated ($p < 0.05$) and the Greenhouse-Geisser adjustment (Greenhouse & Geisser, 1959) was used to accommodate the violation of sphericity. The analysis revealed no significance difference in performance due to number of approaches conducted by a participant ($F(3, 219) = 0.606$, $p = 0.594$, $\eta_p^2 = 0.008$), though the observed power indicated that there was only a 17.5% chance of detecting a difference in performance.

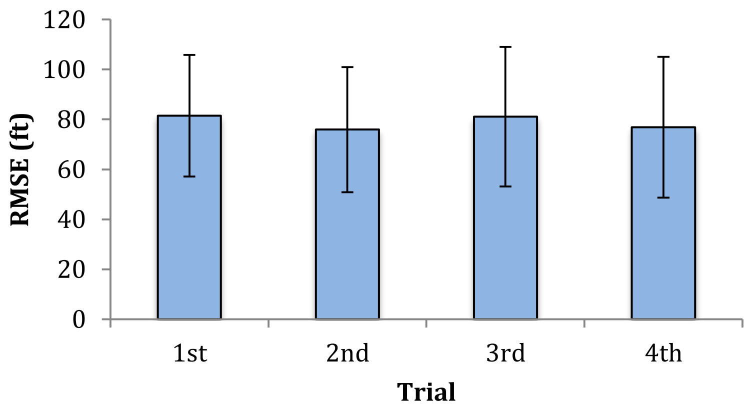

Figure 20. Mean RMSE and standard deviations per trial across lag and wind conditions.

Entire Approach Control

A Friedman Ranks Test (Siegel & Castellan, 1988) with an alpha level of 0.05 was conducted to compare the effect of various levels of lag and wind on distance AWSCM demonstrated by participants. There was a significant difference in distance AWSCM due to experimental manipulations ($\chi^2(5) = 53.559$, $p < 0.001$). The descriptive statistics are in Table 6 and Figure 21.

Table 6

Distance AWSCM Descriptive Statistics (thousands of ft)

		Ideal wind		Adverse wind	
Lag	*n*	*M (SD)*	*Mdn (MAD)**	*M (SD)*	*Mdn (MAD)**
No	13	8.30 (1.71)	8.63 (1.19)	8.19 (1.62)	8.06 (0.94)
Low	13	6.57 (2.48)	7.00 (2.06)	5.46 (2.59)	5.81 (1.68)
High	13	1.41 (1.98)	0.38 (0.38)	1.01 (1.34)	0.38 (0.38)

**Both mean and median are provided, though median is a more meaningful number because this is ordinal data. MAD is Median Absolute Deviation and helps with understanding the variability of the median.*

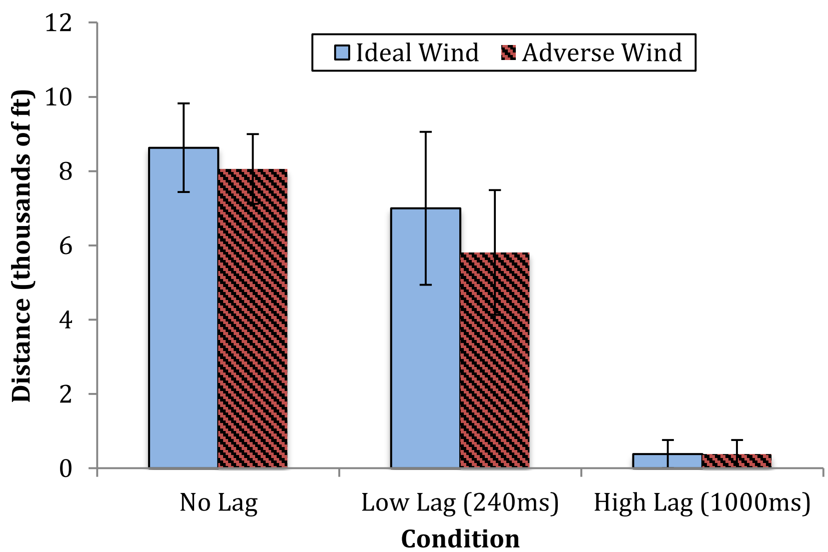

Figure 21. Median distance spent in a safe control mode. Note that the error bars represent the MAD.

The a priori null hypotheses were tested using a Wilcoxon Signed-Rank Test with a Bonferroni correction of $p = 0.01$ to control for inflation of family-wise error. The pairwise comparisons revealed an effect due to lag. It showed less distance AWSCM for the low lag, ideal wind condition (*Mdn* = 7,000 ft) than the no lag, adverse wind condition (*Mdn* = 8,060 ft; Z = -2.622, $p = 0.009$). It also revealed less distance AWSCM for the high lag, ideal wind condition (*Mdn* = 380 ft) than the low lag, adverse wind condition (*Mdn* = 5,810 ft; Z = -3.110, $p = 0.002$). No other significant differences were found.

Landing control

The possible control modes were scrambled (1), opportunistic (2), tactical (3), and strategic (4). A Friedman Ranks Test with an alpha level of 0.05 revealed that there was a difference in landing control modes due to the experimental manipulations ($\chi^2(5) = 173.030$, $p < 0.001$). The descriptive statistics are in Table 7 and Figure 22.

Table 7

Landing Control Descriptive Statistics

		Ideal wind		Adverse wind	
Lag	*n*	*M (SD)*	*Mdn**	*M (SD)*	*Mdn**
No	13	3.36 (0.40)	3.50	3.36 (0.44)	3.50
Low	13	2.94 (0.68)	3.00	3.10 (0.69)	3.00
High	13	1.71 (0.65)	2.00	1.67 (0.64)	1.75

**Both mean and median are provided, though median is a more meaningful number because this is ordinal data. Furthermore, all median absolute deviations were non-zero but less than one.*

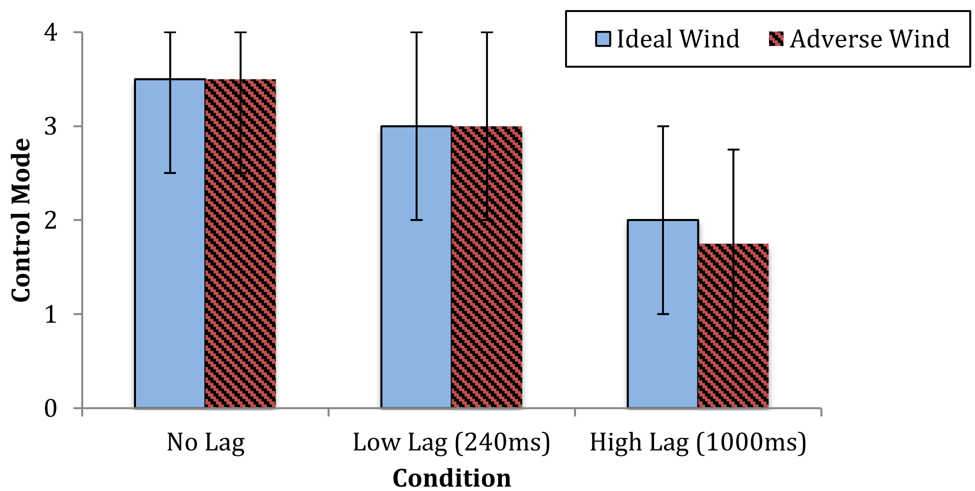

Figure 22. Median landing control mode, where: 1=scrambled, 2=opportunistic, 3=tactical, and 4=strategic. Note that the error bars represent the highest MAD value.

The a priori null hypotheses were tested using a Wilcoxon Signed-Rank Test with a Bonferroni correction of $p = 0.01$ to control for inflation of family-wise error. The pairwise comparisons revealed an effect due to lag. It showed a lower landing control mode for the low lag, ideal wind condition (*Mdn* = 3.00, tactical) than the no lag, adverse wind condition (*Mdn* = 3.50, between tactical and strategic; Z = -4.341, $p < 0.001$). It also revealed a lower landing control mode for the high lag, ideal wind condition (*Mdn* = 2.00, opportunistic) than the low lag, adverse wind condition (*Mdn* = 3.00, tactical; Z = -5.747, $p < 0.001$). No other significant differences were found.

A post-hoc analysis of the effect of lag collapsed across wind conditions on the number of safe landing codes was conducted (see Figure 23 and Table 8). The chi-square goodness of fit tests with a Bonferroni adjustment revealed an effect on safe landing code frequency due to lag. High lag (1000 ms) had fewer safe landing codes than low lag (240 ms; $\chi^2(1) = 99.604$, $p < 0.001$), and low lag (240 ms) had fewer safe landing codes than no lag ($\chi^2(1) = 22.515$, $p < 0.001$).

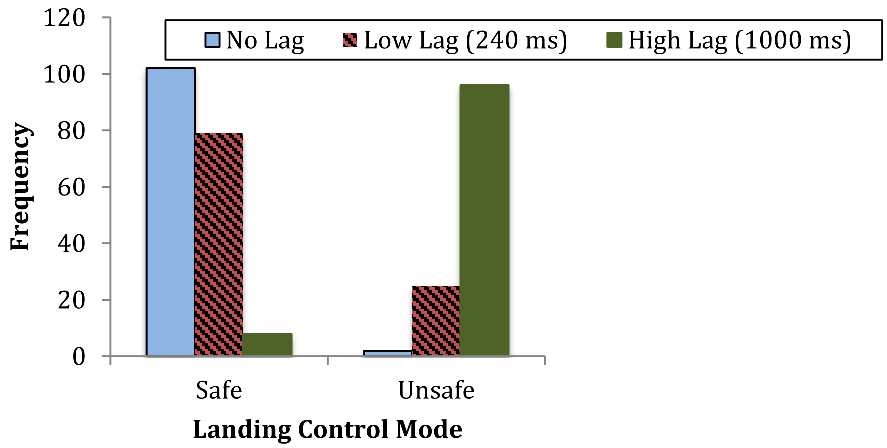

Figure 23. Safe landing control mode code frequencies collapsed across wind conditions.

Table 8

Landing Code Frequencies

		Control Mode		
Condition	*n*	Safe	Unsafe	% Unsafe
No Lag	13	102	2	2.0
Low Lag	13	79	25	24.0
High Lag	13	8	96	92.3

Landing Success

A Friedman Ranks Test (Siegel & Castellan, 1988) with an alpha level of 0.05 was conducted to compare the effect of various levels of lag and wind on the landing success of participants. There was a statistically significant difference in landing success due to experimental manipulations ($\chi^2(5) = 47.574$, $p < 0.001$). The descriptive statistics are in Table 9 and Figure 24.

Table 9

Number of Landing Successes Descriptive Statistics

		Ideal wind		Adverse wind	
Lag	*n*	*M (SD)*	*Mdn (MAD)*	*M (SD)*	*Mdn (MAD)*
No	13	3.77 (0.83)	4.00 (0)	3.77 (0.44)	4.00 (0)
Low	13	3.69 (0.63)	4.00 (0)	3.77 (0.60)	4.00 (0)
High	13	1.38 (1.19)	1.00 (0)	1.54 (1.33)	2.00 (1)

**Both mean and median are provided, though median is a more meaningful number since this is ordinal data.*

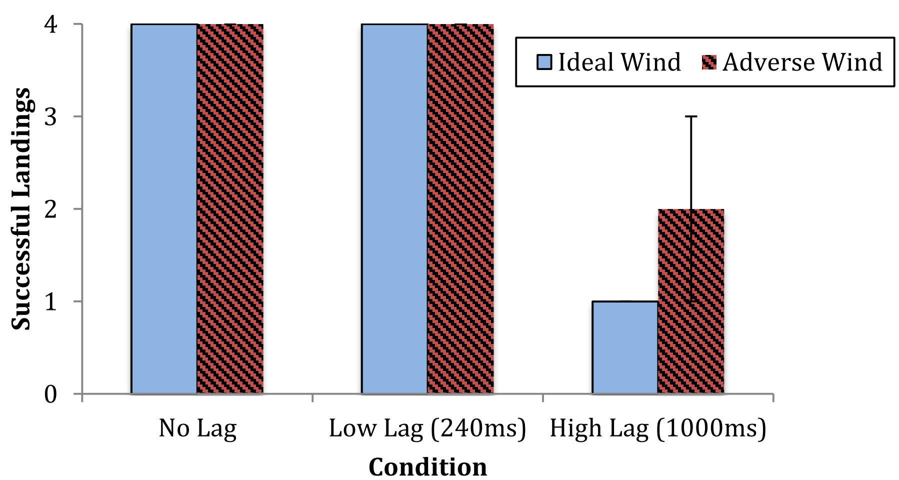

Figure 24. Median number of successful landings in four attempts. Note that the error bars represent the MAD.

The a priori null hypotheses were tested using a Wilcoxon Signed-Rank Test with a Bonferroni correction of $p = 0.01$ to control for inflation of family-wise error. The pairwise comparisons revealed an effect due to lag. It showed fewer landing successes for the high lag, ideal wind condition (*Mdn* = 1) than the low lag, adverse wind condition (*Mdn* = 4; Z = -3.130, $p = 0.002$). No other significant differences were found.

Intercoder Reliability

Using SPSS, kappa was found to be 0.663. Cohen's kappa is a conservative measure (Lombard et al., 2010), so a value above 0.60 can be considered reasonably reliable (e.g., Landis & Koch, 1977; Wood, 2007). Furthermore, the use of COCOM control modes to evaluate pilot performance is a novel activity and as such, lower agreement levels can be acceptable (Lombard et al., 2010). Note that the greatest number of disagreements occurred between tactical and strategic coding (see Figure 25).

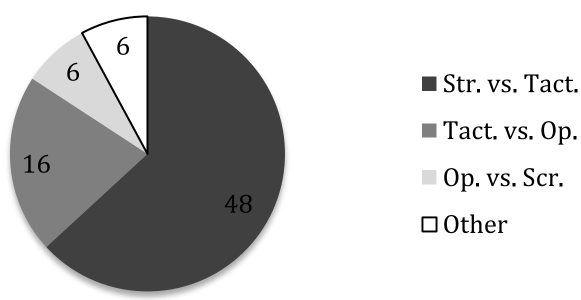

Figure 25. The nature of intercoder disagreements.

Post Session Interviews

A detailed synthesis of the post-experiment interviews can be found in Appendix E. In reviewing the interviews, several common themes emerged. These are summarized in Table 10. The data were not gathered though a proper knowledge elicitation process. As such, it was used to support the analysis of the other measures and to help interpret the results.

Table 10

Interview Themes

Theme	Share of Participants Mentioning Theme
The simulator hardware was unusual and/or challenging to use.	69%
Adopted a “wait and see” control input strategy.	61%
Used especially small and/or smooth inputs.	54%
Thought there was an adverse wind condition when it was only a 5 knot wind.	46%
Changed how they used the rudder.	46%
Did not manage aircraft speed at the same time as other goals such as horizontal alignment.	46%
Used unusual control input strategies such as timed inputs looking and away from display.	46%
Avoided large nose down aircraft attitude and/or accepted a high glideslope.	38%
When deviations became large in the high lag setting, they knew they would “lose it” (i.e., control of the aircraft).	38%
Short final was especially challenging with lag.	38%
Would take care of horizontal deviations and then vertical deviation, or vice versa.	15%

Experience. This experiment did not set out to determine the relationship between experience and ability to accommodate lag, however participant flight experience was recorded. While no data analysis was run to see if experience was a covariate, casual observation of participant experience and performance, Table 11, would suggest it is not. Notice that some high experienced participants had low performance while some low experience participants had high performance. There does not seem to be a clear correlation between flight experience and performance.

Table 11

Participant Flight Experience and Performance

Experience Rank*	Average RMSE (Rank)	Landing Rate (Rank)
1	75.5ft (8)	0.71 (8)
2	160.8ft (13)	0.63 (11)
3	59.9ft (2)	0.75 (7)
4	70.9ft (6)	0.79 (5)
5	67.4ft (4)	0.88 (3)
6	46.4ft (1)	0.92 (2)
7	128.7ft (12)	0.58 (12)
8	108.7ft (11)	0.83 (4)
9	68.4ft (5)	0.71 (8)
10	81.1ft (9)	0.79 (5)
11	71.5ft (7)	0.67 (10)
12	60.2ft (3)	0.96 (1)
13	107.1ft (10)	0.46 (13)

**Note that this column ranks participants from highest number of flight hours to lowest number of flight hours. The actual number of flight hours is not listed in order to protect the identity of participants.*

Discussion

The primary purpose of this study was to explore the effect system lag has on UAS operator performance. This was done by introducing this characteristic of UASs to a simulated manned system, which provides a rough approximation of an unassisted, internally piloted, manual landing UAS.

Effects on Performance

Table 12 summarizes the effects of lag and wind on control performance found in this study. Compared to ideal wind, adverse wind increased RMSE. Compared to no lag, low lag (240 ms) increased RMSE, reduced entire approach control, and reduced landing control. Compared to low lag (240 ms), high lag (1000 ms) increased RMSE, reduced entire approach control, reduced landing control, and reduced the number of landing successes.

Note that four of the hypotheses were not supported by the results. Three of these hypotheses were that wind would cause an effect on entire approach control, landing control, and landing success; however they did not reveal an effect due to wind at any level of lag. The fourth was the hypothesis that low lag (240 ms) would have an effect on landing success when compared to no lag, but it did not.

Table 12

Significant Differences

	Flight Segment			
	Entire Approach		Landing	
Conditions Compared	RMSE	Coding	Coding	Success
No Lag vs. Low Lag	X	X	X	
Low Lag vs. High Lag	X	X	X	X
No Lag vs. High Lag	X	X*	X*	X*
Ideal Wx vs. Adverse Wx	X			

**This comparison was not one of the a priori comparisons, but this difference is based on inference from the other a priori comparisons.*

This study revealed that 1000 ms of lag resulted in degraded performance and unsafe control. The lower level of lag, 240 ms, also resulted in degraded performance, but with high landing success rates. No interaction between wind and system lag was observed. The wind manipulation itself was relatively benign, though warning participants about the possibility of crosswind gusts affected their perception of the simulation and may have affected their control strategies. A majority of participants did report altering their piloting techniques to accommodate lag. They adopted strategies such as: "wait and see," the use of small control inputs, managing one or two goals at time, extra reliance on or abandonment of the rudder, and focusing on yoke position instead of the simulator's display. A more detailed discussion on control strategies and the results of the interviews can be found in Appendix F.

Using COCOM

This study introduced a novel measurement to aircraft performance: COCOM control mode coding (entire approach control and landing control). These measures had coders assess the JCS level of control by reviewing the aircraft's position with respect to the target glide slope. The use of control mode coding complemented the other, more traditional measures (RMSE and landing success) and agreed with them in demonstrating that 1000 ms of lag resulted in unsafe flight. It also provided additional insight into the safety of flight in other conditions. It helped explain why RMSE revealed an effect due to 240 ms when the landing success measure did not: 240 ms of lag degraded performance, but not so much so that landing success was affected. Landing control helped quantify this degraded performance by indicating that 79 out of the 104 landing attempts at 240 ms of lag were conducted in a safe control mode. Contrast this with 102 out of 104 for no lag and 8 out of 104 for 1000 ms of lag. In conjunction with RMSE and landing success, the use of control mode codes also helped with understanding the effect of the wind manipulation, which would blow the aircraft off course, but did not affect participant control of the aircraft. A further discussion of the measures used in this study can be found in Appendix G.

In addition to being a useful measure, COCOM is helpful in describing participant control behavior in this study. Participants had little understanding of the system's lag characteristics and had difficulty evaluating the outcome of their control inputs. This resulted in large overshoots, which at best manifested itself as inefficient PIOs and at worst add hoc trial and error. Participant attention was focused on lag and its effects, thus their event horizon was very close and they had little available time to consider effects beyond the immediate time frame. They also had difficulty managing the multiple goals of aircraft speed, horizontal alignment, and glide slope. The above description of participant behavior closely approximates the COCOM definitions of scrambled and opportunistic control (Hollnagel & Woods, 2005), which can be found in Appendix A.

Now contrast the applicability of COCOM to the inappropriateness of applying COM or OCM. Beyond the fact that both models were developed assuming minimal system lag, other considerations limit their applicability as well. While the landing task was a relatively simple one, there were still several control loops. This study involved four controls (yoke, rudder, flaps, and throttle) meant to control pitch, yaw, roll, and velocity, but COM was developed based the control of one axis (McRuer & Jex, 1967) and is thus less appropriate. In this regard, OCM is more applicable because it takes into account multiple feedback loops (Baron & Kleinman, 1969). However participants had difficulty controlling the aircraft in the high lag setting because they could not accurately predict the effect of control inputs. This operator "predictive error" is not accounted for by OCM, which also assumes the controller has complete understanding of the system's characteristics (Baron & Kleinman, 1969; Kleinman, Baron, & Levison, 1971). In the context of this experiment, COCOM is much more applicable and useful in describing controller behavior, especially for the challenging 1000 ms lag setting.

Study Limitations

Applicability. This study implemented a characteristic of BLOS UASs in a simulated Cessna 172N aircraft. The assumption was that it would affect a trained 172 pilot in the same way that it would affect a UAS pilot that had not experienced this characteristic before. This assumption was made because the researcher did not have access to a readily available pool of UAS pilots. This limits the applicability of the results to certain UAS configurations; the results and conclusions are most applicable to internally piloted manual landing UASs.

Simulator fidelity. Another limitation of this study was the low fidelity of the simulator. It was fixed base with a relatively small display. Participants also remarked that the controls were overly sensitive and the springs in the yoke pushed it to a neutral position when released, which resulted in the aircraft having a nose up attitude. To overcome the nose up attitude and maintain the target glide slope, participants had to constantly push forward on the yoke. At times, the necessary correction required a large nose down correction, however this resulted in the runway and horizon moving off of the small monitor. Furthermore, the ground in X-Plane is not well defined, as shown in Figure 26. Based on the appearance of the ground, it is difficult to determine the altitude and attitude of the aircraft. Closer inspection of the altimeter and the attitude indicator in Figure 26 shows that the aircraft is 900 ft in the air and only slightly nose down.

Figure 26. A screenshot of the aircraft in a shallow dive. Notice the attitude indicator and altimeter show the aircraft is not in danger, but this it hard to tell based on the view.

Conversely, it is likely that the experiment benefited from the nose up trim of the aircraft. With the yoke, participants were not able to “set it and forget it” as they did the throttle. It required the participants to actively apply control inputs throughout the entire approach in order to maintain a negative glide slope. This meant that they had to constantly demonstrate their ability to control the aircraft, which was the focus of this experiment.

Variance. Recall the figure that was first shown in results section, reproduced below as Figure 27. Notice the large amount of variance between subjects. This is not a problem for the data analysis because the experiment was a repeated measures design, however it does merit further investigation and could suggest a confounding variable. It may be that lag does not affect all participants equally. This could be further investigated with a more detailed exploration of experience. The variability may also be the result of the simulator's low fidelity. Some participants may have been able to adapt to the single small screen and the crude yoke better than others. It is possible that a higher quality simulator would reduce the variability between participants. Other common strategies to reduce variability, such as increasing sample size and increasing the number of data points per participant, could also be adopted.

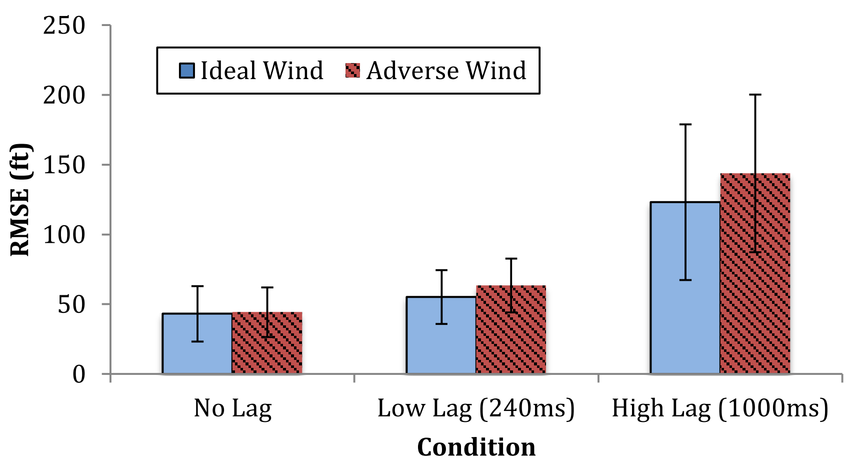

Figure 27. Mean RMSE and standard deviations.

Future Directions

This study highlights many areas for future research. Firstly, the wind manipulation was relatively benign and the 240 ms level of lag should be investigated further with more challenging wind phenomena. Setting wind aside, the range of lag between 240 ms and 1000 ms should be investigated as well. It is possible that somewhere in this range of lag there is a threshold beyond which the manual control of aircraft is unsafe, much like the one that Lane et al. (2002) suggested for zero buoyancy vehicles. In addition to constant lag, variable lag presents another opportunity for future research. Though the baseline lag in this study was negligible, it demonstrated that network lag is not constant and affects all networks.

It is recommended that future research similar to this study include control mode coding as a supporting analysis in order to help determine the level of lag at which flight becomes unsafe. Additionally, the use of COCOM control mode coding to evaluate pilot performance needs to be further refined. Notably, distinguishing between strategic and tactical control could be explored and clarified. Further development of COCOM could help establish an aviation-focused human reliability assessment framework that accurately predicts and quantifies risks.

As previously mentioned, the wind manipulation was relatively weak compared to the lag manipulation. However, the mere mention of possible adverse wind phenomena during the pre-flight brief may have an effect on participant performance during lag conditions. It would be interesting to investigate the effect the pre-flight warning of adverse wind has on pilot performance. Weather conditions are a very real aspect of flight. In order to adapt to lag, pilots must be able to distinguish between the effects of wind and those of lag because the appropriate techniques to both phenomena may not be the same. It is recommended that future studies include weather manipulations in order to better reflect reality.

It is evident that participants adopt various control strategies to accommodate system lag. This could be further investigated in order to determine which strategies improve flight performance and safety. This could potentially lead to the development and evaluation of training plans that help pilots learn to accommodate system lag. In addition to researching training, future direction should also involve investigating the effects of an enhanced simulator. Enhancements could include: larger displays, advanced instrumentation, predictive displays, improved control hardware, and assisted-recovery autopilot.

Lastly, this study revealed a large amount of variance between participants. While this does not appear to be due to the number of flight hours flown, there may be some other aspect of experience that contributed to this and it warrants further investigation. Future research should look into the effect of simulator experience, video game experience, and experience with controlling systems over networks. It is also possible that this was due to the low fidelity of the simulator. Conducting a similar study with an improved and/or UAS simulator would help build upon the findings of this study. Lessons learned and summary of recommendations for future work are included in Appendix H.

Conclusions

This study showed that trained pilots, with limited exposure to lag, could not safely land a simulated aircraft with 1000 ms of system lag. When system lag was 240 ms, pilots could successfully land the simulated aircraft, but both performance and level of safety were degraded. This study also showed that a 20 kt crosswind gust that lasts 2 s and occurs 1.6 NM from the runway affected approach performance, but not landing success rates or pilot control of the simulated aircraft. This study demonstrated the applicability and usefulness of COCOM control mode codes in evaluation and understanding the effect of lag on flight performance and safety. Lastly, this study showed that pilots adopt various control strategies to overcome system lag.

References

Baron, S. (1973). Application of the optimal control model for the human operator to reliability assessment. *IEEE Transactions on Reliability, R-22*(3), 157-164. doi: 10.1109/TR.1973.5215931

Baron, S., & Kleinman, D.L. (1969). The human as an optimal controller and information processor. *IEEE Transactions on Man-Machine Systems, 10*(1), 9-17. doi: 10.1109/TMMS.1969.299875

Casali, J.G., & Wierwille, W.W. (1980). The effects of various design alternatives on moving base driving simulator discomfort. *Human Factors, 22*(6), 741-756.

Cessna Aircraft Company (1977). Pilot's operating handbook: Skyhawk Cessna model 172N. Wichita, Kansas: Cessna Aircraft Company.

Chan, R.B., Jhoun, J.H., & Childress, D.S., (1997). An investigation of tracking behavior exhibited during balancing and standing. *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1667-1670). Chicago, Illinois: IEEE. doi: 10.1109/IEMBS.1997.757040

Chen, J.Y.C., Haas, E.C., & Barnes, M.J. (2007). Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics, 37(6),* 1231-1245. doi:10.1109/TSMCC.2007.905819

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Conklin, J. (1957). Effect of control lag on performance in a tracking task. *Journal of Experimental Psychology, 53,* 261-268.

Cooper, G.E., & Harper, R.P. (1969). *The use of pilot rating in the evaluation of aircraft handling qualities.* (Technical Note D-5153). Neuilly-Sur-Seine, France: Advisory Group for Aerospace Research and Development. Retrieved from NASA Technical Reports Server website: http://ntrs.nasa.gov/search.jsp

Craighead, J., Murphy, R., Burke, J., & Goldiez, B. (2007). A survey of commercial & open source unmanned vehicle simulators. *Proceedings from IEEE International Conference on Robotics and Automation* (pp. 852-857) Rome, Italy: IEEE.

Day, C.N. (1977). Modern control application to manual control-historical perspective and future direction. *IEEE Conference on Decision and Control* (pp. 226-229). Ohio, USA: IEEE. doi: 10.1109/CDC.1977.271576

Davis, J., Smyth C., & McDowell K. (2010). The effects of time lag on driving performance and a possible mitigation. *IEEE Transactions on Robotics, 26(3),* 590-593. doi:10.1109/TRO.2010.2046695

De Vries, S.C. (2005). *UAVs and control delays*. (Report TNO-DV3 2005 A054). Soesterberg, The Netherlands: TNODefence, Security and Safety.

DiStefano, J.J., Stubberud, A.R., & Williams, I.J. (1990). *Schaum's outline of theory and problems of feedback and control systems*. New York: McGraw-Hill.

Dong, H., Hsiang, S.M., & Smith, J.L. (2009). An optimal-control model of vision-gait interaction in a virtual walkway. *IEEE Transactions on Systems, Man, and Cybernetics, 39*(1), 156-166. doi: 10.1109/TSMCB.2008.2003427

Draper, M.H., Viire, E.S., Furness, T.A., & Gawron, V.J. (2001). Effects of image scale and system time delay on simulator sickness within head couple virtual environments. *Human Factors, 43*(1), 129-146.

Ellis, S.R., Mania, K., Adelstein, B.D., & Hill, M.I. (2004). Generalizeability of latency detection in a variety of virtual environments. *Proceedings of the Human Factors and Ergonomics Society 48th annual meeting* (pp. 2632-2636). Santa Monica, USA: Human Factors and Ergonomics Society.

Ferrel, W. (1965). *Remote manipulation with transmission delay.* (Technical Note D-2665) Boston, MA: Massachusetts Institute of Technology. Retrieved from NASA website: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19650007728_1965007728.pdf

Feth, D., Groten, R., Peer, A., & Buss, M. (2009). Control-theoretic model of haptic human-human interactions in a pursuit tracking task. *The 18th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 1106-1111). Toyama, Japan: IEEE. doi: 10.1109/ROMAN.2009.5326316

Franklin, G.F., Powell, J.D., & Emami-Naeini, A. (1994). *Feedback control of dynamic systems*. Reading, Massachusetts: Addison-Wesley.

Frost & Sullivan (2007). *Study Analysing the current activities in the field of UAV* (Report Part 1). European Commission Enterprise and Industry Directorate-General. Retrieved from http://ec.europa.eu/enterprise/policies/security/files/uav_study_element_2_en.pdf

Fujita, Y., & Hollnagel, E. (2004). Failures without errors: quantification of context in HRA. *Reliability Engineering and System Safety, 83,* 145-151. doi: 10.1016/j.ress.2003.09.006

Ganesh, S., & Bajcsy, R. (2008). Recognition of human actions using an optimal control based motor model. *IEEE Workshop on Applications of Computer Vision* (pp. 1-6). Copper Mountain, Colorado: IEEE. doi: 10.1109/WACV.2008.4544021

Greenhouse, S.W., & Geisser, S. (1959). On the methods in the analysis of profile data. *Psychometrika, 24,* 95-112.

Hess, R.A. (1981). Aircraft control-display analysis and design using the optimal control model of the human pilot. *IEEE Transactions on Systems, Man, and Cybernetics, 11*(7), 465-480. doi: 10.1109/TSMC.1981.4308724

Hollnagel, E. (1997). Reliability analysis and operator modeling. *Reliability Engineering and System Safety, 52(3),* 327-337.

Hollnagel, E. & Woods, D.D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies, 18(6)*, 583-600.

Holnagel, E. & Woods, D.D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: CRC Press.

Hopcroft, R., Burchat, E., & Vince, J. (2006). *Unmanned Aerial Vehicles for Maritime Patrol: Human Factors Issues* (Report DSTO-GD-0463). Victoria, Australia: Australian Government Department of Defence. Retrieved from Australian Government Defense Science and Technology Organization website: http://dspace.dsto.defence.gov.au/dspace/bitstream/1947/4404/1/DSTO-GD-0463%20PR.pdf

Hsiang, S.M., Dong, H., & Karakostas, T. (2006). Optimal control in a virtual walkway. *IEEE Mountain Workshop on Adaptive and Learning Systems* (pp. 104-109). Logan, Utah: IEEE. doi: 10.1109/SMCALS.2006.250699

Inaba, T., & Matsuo, Y. (1997). Loop-shaping characteristics of a human operator in a force reflective manual control system. *IEEE International Conference on Systems, Man, and Cybernetics* (pp. 3621-3625). Orlando, Florida: IEEE. doi: 10.1109/ICSMC.1997.633230

Jennings, S., Reid, L.D., Craig, G., & Kruk, R.V. (2004). Time delay in visually coupled systems during flight test and simulation. *Journal of Aircraft, 41(6),* 1327-1335.

Junker, A.M., & Levison, W.H. (1977). Application of control theory to the investigation of roll motion effects on pilot control behavior. *IEEE Conference on Decision and Control* (pp. 247-254). Ohio, USA: IEEE. doi: 10.1109/CDC/1977.271580

Jury, E., & Pavlidis, T. (1963). A literature survey of biocontrol systems. *IEEE Transactions on Automatic Control, 8*(3), 210-217. doi: 10.1109/TAC.1963.1105562

Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology*, *3(3),* 203-220.

Klein, G., Ross, K.G., Moon, B.M., Klein, D.E., Hoffman, R.R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems, 18(3),* 81-83. doi: 10.1109/MIS.2003.1200735

Kleinman, D., Baron, S., & Levison, W. (1971). A control theoretic approach to manned-vehicle systems analysis. *IEEE Transactions on Automatic Control, 16*(6), 824-832. doi: 10.1109/TAC.1971.1099842

Krstic, M. (2008). Optimal adaptive control-contradiction in terms or a matter of choosing the right cost functional. *IEEE Transactions on Automatic Control, 53*(8), 1942-1947. doi: 10.1109/TAC.2008.929464

Lane, J.C., Carignan, C.R., Sullivan, B.R., Akin, D.L., Hunt, T., & Cohen, R. (2002). Effects of time delay on telerobotic control of neutral buoyancy vehicles. *Proceedings of the IEEE International Conference on Robotics and Automation* (pp 2874-2879). Washington, D.C.: IEEE. doi: 10.1109/ROBOT.2002.1013668

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Lombard, M., Snyder-Duch, J, & Braken, C.C. (2010). Intercoder reliability: practical resource for assessing and reporting intercoder reliability in content analysis. Retrieved April 1, 2012 from http://astro.temple.edu/~lombard/reliability/

Mahoney, R.M. (1994). Towards a control theory model of human target-directed movements. *IEEE International Conference on Systems, Man, and Cybernetics* (pp. 30-35). San Antonio, Texas: IEEE. doi: 10.1109/ICSMC.1994.399807.

McCarley, J.S., & Wickens, C.D. (2005). *Human factors implications of UAVs in the national airspace*. (FAA Technical Report AHFD-05-05/FAA-05-01) University of Illinois, Aviation Human Factors Division. Retrieved from University of Illinois website: http://www.humanfactors.illinois.edu/Reports&PapersPDFs/TechReport/05-05.pdf

McRuer, D.T., & Jex, H.R. (1967). A review of quasi-linear pilot models. *IEEE Transactions on Human Factors in Electronics, 8*(3), 231-249. doi: 10.1109/THFE.1967.234304

McRuer, D., & Weir, D.H. (1969). Theory of manual vehicular control. *IEEE Transactions on Man-Machine Systems, 10*(4), 257-291. doi: 10.1109/TMMS.1969.299930

Moss, J.D., Muth, E.R., Tyrrell, R.A., & Stephens, B.R. (2010). Perceptual thresholds for display lag in a real visual environment are not affected by field of view or phychophysical technique. *Displays, 31,* 143-149. doi: 10:1016/j.displa.2010.04.002

Mulder, M., Van Paassen, M.M., Flach, J.M., & Jagacinski, R.J. (2006). Fundamentals of manual control. In W.S. Marras, & W. Karwowski (Eds.), *Fundamentals and Assessment Tools for Occupational Ergonomics* (pp. 12.1-12.26). Boca Raton, Florida: CRC Press.

Onwuegbuzie, A.J., & Leech, N.L. (2004). Post hoc power: a concept whose time has come, *Understanding Statistics, 3*(4), 201-230.

Parush, A. (2006). Human errors in UAV takeoff and landing: theoretical account and practical implications. In N. Cooke (Ed.), *Human Factors of Remotely Operated Vehicles Volume 7* (pp. 91-103). Amsterdam, Netherlands: JAI Press.

Patterson, R., Winterbottom, M.D., & Byron, J.P. (2006). Perceptual issues in the use of head mounted visual displays. *Human Factors, 48*(3), 555-573.

Ricard, G.L. (1994). Manual control with delays: A bibliography. *Computer Graphics, 28*(2), 149-154.

Ricard, G.L. (1995). Acquisition of control skill with delayed and compensated desplays. *Human Factors, 37*(3), 652-658.

Siegel, S., & Castellan, N.J. (1988). *Nonparametric statistics for the behavioral sciences*. Boston, Massachusetts: McGraw-Hill.

Staplefor, R.L., McRuer, D.T., & Magdaleno, R.E. (1967). Pilot describing function measuremnets in a multiloop task. *IEEE Transactions on Human Factors in Electronics, 8*(2), 113-125. doi: 10.1109/THFE.1967.233628

Sullivan, B., Ware, C., & Plumlee, M. (2006). Predictive displays for survey vessels. *Proceedings of the Human Factors and Ergonomics Society 50th annual meeting* (pp. 2424-2428). San Francisco, CA: Human Factors and Ergonomics Society.

Tadema, J., Theunissen, E., & Koeners, J. (2007). Using perspective guidance overlay to improve UAV manual control performance. *Proceedings of SPIE 6559: Vol. 65590-C.* (pp. 1-11). Orlando FL: SPIE. doi:10.1117./12/719238

Thurling, A.J. (2000). *Improving UAV flying qualities using time delay compensation*. Masters of Science Thesis, United States Air Force Institute of Technology, USA. Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA378376

Van Erp, J.B.F., & Van Breda, L. (1999). *Human factors issues and advanced interface design in maritime unmanned aerial vehicles: A project overview.* (Report TNO TM-99-A004). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

Van Paassen, R., & Mulder, M. (2006). Identification of human control behavior. In W. Karwowski (Eds.), *International encyclopedia of ergonomics and human factors volume 1* (pp. 400-407). Retrieved from http://www.crcnetbase.com

Watson, B., Walker, N., Ribarsky, W., & Spaulding, V. (1998). Effects of variation in system responsiveness on user performance in virtual environments. *Human Factors, 40(3),* 403-414.

Whiteley, A.L. (1946). Theory of servo systems, with particular reference to stabilization. *Journal of the Institution of Electrical Engineers, 93*(34), 353-367. doi: 10.1049/ji-2.1946.0083

Wickens, C.D. (1992). *Engineering psychology and human performance*. New York, New York: Harper Collins Publishers.

Williams, K.W. (2004). *A summary of unmanned aerial aircraft accident/incident data: Human factors implications*. (FAA Technical Report AM-04/24). Oklahoma City, OK: FAA.

Wood, J. M. (2007, October 3). Understanding and Computing Cohen's Kappa: A Tutorial. *WebPsychEmpiricist*. Retrieved May 10, 2012 from http://wpe.info/vault/wood07/Wood07.pdf

Worm, A. (1999). Mission efficiency analysis of tactical joint cognitive systems. *Proceedings of the NATO International Symposium on Modeling and Analysis of Command and Control*, Paris, France.

ZTI (2012, June 2). NetDisturb. Retrieved June 2, 2012 from http://www.zti-telecom.com/EN/ NetDisturb.html

Appendix A

Coder Training Plan

Hello and thank you for being a coder! This will involve looking at a flight path that has been plotted on a graph and evaluating the level of control the pilot is demonstrating. This document will help guide you. You will learn about "Control Modes," how they relate to this task, a description of this task, and then some tips that will help you evaluate the data.

Control Mode Descriptions

1. What are control mode codes?

Contextual Control Model (COCOM) identifies four different modes of control: scrambled, opportunistic, tactical, and strategic, which are referred to as control modes (Fujita & Hollnagel, 2004; Hollnagel & Woods, 2005). Control modes are described as:

- *scrambled*, which consists of random trial and error with little to no planning or thinking;
- *opportunistic*, which involves only limited planning due to lack of understanding or limited time, which results in inefficient actions and wasted attempts;
- *tactical*, which takes some delayed effects into account and actions often follow known rules, though they can still be ad hoc; and
- *strategic*, which considers high level goals and understands the dependencies between actions and multiple goals (Hollnagel & Woods, 2005)

Hollnagel & Woods' (2005) description of the control mode characteristics is outlined in Table 13.

Table 13

Control Modes. Adapted from Hollnagel & Woods (2005)

Control Mode	Number of goals	Subjectively available time	Evaluation of outcome	Selection of action
Strategic	Several	Abundant	Elaborate	Based on models/predictions
Tactical	Several (limited)	Adequate	Detailed	Based on plans/experience
Opportunistic	One or two (competing)	Just adequate	Concrete	Based on habits/association
Scrambled	One	Inadequate	Rudimentary	Random

Control Mode Application

2. How do they relate to this task?

The aforementioned control modes have been operationalized to suit the needs of this research. Below is a description of how the basic definitions of each control mode relates to participant performance in this research. Each control mode is first described with respect to the model (COCOM). After this description, the following paragraph then describes the operational definition of each control mode as it pertains to this study.

Scrambled control. Hollnegal and Woods (2005) describe this control mode as "a blind trial-and-error performance." The actions of the operator demonstrate little understanding of the system and can appear to be completely random. The operator does not understand the situation and does not have enough time to.

In this experiment, a participant would be demonstrating scrambled control if their control actions appear random and lead to less control. Figure 28 shows an example of what the altitude of an aircraft in scrambled control may look like. The red circles are instances where the operator had an opportunity to use control actions that would bring the aircraft closer to the target glide, but instead used control actions that sent the aircraft further from the target glide slope. The flight path appears random when compared to the target glide slope and the participant eventually loses control of the aircraft.

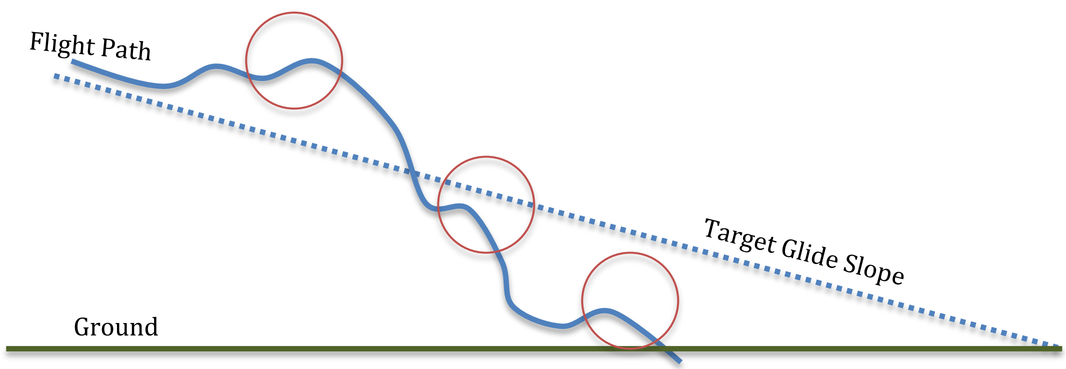

Figure 28. Example of scramble control.

Opportunistic control. This control mode displays more understanding of the situation that scrambled, but it is still incomplete. Actions are taken for their immediate outcome, while not taking into account the long-term effects, which often results in inefficiencies (Hollnagel and Woods, 2005).

In this experiment, pilot induced oscillations (PIOs) would be evidence of opportunistic control. A PIO is when repeated overshoots of similar magnitude occur. The participant understands what immediate action is necessary to return to the target glide slope, but overcompensates since they do not fully understand or appreciate the delayed effect of their control inputs due to system latency. Figure 29 shows an example of what the altitude of an aircraft in opportunistic control may look like. Notice when there is a deviation, the flight path does return to the glide slope, but it overshoots the glide slope due to the pilot's lack of understanding that control inputs have a delayed effect when there is additional system latency.

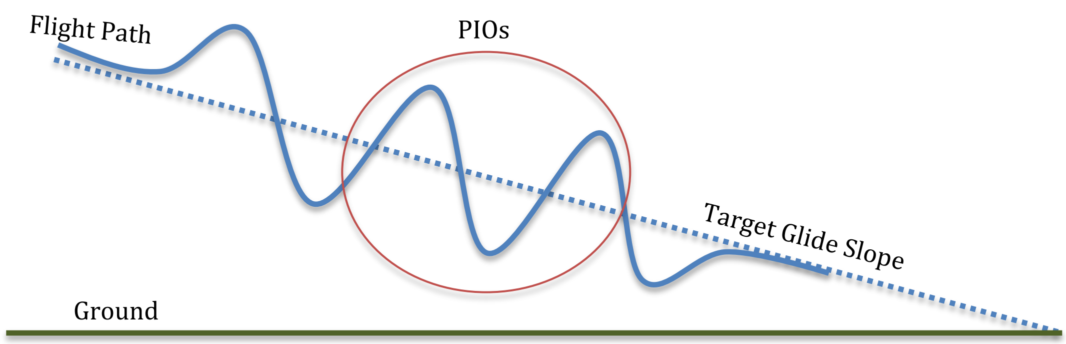

Figure 29. Example of opportunistic control.

Tactical control. This control mode will take some delayed effects into account. Whereas opportunistic control is only concerned with immediate needs (e.g., aircraft is above the target glide slope, the pilot descends), the tactical control timeline for planning extends further into the future (e.g., aircraft is above target glide slope, pilot descends, but reduces the rate of descent as the aircraft approaches target glide slope) (Hollnagel and Woods, 2005).

In this experiment, a participant who is able to compensate for system latency and minimize overshoot when reacting to disturbances would be exhibiting tactical control. This may look like PIOs in frequency, but the amplitude would diminish. Figure 30 shows an example of what the altitude of an aircraft in tactical control may look like. Notice that some overshoot does occur, but that it rapidly diminishes since the pilot understands the delayed effect of control inputs.

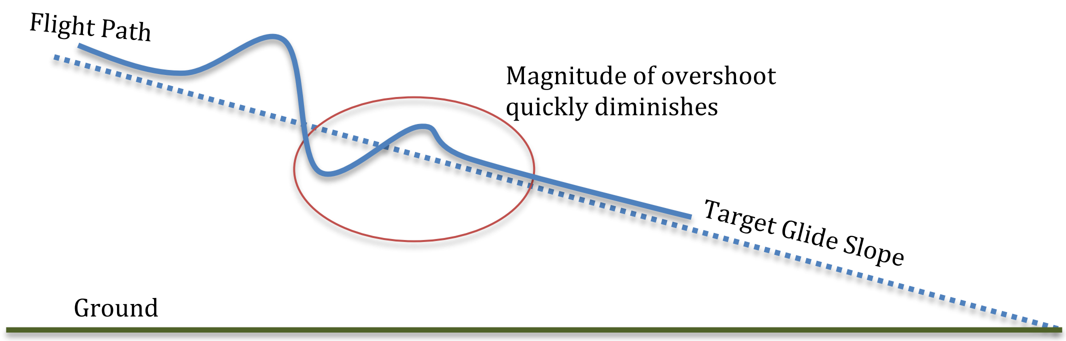

Figure 30. Example of tactical control.

Strategic control. This control mode takes into account higher-level goals and the horizon for planning extends even further into the future than that of tactical control. This mode considers the interaction between multiple goals and manages them in order to obtain the most efficient and effective use of resources (Hollnagel and Woods, 2005).

In this experiment, correcting for disturbances without causing any overshoot would be considered evidence of strategic control. Figure 31 shows an example of what the altitude of an aircraft in strategic control may look like. This would also have to be compared to other performance parameters to see if the other goals (e.g., desired airspeed, aircraft alignment with the runway) were being managed as well. If the altitude profile for a particular approach was similar to Figure 31, but the alignment with the runway fluctuated randomly (i.e., scrambled control), then it cannot be said that that particular approach demonstrated strategic control.

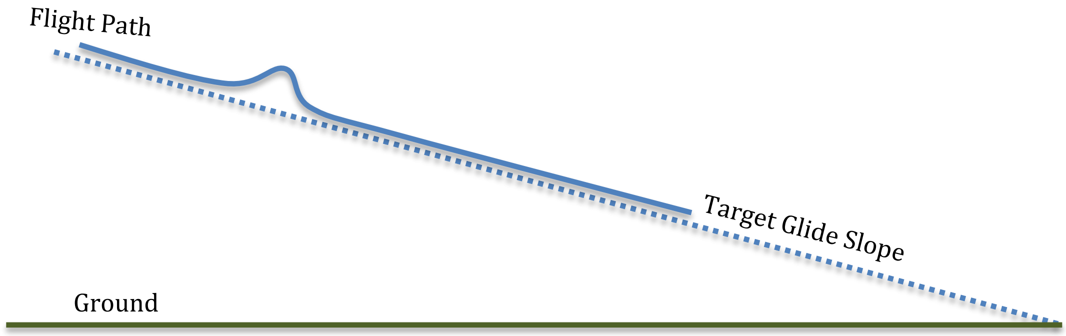

Figure 31. Example of strategic control.

Strategic versus tactical control. In COCOM, features that distinguish strategic from tactical control are number of goals, understanding of dependencies between actions, and the planning horizon (Hollnagel & Woods, 2005). Strategic control occurs when many goals are managed efficiently such that long-term effects of actions are accounted for and incorporated into the execution of an efficient plan. Distinguishing between strategic and tactical control in the context of a 2-min landing approach is exceptionally difficult. The pilot ultimately has a single goal (safely landing the aircraft), which can at most be decomposed into a few closely related sub-goals (airspeed, altitude, alignment, angle of attack). Furthermore, the timeframe is limited in that there are no truly long-term goals or effects beyond the 2 minutes. Distinguishing between these two control modes would involve altering the task and or time consuming knowledge elicitation techniques. Because of this, it was decided to use the operational definitions that are outlined above. For the purposes of this experiment, these operational definitions allow coders the opportunity to efficiently distinguish between two levels of control that are on the high end of the control continuum. This approach is considered reasonable since this research is more concerned with the occurrence of opportunistic and scrambled control than the distinction between strategic and tactical control.

Coding

3. The coding task

The coder will be handed the vertical and horizontal flight profile of an approach and will evaluate and assign control mode codes to the approach based on the definitions above. The coder may assign as many codes to a single profile as they see fit. Consider both the vertical and horizontal profiles as they evaluate an approach and assign the lowest control mode. For example, the horizontal profile in Figure 32 would suggest Strategic and Tactical Control, however the vertical profile in Figure 33 shows evidence of Opportunistic and Tactical Control. In this case the evidence of lesser control modes in the vertical profile overrules the lack of evidence of similar control modes in the horizontal profile.

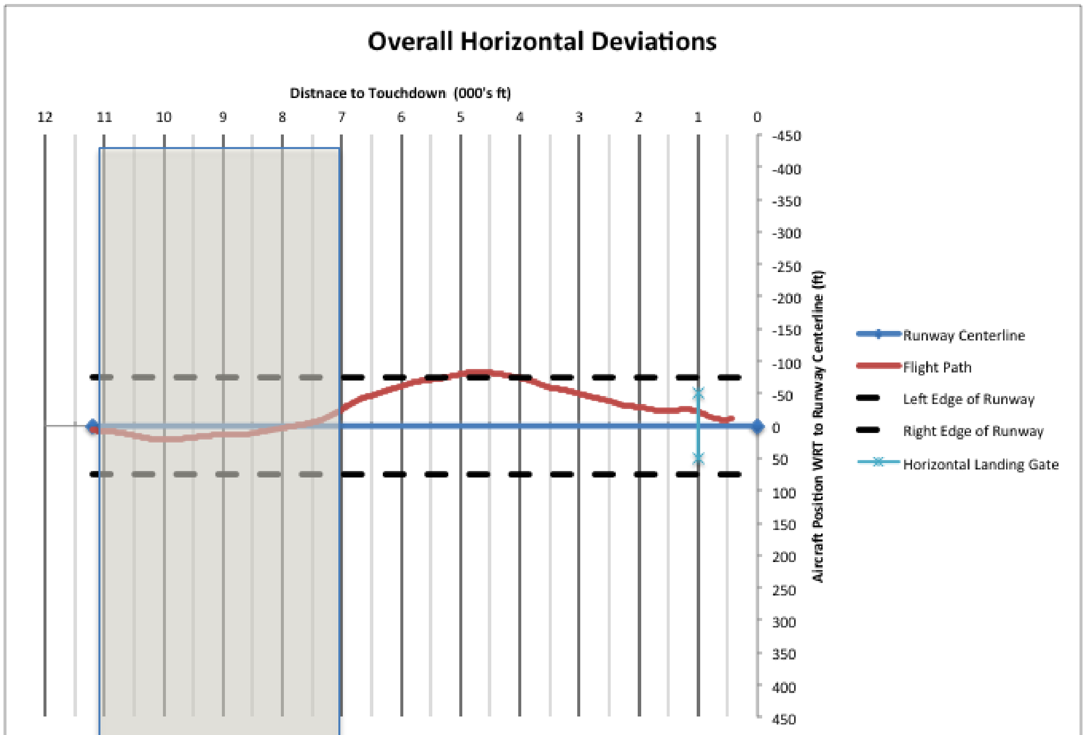

Figure 32. Example of horizontal profile.

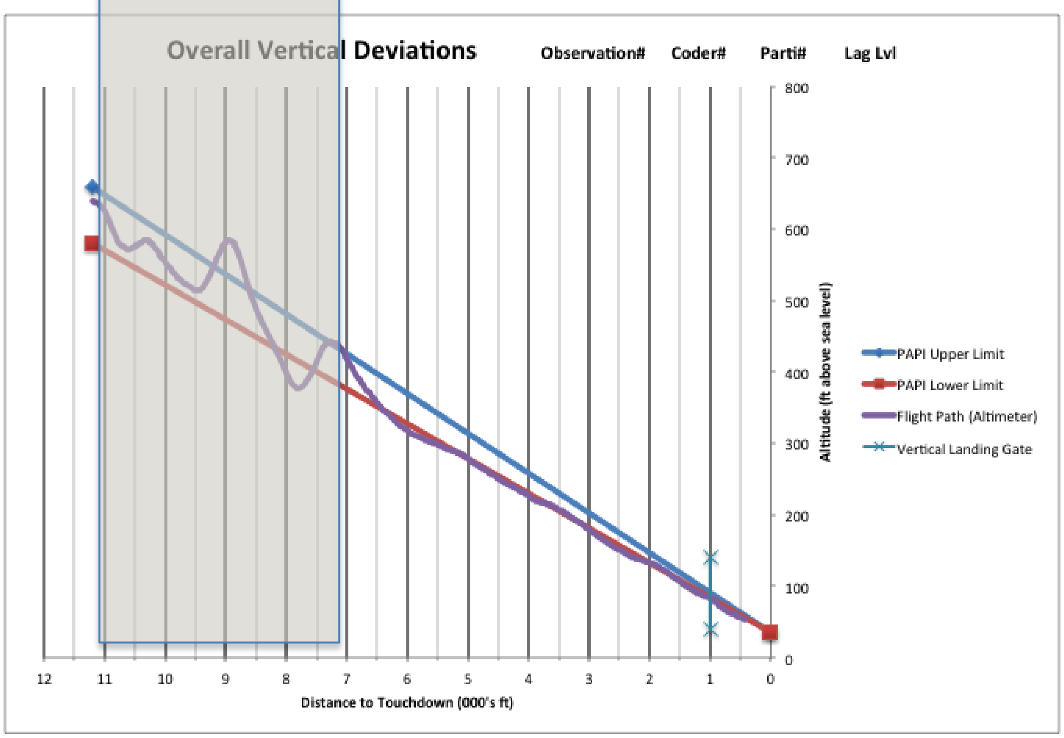

Figure 33. Example of vertical profile.

As mentioned, the approach can be divided up as the coder sees fit. In the case of Figure 32 and Figure 33, from 11000 ft to 7500 ft the occurrence of PIOs indicate Opportunistic Control. The section from 7500 ft to 4000 ft indicates Tactical Control since there are overshoots, but their magnitude rapidly diminishes. The final 4000 ft suggests Strategic control since there are only minor deviations from the ideal.

Figure 34 and Figure 35 demonstrate another instance where the lesser control mode overrules what could be seen as a higher control mode. The vertical profile might suggest Tactical Control from 7000 ft to 5000 ft, but the horizontal profile clearly shows that it is Scrambled and Opportunistic Control.

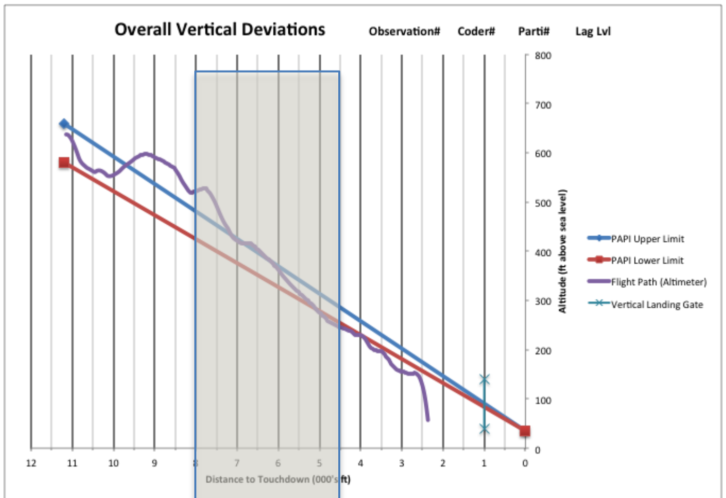

Figure 34. Supplementary vertical profile example.

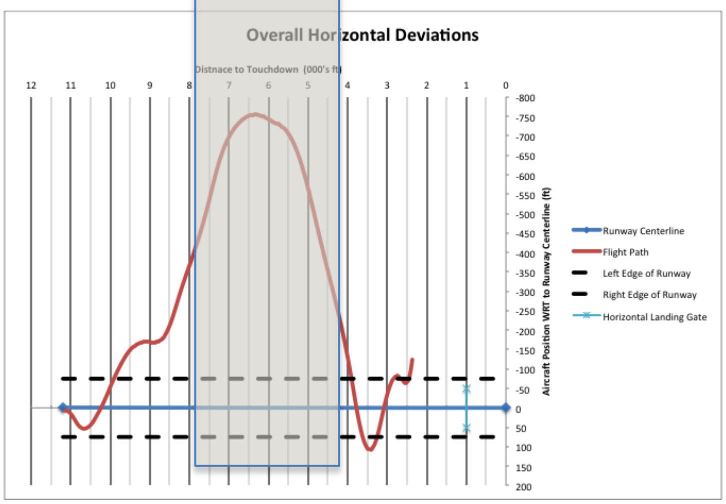

Figure 35. Supplementary horizontal profile example.

A second coding task will be to assign a control mode code to the final 2000 ft of the approach. In this instance, the coder must choose one, and only one control mode code that best describes the final 2000 ft. If there is evidence of two different control modes, then the coder will assign the control mode that makes up the majority of the final 2000 ft. This is considered an unlikely situation because 2000 ft prior to touchdown is a limited window of opportunity for multiple control modes to be present. Figure 36 and Figure 37 are examples of final 2000 ft profiles that coders will be asked to evaluate.

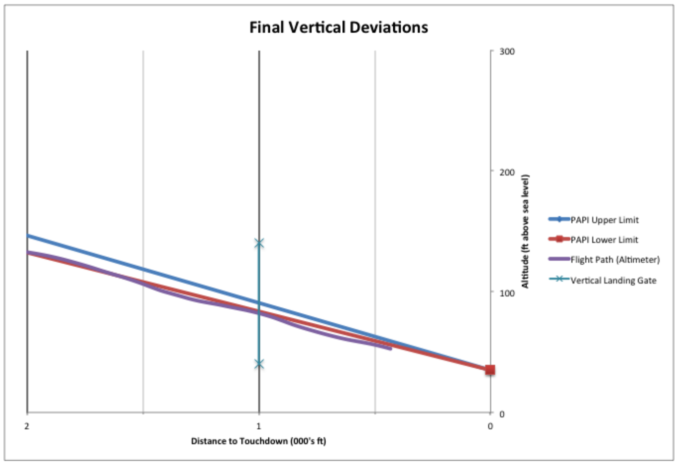

Figure 36. Landing control vertical profile example.

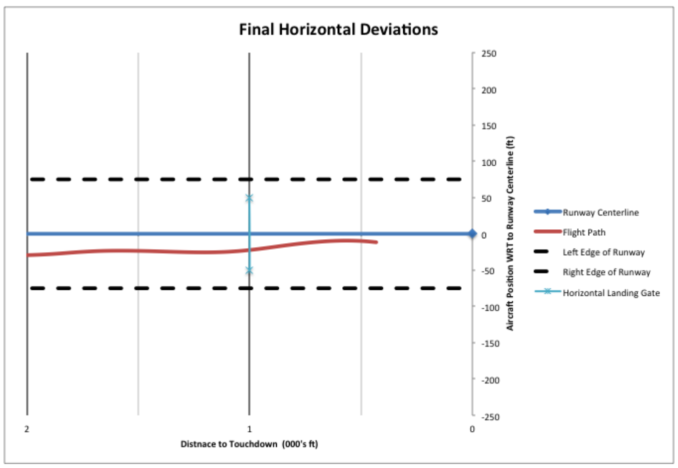

Figure 37. Landing control horizontal profile example.

Note that the flights paths end at 400 ft. This is when the simulation ended during the experiment. Also, if no flight path is shown, then that means that the aircraft was nowhere near the target glideslope for the final 2000 ft. In this instance, mark the approach as "failed."

4. Miscellaneous notes

Scale. The full flight profile is 10,000 ft yet it is presented on an 8.5x11" sheet of paper with only 800 ft of altitude. This means that a modest change in trajectory can appear extreme, as shown in Figure 38. This is not to discount what appear to be extreme trajectory changes; these changes are an important aspect of coding, however if a certain flight path appears insane, it is partially due the scale's amplifying effect.

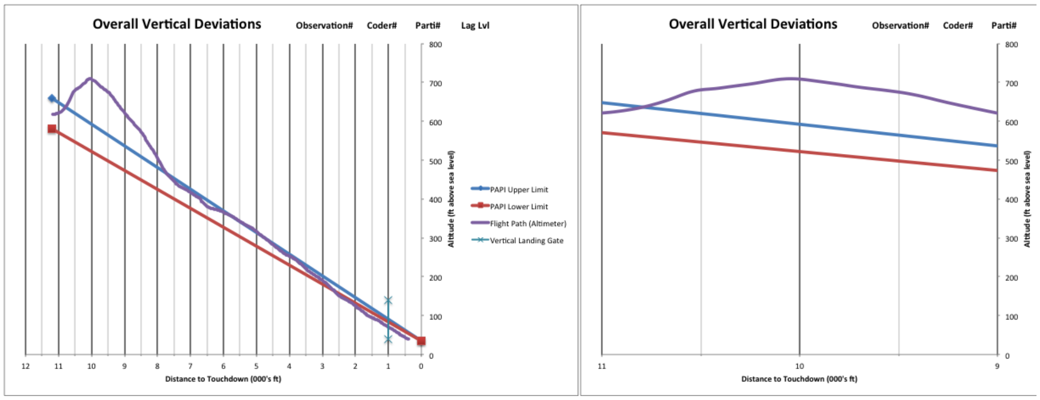

Figure 38. Different scales for the same approach data.

Trajectory. It is important to consider the trajectory of the flight path with respect to the target glide path and not simply the magnitude of the deviation. A large deviation that is progressively getting smaller is evidence of a pilot who is in control and attempting to correct the flight path. Consider Figure 39, the aircraft begins well off of the target glide slope, yet the trajectory is steadily converging with the target glideslope. This would indicate a higher level of control such as Tactical or Strategic.

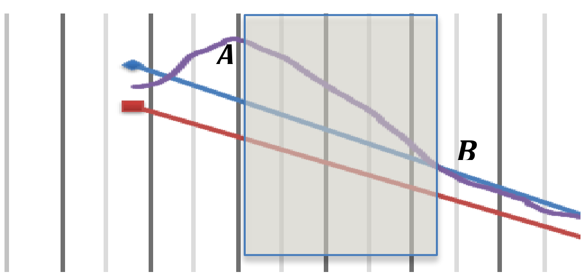

Figure 39. An improving deviation.

Trajectory changes. An important aspect of the COCOM is the operator's understanding of the system. It is difficult to know what a pilot is thinking and whether they understand the system, but changes in aircraft trajectory can provide some insight. For instance, at point A in Figure 39 the pilot changes from a trajectory that takes the aircraft away from the target slope to one that brings the aircraft closer to the target slope. This suggests that the pilot has some understanding of how the system works and it is unlikely that they are using trial and error (thus, not demonstrating scrambled control). Point B is another trajectory change that suggests the pilot understands how the system works. Furthermore, the absence of overshoot (and PIOs) and point B suggest an even higher level of understand and control (tactical or strategic).

Coders are to use trajectory changes in conjunction with the trajectory in order to refine their coding. The trajectory in Figure 39 suggest Tactical or Strategic control. Examining how that trajectory began and eventually changed (apparent and distinct change in pilot control inputs with virtually no overshoot once the target was achieved) is evidence of Strategic control.

PIOs. Identifying PIOs is critical to identifying the control mode. As with all oscillations, the chief characteristics are magnitude and frequency. If the flight path appears to be a wave, this could indicate a PIO. However, when the frequency is low (i.e., large wavelength/a really long wave), Figure 40, then the number of overshoots is minimal and the appearance of a wave is only minor corrections over a long distance.

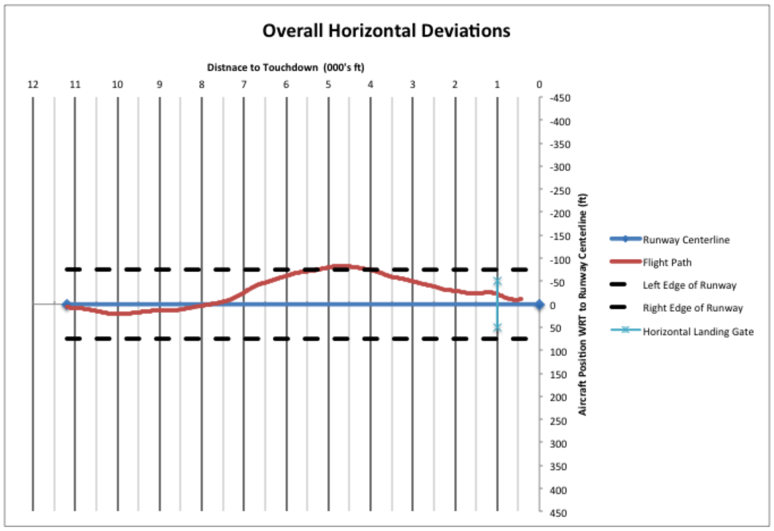

Figure 40. A low frequency deviation is not strong evidence of a PIO.

A single wavelength over the entire approach is not a PIO, however a wavelength of approximately one third of the approach (~3,500 ft) or smaller is a strong indication of a PIO. The PIO in Figure 40 has a wavelength of over 10,000 ft in that it takes 10,000 ft for it to dip below the target path, overshoot, and then return to the target path. This is not a strong indication of a PIO.

Likewise, if there are small amplitude waves, Figure 41, then the magnitude of the overshooting is small. These are only minor corrections are not a strong indication of PIOs. Generally speaking, if the wave height (distance from the bottom/dip to the top/peak), if less than 50 ft, then that is not a strong indication of a PIO.

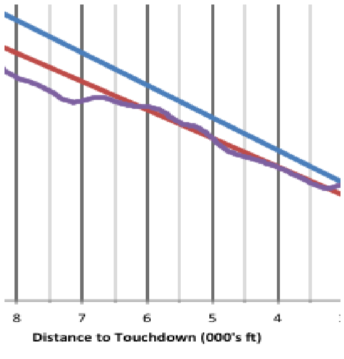

Figure 41. Small magnitude overshoot.

When identifying PIOs, it's important to consider both wave-length and amplitude. Figure 42 shows a flight path with a large wavelength, which might suggest that there aren't any PIOs, however the amplitude is very large, which does suggest PIOs. Furthermore, at points C and D are trajectory changes that suggest the pilot does not understand the system (i.e., taking "corrective" action that is in fact making the situation worse).

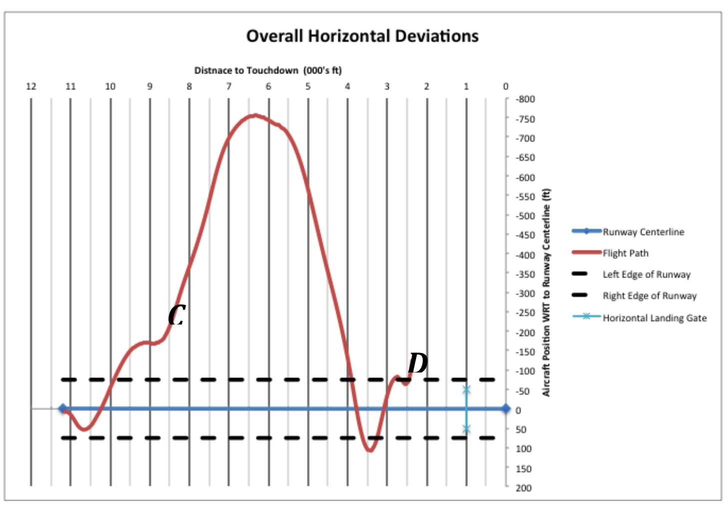

Figure 42. Large wavelength and amplitude.

Partitioning. It is possible (and likely) that the pilot may demonstrate varying control modes over the course of an approach. For this reason, when evaluating the "Entire Approach" (as opposed to the "Landing Control"), the coder is free to partition the approach and assign different control mode codes to different parts of the approach (though only one for any given section). While the coder is free to partition the approach, they should not do so to the extent that they create miniscule sections. Generally speaking, the approach should not be partitioned into sections smaller than 2000 ft, though if there is compelling evidence of different control modes then sections can be as small as 1000 ft. No sections will be smaller than 1000 ft.

Initial aircraft trajectory. The simulations began with the aircraft traveling parallel to the ground (as opposed to along the glide slope) and with a nose-up trim that causes the aircraft to climb. The result of this is that most approaches begin with the aircraft climbing and deviating from the glide path.

Missing data. For some approaches, the flight path may abruptly end. This is an aircraft crash, which is a strong indication that the control mode was Scrambled or Opportunistic just prior to the crash.

5. Final thoughts

If you are having trouble coding, refer back to the original theory and definitions to help guide your evaluation. You can also use your previously coded approaches as a reference. This will help you remain consistent when coding. Below is a quick guide that should help if you keep it handy as you code.

Control Mode Codes Quick Guide

Item	Description	Tips
Scrambled	-random trial and error -operator does not understand the system	-look for actions that don't make sense, eg: aircraft is too low, but the pilot alters the flight path by pointing the nose down
Opportunistic	-inefficient -operator has limited understanding of the system	-the pilot tries to correct for deviations (when aircraft is too low, pull up), but doesn't understand how to handle delay -inefficient, aircraft frequently overshoots the target glideslope (Pilot Induced Oscillations)
Tactical	-effective but somewhat ad hoc -operator understands and considers some delayed effects	-pilot corrects deviations and somewhat takes into account delayed effects, as demonstrated by diminishing PIOs -there are overshoots, but they get smaller
Strategic	-high level of control -operator completely understands how to control the system	- no PIOs, deviations are corrected with virtually no overshoot after a flight path correction and flight path closely reflects the target path

Coding Task Guidelines

Item	Guideline
Evidence	-Use both the horizontal and vertical deviations. When they suggest different control modes, the evidence of a lower control mode outweighs the lack of evidence of a lower control mode
Trajectory	-Is the trajectory similar to the target? Is it slightly different in order to reduce deviations? These are indications of higher control modes, even if there are noticeable deviations.
Trajectory Changes	-Use these to help refine coding -Is the change of the better? Is the change well timed (i.e., at the correct moment)? These are indications of higher control modes
PIOs	-Typified by overshoots and the appearance of "waves" in the flight path
PIO wavelength and amplitude	-Wavelength of more than ~3,500 ft is not a strong indication of a PIO, less than that suggests PIOs -Amplitude less than 50 ft is not a strong indication of PIO, greater than 50 ft suggests PIOs -Wavelength and Amplitude need to be considered together, so a wavelength that is large can be a PIO if the amplitude is also large
Partitioning	-Only create sections less than 2000 ft when there is compelling evidence and do not create sections smaller than 1000 ft

Appendix B
Lag in UAS Control

Conducted by Marshall Lloyd
Advisor: Dr. Neville
Embry-Riddle Aeronautical University
Daytona Beach, FL, 32114

The experiment you are volunteering to take part in is investigating the effect of lag on simulator performance. Lag in this experiment is a delay in control response to pilot inputs in the simulator. The results of this study are to be applied to Unmanned Air Vehicles (UASs), also known as drone aircraft. UASs are remotely piloted by operators that can be thousands of miles away. This can cause a delay between an operator's command and when the aircraft executes the command.This experiment is investigating how much this lag can affect performance on landing.

Your involvement in this study will consist of three simulator sessions. Each will be approximately 45 minutes and they will take place on different days. You will receive $5 during your first two sessions for a total of $10 as compensation for your participation. A prize of $100 will be awarded to the participant with best performance and $20 will be awarded for the second best performance.

During the experiment, there is a slight possibility that you may experience simulator sickness, which is similar to motion sickness. You will be closely monitored and tested for symptoms. In the unlikely event that you experience simulator sickness, the experiment will discontinue. You are free to discontinue you participation at any time. The results collected in this study will remain anonymous and your name will not be published. You may contact myself, Marshall Lloyd 386-872-0066, or my supervisor, Dr. Neville 407-461-1277, at anytime if you have any questions or concerns. You will be provided with a copy of this form. When the study is complete, you will receive a debrief about the results and a copy of the report.

Statement of Consent

I am an informed participant of this experiment. I have read the above information and have asked any questions that I may have had about my involvement in the experiment and the experiment in general. I have been informed of the purpose of this experiment and I am aware that I will receive $5 during the first two sessions. I am aware that I am free to leave the experiment at any time, but doing so will eliminate me from contention for the $100 first place and $20 second place prizes. I am aware that the data collected on my performance in this experiment will not be associated with my name in the publication of results.

Participants Name (please print):________________________________

Signature of Participant:____ Date:_________

Signature of Experimenter:_______________________________________ Date:_________

Appendix C

Detailed Procedure and Plugin Instructions

Turn on *Da* computer, start NetDisturb, select level of delay, and click on "*Run All*." Turn on three X-Plane computers (Mr, Iy, and Pt). Ensure that NetDisturb is running before turning on the other computers. Install the plugin on Mr by copying *XWindStudy.xpl* and *XWindStudy.cfg* into the X-Plane plugin folder, which can be found in *…XPlane\Resources\plugins*. Open the *.cfg* file using notepad and enter desired experiment parameters, refer to Table 14 for instructions and Figure 43 for a screenshot of the plugin settings used for this study.

Table 14

Plugin Configuration Instructions

Line	Instructions
DataLoggingPath	-Set this variable to the path of the folder that will contain the data files for each participant. If the folder does not exist, it will be created.
DataLoggingRate_Hz	-Set this to the desired data recording rate (in Hz).
Lat and *Long*	-Initial aircraft GPS position. -Note that KDAB runway 34 centerline axis is defined by y=(-2.08164890391)x - 139.53994988802
Alt, IAS, and *Heading*	-Remaining initial aircraft conditions. -Alt is measured in feet, IAS in Knots, and heading in degrees.
HeadWind	-The headwind for the entire approach in knots.
GlidePathAngle_Hi and *GlidePathAngle_Lo*	-The glideslope from which vertical deviations will be measured. This is in degrees. -Set the upper and lower boundaries of the PAPI.
StopDistance_ft	-The simulation termination condition -The simulation is paused and data recording stops when the aircraft is this far from the touchdown markers.
DistanceFromRwy_nm	-Crosswind timing for left or right crosswind. -The adverse crosswind will be triggered at this distance from the runway touchdown point. -This is measured in nautical miles.
CrossWind	-Magnitude of the crosswind in knots.
Duration_ms	-Duration of the crosswind in ms.
Lat, Long, and Elev	-Targeted touchdown point on the runway, adjacent to the PAPI lights. This is used as a reference point to determine deviations. -The default is for KDAB runway 34, however it can be changed in the X-Plane plugin menu. Move the aircraft to the desired location and select *XWindStudy->Configure->Set Runway Begin Point*
Lat, Long and Elev.	-End of runway. Serves a similar purpose as Runway Begin and can be modified in the same way.

XWindStudy.cfg

```
[XWindStudy]
DataLoggingPath=C:\Documents and Settings\BOEING\Desktop\Participant Data
DataLoggingRate_Hz=4.000000

[ScenarioBegin]
Lat=29.1494645331
Lon=-81.0364390000
Alt=638.890000
IAS=75.000000
Heading=340.000000

[Common]
HeadWind=5.000000
GlidePathAngle_Hi=3.210000
GlidePathAngle_Lo=2.810000
StopDistance_ft=400.000000

[AdverseLeft]
DistanceFromRwy_nm=1.600000
CrossWind=20.000000
Duration_ms=2000.000000

[AdverseRight]
DistanceFromRwy_nm=1.600000
CrossWind=20.000000
Duration_ms=2000.000000

[RunwayBegin]
Lat=29.1777633719
Lon=-81.0500334341
Elev=30.7624426825

[RunwayEnd]
Lat=29.1903982249
Lon=-81.0561030710
Elev=29.5632965078
```

Figure 43. Plugin configuration file screenshot.

Start X-Plane on Mr, Iy, and Pt. Once X-Plane is running on all three computers, export Mr's display to Iy in X-Plane's *Net Connections*. Here you will specify Iy's IP address, as shown in Figure 44. Now Iy's Inet 2 page in Net Connections should look like Figure 45.

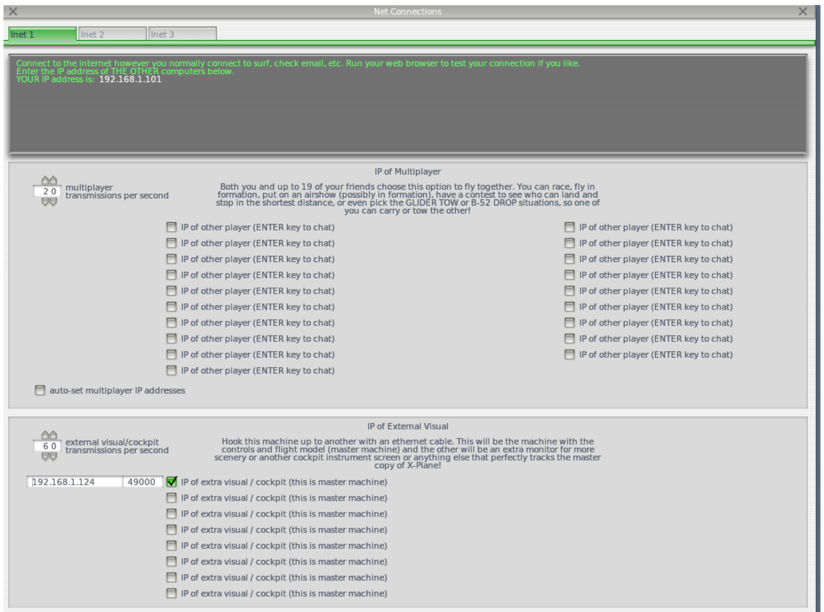

Figure 44. Screenshot of *Net Connections* on exporting display.

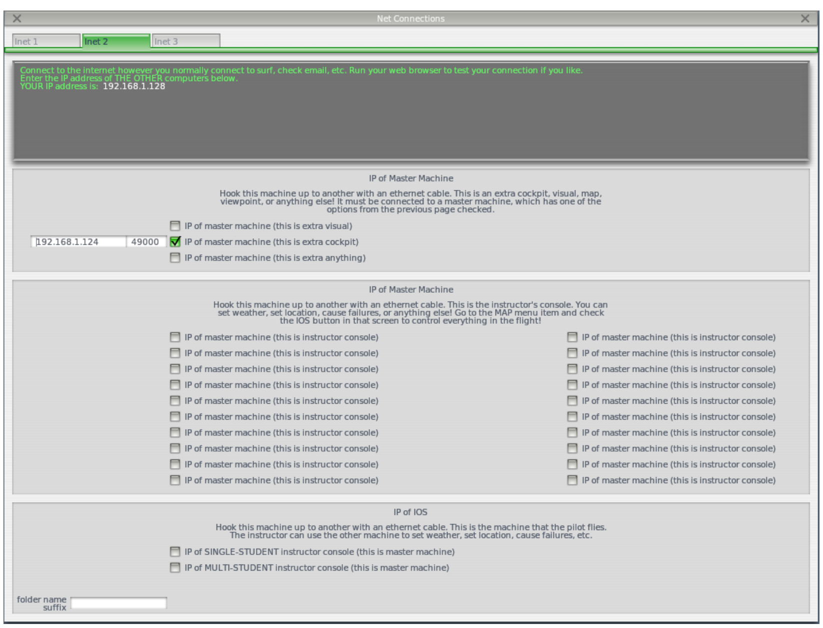

Figure 45. Screenshot of *Net Connections* on recipient display.

Export Iy's display to Pt using the same procedure. The display on all three X-Plane computers should now look the same. Input controls on the Mr yoke and visually check that the monitors respond in unison while paying attention to the amount of delay.

Welcome the participant and have them complete the appropriate paperwork. Read the pre-brief script. Have the participant sit at the participant station. From the experimenter station:

- pause X-Plane;
- start a new participant data file by using *Plugin>XWindStudy>Control>New Participant* and entering participant information, then close the window using *X*;
- bring the aircraft to a 3 NM approach using *Position>Local Map>Runway 34 3nm,* as shown in Figure 46;

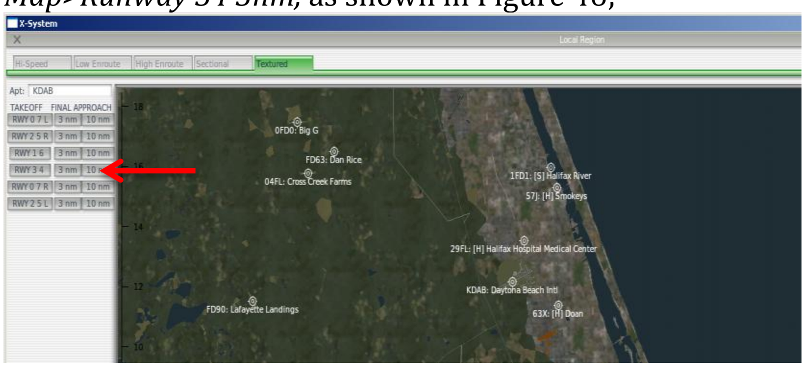

Figure 46. Moving aircraft to approach position.

- instruct the participant to not touch any controls as you un-pause the simulation. This is to zero the controls;
- ready the plugin by opening *Plugin>XWindStudy>Control*
- un-pause the simulation for a moment and then click on the desired scenario (*Ideal, Adverse Left, or Adverse Right*);

- move the aircraft to 618 ft using *Position>Local Map>Altitude*. This levels the aircraft and has it start at the same altitude;
- instruct the participant that you are about to un-pause the simulation and that this time, they are to take control;
- un-pause the simulation and observe as the participant conducts the approach;
- the simulation will automatically pause and stop recording near the touchdown markers, however if the aircraft crashes, pause X-Plane and then stop recording data by using *Plugin>XWindStudy>Control>Stop Recording*; and
- bring the aircraft to a 3 NM approach again and repeat previous steps for the next trial.

If the participant crashes, reassure them that there are very difficult settings and that they are to try their best. After two approaches, administer the SSQ and stop the experiment if there are signs of simulator sickness. Administer the SSQ again at the end of the session. End the session by conducting the post experiment interview and recording participant responses.

Appendix D

Participant Pre-Brief Script

-Give them $10 and have them sign the Informed Consent form if they haven't already.

-"Thank you, this is a script I read to every participant."

-"This experiment simulates the lag that UAS operators can experience, though sometimes you will fly in a normal aircraft configuration. You will experience the lag condition during some of your sessions."

-"Sometimes lag can induce simulator sickness, which is like motion sickness, but less severe. I will have you conduct two approaches and test you, then again at the end of the experiment. If you are feeling sick, let me know and we will stop. Please take a look at the questionnaire and let me know if you are feeling any of those symptoms right now."

-"You will conduct a series of approaches, and the simulation will end right before touchdown."

-"Stay in line with the center of the runway, and use the PAPI to glide your glide slope. Your goals are:

- -stay aligned with the centerline,
- -maintain a glide slope according to the PAPI (two white lights)
- -hit the runway touchdown point next to the PAPI
- -land with airspeed around 60-65knots.
- -Note that the runway is roughly 38ft above sea level, so that is what the altimeter reads on the ground at the touchdown markers."

-Show the participant the controls, flaps, throttle, landing gear (fixed), rudder, and column. Inform them that cockpit view is fixed and that they will not be able to trim the aircraft.

-"You start at 618ft, with 75KIAS, and heading towards the runway. There will always be a 5 knot headwind and there is a chance of crosswind gusts during the approach."

-Explain the experimental procedure. Let them know that there will be several warm ups, that everyone has trouble at certain settings and that they should simply try their best.

-"Any questions?"

Appendix E

Participant Interview Synthesis

Table 15, Table 16, and Table 17 are summaries of the participant responses to the questions:

1. How do you think you performed during this session?

2. Were these approaches difficult?

3. Did you change how you fly the aircraft in order to compensate for the system settings?

Some cells in the *responses to questions* column have a fourth line item. This is simply interesting responses or observations that do not fit in with any of the other questions. Participant RMSE for each lag condition is also included in this column.

Table 15

No Lag Conditions Feedback

Participant Experience	Participant Response to Questions
810 hrs	1. Performed alright, didn't maintain the right glideslope and often ended high. 2. No 3. Had to counter the pitch up tendency. Would set power and flaps at the start. RMSE: 51ft
350 hrs	1. Alright, not as good as regular flying. 2. No. 3. Yes, used full power and flaps. RMSE: 39ft
16,000 hrs	1. Better than level 1 lag. Steep learning curve on the simulator. 2. Somewhat due to lack of trim, having to nose down, and seeing the PAPI from a distance. 3. Yes, learning the sensitivity and tricks to the simulator including always applying forward pressure to get to neutral. RMSE: 32ft
1,500 hrs	1. Okay, same as the low lag. 2. More difficult than a real AC, no AC feel and limited view. 3. No RMSE: 48ft
2,200 hrs	1. Good. 2. No. 3. No. RMSE: 27ft
254hrs	1. Much better than the high lag setting. Didn't feel lag in the control, AC was more responsive. 2. Not really. 3. Due to nose up trim, had to nose down right away. RMSE: 64ft
750 hrs	1. Pretty well after warm up. 2. No 3. Found out what worked with the yoke: yoke forward is neutral. 4. Requested trim capability. RMSE: 19ft

230 hrs	1. Okay, not as on target as would like to be. 2. Very easy and wind wasn't a challenge. 3. Used flaps the entire time, but was pretty much how they always fly. RMSE: 37ft
415hrs	1. Better than others. 2. No. 3. No. RMSE: 34
600 hrs	1. Best performance. 2. No. 3. Yes, the controls were very sensitive so used smaller movements. RMSE: 52ft
450 hrs	1. Well. 2. No. 3. No. RMSE: 88ft
130 hrs	1. Not the as well as flying an actual AC, but better than the high lag condition. 2. The nose up trim was tricky. 3. Flew most approaches a little high due to the nose up trim of the AC. Used the flight instruments more than they normally would have. RMSE: 35ft
13,000 hrs	1. Alright, could manage to handle the simulator settings. 2. No. 3. The warm ups helped adapt to the simulator. Had to adapt to the round dials. RMSE: 124ft

Table 16

Low (240 ms) Lag Conditions Feedback

Participant Experience	Participant Response to Questions
810 hrs	1. Performed horribly, missed the centerline and glidepath, especially on short final. 2. Yes, a bit. Felt prone to PIOs, especially on short. Pulling back (flaring) was hard to time correctly. 3. Put in a correction and wait to see effect. 4. Eventually figured out there was a delay, but kept talking about a difficult crosswind and shear well after the weather event. RMSE: 62ft
350 hrs	1. Fine. 2. No. 3. No. RMSE: 40 ft
16,000 hrs	1. Took a little while to get the feel of the system. PAPI was hard to see at first (had glasses). 2. Yes, rudder and yoke were difficult to use (sensitive). Hard to do nose down corrections because horizon was lost. Found control column more difficult than the lag. 3. Yes, lack of trim was a challenge (had to constantly force nose down). Tried to anticipate lag, but found more success by flying what they saw. Felt like an unresponsive glider. RMSE: 68ft
1,500 hrs	1. Pretty good. 2. Difficult at the very end with the wind shear (though there wasn't any shear, he must be referring to the lag). 3. No, flew just as he would a real AC. 4. Throttle distance to power response ratio isn't the same as a real 172. RMSE: 48ft
2,200 hrs	1. Good 2. No 3. The spring force feedback on the yoke pushed to a non neutral position. Had to exert force to get to neutral and couldn't easily find and/or maintain it. RMSE: 52ft

254hrs	1. Middle performance, noticed some lag in the controls. 2. Wind shear was a bit of a problem. The last 1000 ft were the most difficult, the result of any input was intense. 3. Yes, used small corrections and took feet off of the rudder because controls are very sensitive. RMSE: 63ft
750 hrs	1. Worse, but decent. 2. Yes, a lot more concentration and effort. 3. Used small inputs, would wait to see the effect. Line up the centerline right away and then worry about glide slope. Point nose down right away to accommodate the nose up tendency. RMSE: 25ft
230 hrs	1. Better than high lag. 2. No. 3. Not really, flown like a normal approach. RMSE: 54ft
415hrs	1. Okay, but not great. 2. No. 3. Yes, smoothed out inputs. Wait and see the effect before making next move. RMSE: 60 ft
600 hrs	1. Poorly, erratic controls. 2. No. 3. Couldn't quite tell exactly what was off, so just flew normal and was aware of the sensitive controls. 4. Noticed nose up trim. RMSE: 84ft
450 hrs	1. Better than high lag, but worse than a normal flight or sim. 2. Yes, it required a little extra attention and thought. 3. Flew normal, related the situations to normal aircraft flight. Was able to pay attention to speed. RMSE: 95ft
130 hrs	1. Seemed the same feel and performance as the last (no lag) setting. 2. Not entirely difficult. Felt they did a good job of maintaining the correct glideslope. Avoiding oscillations on short final was a bit of a challenge. 3. Avoided pitching down on short final because it resulted in oscillations during the warm ups. Didn't use the rudder, partially due to the fact that it was a simulator.

	RMSE: 60 ft
13,000 hrs	1. Good 2. No. 3. Same as low lag 4. Wind shear is usually for and aft, not left and right. RMSE: 164ft

Table 17

High (1000 ms) Lag Conditions Feedback

Participant Experience	Participant Response to Questions
810 hrs	1. Started completely horrible, but then less so. 2. Yes, it was unfamiliar and can't anticipate system behavior/reaction. Hare to flare the AC, cannot judge timing with delay. 3. Avoided over controlling, used very small inputs (put it in, take it out, and see what happens). Used rudders to line up with centerline – ailerons (yoke) resulted in loss of control. 4. Short final is very difficult, it was stressful and frustrating. Thought there was a bunch of wind shear. RMSE: 89ft
350 hrs	1. Felt like spirit was crushed. Very frustrating. 2. Yes, very difficult. Couldn't accommodate the lag. 3. Yes, stopped using/trusting the screen. Kept the screen in the periphery, but paid attention to how far the controls were moved. RMSE: 164ft
16,000 hrs	1. Not as good as wanted to. Thought there would be a more steady improvement. 2. Yes, the lag (accumulation of inputs without seeing an effect) and the neutral nose up tendency. Any more than 30deg bank and they would just "lose it." 3. Tried different techniques (ailerons vs. rudder), had some luck with "cheating with the rudder" by using it to correct horizontal deviations. Did some "wait and see" but this wasn't possible at lower altitudes. Used smaller corrections and "just flew" – avoided thinking about it. 4. "It's nice to finally find the centerline" (after many misses). Accepted being off of the ideal glide path because they had learned that major corrections resulted in loss of control ("losing it"). RMSE: 126ft
1,500 hrs	1. Terribly, lots of crashing or hard landings. 2. Short final corrections were very difficult. Not enough time to make corrections. 3. Yes, controlled using 'timing', ie: input a control for a short period of time (as opposed to until you see it lining up. Gusts weren't a big deal, it's a stable AC and the gusts occur at a high altitude. RMSE: 117ft

2,200 hrs	1. Awful, very frustrated. Thought they would learn to accommodate the lag more quickly. 2. Awareness of AC speed went down, though it didn't really vary that much. 3. Yes, more focus on the centerline right away. Small inputs to avoid PIOs. Correcting to the nose up trim (by pointing nose down) was difficult because they would lose the horizon and the coarsely pixilated ground wasn't helpful. 4. Spring force feedback was unusual. RMSE: 101ft
254hrs	1. Medium performance, and very frustrating. 2. Not after warming up. Lag and the lack of feedback were annoying. The wind shear caused trouble. The last couple hundred feet were near impossible if not lined up, can't correct. 3. Yes, after realizing lag. Set power at the beginning and used small corrections. Tried to stay ahead/on top of deviations. Used attitude (push on yoke) from previous approach instead of referring to the monitor. RMSE: 88ft
750 hrs	1. Poorly, felt like a student pilot all over again. 2. Very difficult, couldn't get it to stabilize. Recovery from adjustments was impossible. 3. Yes, used very small and smooth control movement. Used timing instead of the monitor to regulate inputs. Input control and then wait. RMSE: 96ft
230 hrs	1. Definitely worse, though felt improvement throughout. 2. Yes, tough to line up. Had to go against instincts. 3. Stopped doing inputs before lining up. Relying more on memory (were to hold controls) than technique (normal flying reactions). Set power at the start based on other approaches. Used the rudder a lot more for lateral corrections. RMSE: 89ft
415hrs	1. Horrible, lots of crashes. 2. Yes, the plane is unresponsive. Could tell the "landing" was going to be a crash well before getting close (eg: if it needed sizable correction past a certain point). Hard to find neutral point on the yoke. 3. Yes, stopped using the rudder, concentrating on three inputs was too difficult. Set power and flaps right away and then used small inputs on the yoke. Stopped caring about perfecting the glideslope and centerline and just aimed for the runway. Started looking at the end of the runway instead of the start.

	4. During the trials said: "I'm going to lose it," and was very frustrated. RMSE: 111ft
600 hrs	1. Terrible, lots of crashing. 2. Yes, couldn't figure out how to maintain control. 3. Yes, tried to time turns and see what the effect was (wait and see). Closer to the ground, needed much finer and accurate controls. 4. Frustrating, doesn't even like normal simulators. RMSE: 250 ft
450 hrs	1. Not very well, terrible in fact. Very unsafe. 2. Challenging, felt like there was a difficult wind shear. Focus was on not crashing and lost track of AC speed. 3. Tried to figure out what worked. Started to change thought process, thought ahead of the aircraft and then flew normally. Recognized the overcorrections. 4. Especially good at finding the centerline and recovering from large horizontal deviations. RMSE: 144ft
130 hrs	1. Bad, but improved with practice. 2. Very difficult, was never sure how much bank to use. Wasn't sure if the power control was having an effect. 3. Tried to think about the lag and the delayed effect of control inputs. Didn't use the rudder because they didn't think it was helping and was worried about causing a stall. 4. Thought that there was wind (beyond the manipulation) and felt that they could land better if it wasn't there, though in fact it wasn't. RMSE: 226ft
13,000 hrs	1. Hard to handle the novel experience. 2. Tough to handle. 3. Tried to used normal techniques to correct the dutch roll. Thought the wind was causing the challenge. Extreme pitch up and down. 4. This participant simply did not understand how to handle what was going on. RMSE: 193ft

Appendix F

Detailed Discussion of Interviews

A brief interview was conducted after each session at a particular level of lag. The questions were broad and allowed participants to freely describe their experience. Many common themes emerged amongst the responses. Since the questions were broad, it is possible that these themes were experienced by participants who did not explicitly mention them. The following paragraphs discuss these themes.

Weather. Nearly half of the participants mentioned that adverse weather conditions made controlling the aircraft in the lag condition very challenging. What is surprising about this is that they mentioned weather phenomena that were not present. Some mentioned constant crosswinds while others mentioned strong and unpredictable gusts during short final, neither of which were part of any trials. The pre-session weather brief included a warning that there may be crosswind gusts during the approach and this may have led participants to believe they would experience strong weather phenomena.

System lag was a novel experience and participants seemed to have difficulty understanding its effect on aircraft handling. When the aircraft was difficult to handle during the lag conditions, they would sometimes attribute this to adverse weather conditions. Because the participants were not in an actual aircraft, but in a fixed base simulator, the "feel" of the wind was not there. They could not determine if sporadic aircraft movements were due to delayed control inputs or adverse whether phenomena, and they had a tendency to attribute part of the challenge to adverse weather. Furthermore, some participants mentioned that they used standard piloting techniques to overcome the imagined weather (e.g., using a standard piloting technique to overcome a dutch roll). It is possible that this approach by the participants constituted a negative transfer of skills, though this is difficult to determine based on the data collected.

Control strategies. In addition to standard piloting techniques to overcome perceived adverse weather, participants adopted several other techniques to overcome the challenge of system lag. The fact that a HITL controller will modify control behavior to adapt to system settings is well established (Baron & Kleinman, 1969, McRuer, Jex, 1967); however several novel techniques to overcome system lag were observed. An interesting observation was participant use of the rudder. Though this was not directly observed, several participants mentioned that they stopped using the rudder altogether. They claimed that it was too sensitive, that they were concerned it may trigger a stall, or that it simply was not helping control the aircraft. Conversely, other participants mentioned that they became increasingly dependent on the rudder. The claimed that they risked losing control if they used the yoke for horizontal deviation corrections and so instead they depended on the rudder for horizontal control. It is not clear which tactic results in better performance and this presents an opportunity for future research.

A much more common tactic, which was mentioned in the literature was the “wait and see” technique (e.g., Chang & So, 1999; Ferrel 1965; Lane et al., 2002; Watson et al., 1998). Instead of applying a control input until the aircraft was in the correct position, the participant would use a smaller control input and wait to see the effect. Participant's claimed that small, timed inputs helped with control, but that this technique was of little use on short final. This is because there is not enough time to "wait and see" during short final. This behavior can be seen in some of the flight profiles. The participant gains control of the JCS using the "wait and see" technique, but as time becomes scarcer on short final, they quickly lose control. An example of this pattern can be seen in Figure 47.

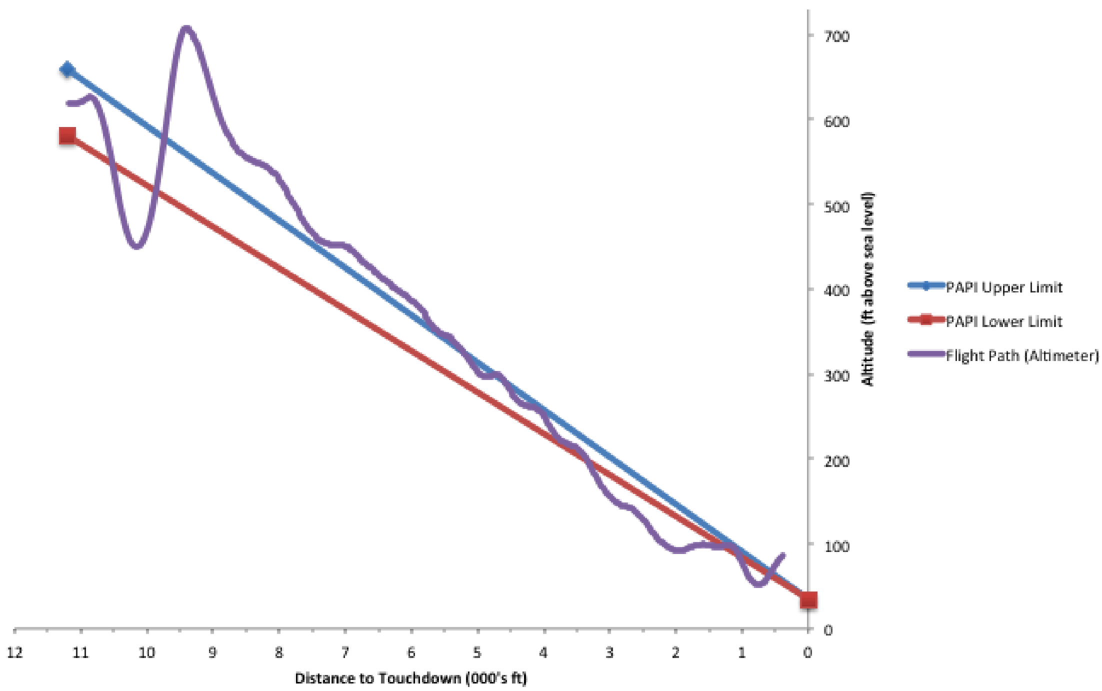

Figure 47. An example of "wait and see" in use. Notice the control inputs become small and iterative at 9,000 ft. However near the end of the approach, when there is not enough time to "wait and see," there are larger deviations and an apparent loss of control.

Rapid loss of control. Participants frequently described this rapid loss of control on short final as "losing it." This occurred when the aircraft's position required large and/or rapid corrections. The participants were not able to apply the appropriate corrections and accommodate for the lag. Any attempt to do so would result in a rapid loss of control. Prior to “losing it,” participants remarked that they when they were not lined up for a successful landing, they knew that the required corrections would result in a certain crash because they didn’t have enough time or space to make them in a controlled manner. This behavior is consistent with the COCOM control modes. As available time decreases, so does the level of control (Hollnagel & Woods, 2005). The high lag condition required a large amount of time to use “wait and see;” thus when that time is not available, the control mode decreases, as was observed.

The small control inputs and the “wait and see” technique are also consistent with traditional control theory. Traditional control theory describes one way to improve stability in an unstable control system is to reduce the sensitivity to stimulations (e.g., Franklin et al., 1994). Participants reduced their sensitivity by reducing the magnitude and frequency of inputs, i.e., small inputs and “wait and see.”

Multiple goals. Another aspect of COCOM is the management of multiple goals (Hollnagel & Woods, 2005). Several participants mentioned that in the high lag condition, they were unable to manage multiple goals at once. They adopted a strategy whereby they would take care of one goal at a time. Participants mentioned intentionally setting airspeed much sooner than do under normal flying conditions. They adopted a tactic of “set it and forget it” in order to avoid the need for managing aircraft speed later on in the approach. Some participants also mentioned focusing first on glideslope of centerline alignment in isolation, and then managing other deviations. This tactic of managing one goal at a time is consistent with the definition of the lower control modes (Hollnagel & Woods, 2005), which were observed in the high lag conditions.

Reduced reliance on display. A final adaptation mentioned by participants was their reduced reliance on the computer display. Some mentioned putting the display in their peripheral vision and only checking it periodically. Others mentioned focusing on the yoke and trying to find the position that provided the desired glideslope on the previous attempt.

Learning. Some participants expressed the view that they would be able to improve performance in the 1000 ms lag condition given enough time to practice. The data did not reflect a noticeable learning effect after the initial warm up, though participants only flew a total of 16 approaches in the high lag condition. It is possible that given a greater opportunity to practice, participants would become more familiar with how the latent system reacts to control inputs. With greater understanding, they may be able to better comprehend the long-term effects of their action and demonstrate safe control.

Appendix G

Discussion of Measures

RMSE. This was the most sensitive of all the measures, though it was not always the most informative. The fact that there is a deviation does not necessarily mean that the pilot is losing control and will not be able to safely land the aircraft. As such, RMSE was able to reveal differences in performances due to a relatively benign wind event, but that did not mean that there was a significant loss of control or reduction in flight safety. Put another way, RMSE is useful for revealing subtle differences and interactions, but not in determining what is safe.

RMSE also failed as a measure when participants had extremely poor performance. This is not necessarily detrimental, because it is obvious when performance is extremely poor and there is no need for a sensitive measure such as RMSE.

Control mode coding. This is a non-parametric measure and is less informative from a statistical standpoint since it cannot provide insight into independent variable interactions. It is also qualitative data based on coder evaluation, which can be subject to inconsistencies, though in this experiment it was found to be acceptably consistent. Beyond these shortcomings, control mode coding can be more informative than RMSE because control has more atributes than simply aircraft position. This is the first time COCOM control modes were used to evaluate pilot performance and so it is reassuring that control mode coding was able to reveal all but one of the significant differences revealed by RMSE. Moreover, the one difference the COCOM coding did not reveal is in fact a demonstration of its strengths. It did not reveal a difference in control due to the wind manipulation, which was a weak manipulation that did not cause participants to lose control.

A challenge when applying COCOM control modes to evaluate approaches is that the difference between strategic and tactical control is very hard to discern. Even with the operationalized definitions, it was difficult to determine the difference between the two, especially during the landing control mode coding since the time frame is so small and participants do not have an opportunity to demonstrate understanding of intricate interdependencies and higher level goals (Hollnagel & Woods, 2005). It can even be argued that there are no higher-level goals and not enough interdependencies to discern between tactical and strategic control at all. This challenge was demonstrated by the fact that the more than half of all intercoder disagreements were between strategic and tactical control modes. While it is difficult to determine the difference between tactical and strategic control modes, in this context, the difference itself is arguably inconsequential. Both control modes are safe and that was the focus of this study.

Lastly, the final control mode coding was more sensitive than the landing success measure. This is due to the fact that it was possible to stay within the bounds of the landing success gates and still exhibit an unsafe control mode, as shown by Figure 48. Furthermore, landing success was a crude measure because it was based on aircraft position and speed. It measured if the aircraft was in a position to make a successful landing, not whether it did (it did not take into account pitch, yaw, or roll).

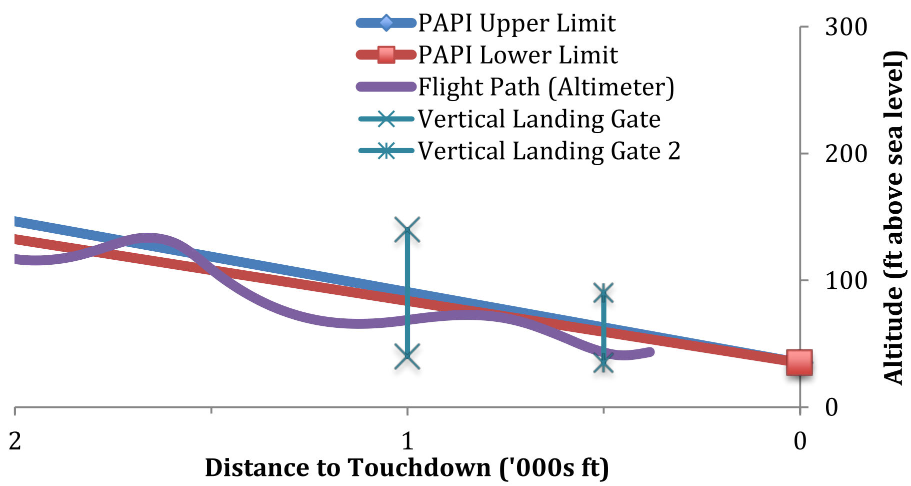

Figure 48. A "successful" landing that exhibits an unsafe control mode.

Intercoder reliability. As mentioned, intercoder reliability was found to be acceptable. While the coders did not discuss every approach together, they did meet frequently to discuss specific approaches and the coding activity in general. These frequent meetings were beneficial and likely had a positive effect on reliability, though there were still a large number of disagreements between tactical and strategic control mode codes. This is one of the weaknesses in using COCOM control modes to evaluate an activity that occurs over a short time span, though for the purposes of flight safety, the difference between tactical and strategic control is marginal.

Appendix H

Lessons Learned

The simulated aircraft had a nose up trim and participants were not provided with the ability to adjust the trim. This reduced the fidelity of the simulator, yet had an unexpected positive impact on the study. An initial concern was that participants would be able to set the yoke to the ideal glideslope and then not need to apply any other control inputs. Due to the nose up tendency of the aircraft, participants had to constantly apply control inputs. This meant that they had to actively control, and thus actively demonstrate control during the entire approach. The constant need to demonstrate control may have benefited this study and should be taken into consideration in future studies.

Participants also mentioned that the controls were very sensitive. The combination of sensitive controls and the nose up trim may have increased the occurrence of pilot induced oscillations (PIOs). In addition to the controls, participants also had issues with the display. It was relatively small and the runway often moved off screen when the aircraft was in a nose down attitude. Future studies should consider improving the fidelity of the simulator.

The wind manipulation was relatively weak compared to the lag manipulation. The lag manipulation was present during the entire approach, but the wind manipulation only lasted 2 seconds during a roughly 2 minute approach. In hindsight, it is not surprising that the wind manipulation was weak compared to the lag manipulation. Future studies should take into consideration the strength and duration of wind manipulations.

Nearly all participants became very frustrated with their performance during the high lag (1000 ms) conditions. They believed they would be better at adapting to the high lag setting and yet they still had a great deal of difficulty controlling the aircraft. When the participants experienced repeated failures during an activity they are normally very good at, they became frustrated. Future studies should take participant frustration into consideration.

The landing success measure looked at the aircrafts position and airspeed at two points just prior to touchdown, but did not consider aircraft attitude. This meant that the aircraft could have hit the runway upside down and it could be considered a successful landing. While this never happened, it does demonstrate the importance of pitch, roll, and yaw when determining landing success. Future studies should increase the number of criterion for landing success in order to improve its validity. In addition to lessons learned, Table 18 provides a summary of the recommended future research that can build of this study.

Table 18

Summary of future research

Observation		Future Research
Lag	Effects	-Research the various control techniques mentioned by participants to see if some are successful in reducing the effect of system lag on operator control -Research learning effects on operator control of systems with lag
Lag	Magnitude	-Research the range between 240 ms and 1000 ms to better understand the relationship between system lag and operator control
Lag	Fixed	-Research the effects of variable lag compared to fixed lag
Wind	Limited effects	-Research effects of other weather phenomena such as: constant crosswinds, multiple crosswinds, and a crosswind that occurs closer to the runway
Wind	Pre-flight brief	-Research the effect of the pre-flight weather brief on participant performance and the possible negative transfer of skills when weather recovery techniques are used to overcome system lag (e.g., dutch-roll techniques)
Simulator fidelity	Rudimentary display	-Research the effect of displays on operator control of systems with lag. This can include: -larger display -multiple displays -additional instrumentation -supplementary information feedback -predictive displays
Simulator fidelity	Rudimentary controls	-Research the effects of control systems on operator control of systems with lag. This can include: -various force feedback configurations -additional controls (e.g., trim controls) -improved control sensitivity -assisted-recovery autopilot
Control modes	Measuring control	-Research and refine the use of COCOM controls modes to evaluate pilot performance. This should focus on improving their use to measure control