

Comparing Three Variants of a Computational Process Model of Basic Aircraft Maneuvering

Jerry T. Ball²
Kevin A. Gluck¹
Michael A. Krusmark²
Stuart M. Rodgers¹

¹Air Force Research Laboratory ²L3 Communications
Warfighter Training Research Division, 6030 S. Kent St., Mesa, AZ 85212-6061 USA
(all email addresses are first.last@williams.af.mil)

Keywords:

Cognitive Modeling, Knowledge Differences, Strategy Differences, UAV

ABSTRACT: *A key objective of cognitive modeling research at the Air Force Research Laboratory's Warfighter Training Research Division is to be able to explore the effects of background knowledge and task strategies on performance and learning of skills relevant to accomplishing the Air Force mission. In pursuit of that objective, this paper compares three variants of a computational process model of basic aircraft maneuvering. All three variants are embodied performance models implemented in the ACT-R cognitive modeling architecture. The model variants operate a Predator UAV Synthetic Task Environment (STE). Each model variant implements a different combination of background knowledge and task strategy for completing the basic maneuvering task. The three variants of the model are called Model Variant P (Performance only), Model Variant CP (Control and Performance) and Model Variant CFP (Control Focus and Performance). Model Variant P lacks the knowledge of control instrument settings typically available to expert pilots and only considers performance indicators in completing the basic maneuvering task. Model Variant CP has knowledge of control instrument settings needed to accomplish various performance objectives and uses that knowledge as part of a crosscheck strategy which includes attending equally to control and performance indicators. Model Variant CFP also has knowledge of control instrument settings, but has a different crosscheck strategy which includes focusing on control instruments until they are correctly set, in addition to normal crosschecking across control and performance indicators. This paper documents efforts to use these model variants to explore the relative effects of differences in knowledge and task strategy on pilot performance in UAV basic maneuvering.*

1. Introduction

A key objective of cognitive modeling research at the Air Force Research Laboratory's Warfighter Training Research Division (AFRL/HEA) is to be able to explore the effects of background knowledge and task strategies on performance and learning of skills relevant to accomplishing the Air Force mission. Currently, the division's Performance and Learning Models (PALM) Research Program is focused on the use of a Synthetic Task Environment (STE) which includes a high-fidelity simulation of a Predator Uninhabited Air Vehicle (UAV) augmented with basic maneuvering, landing and reconnaissance tasks and data collection facilities. We are using the UAV STE as a testbed for conducting empirical research and creating embodied cognitive models of UAV pilot performance and learning. Ultimately, the goal is to use this and other models to develop modeling

guidelines for detailed and psychologically realistic representations of human behavior in complex, dynamic warfighting domains.

This paper will begin by setting the context for our computational cognitive modeling research through some background information on the STE, piloting a UAV, and the ACT-R cognitive modeling architecture which we are using. It then introduces the three model variants, Model Variant P (Performance Only), Model Variant CP (Control and Performance), and Model Variant CFP (Control Focus and Performance), and describes the representations and processes built into each variant of the model. The paper continues with a comparison of the performance of the model variants with each other and in the case of model variants CP and CFP, with human performance data. The paper concludes with a discussion of the relevance of the research for warfighter training.

2. Background on the UAV STE

The core of the STE is a realistic simulation of the flight dynamics of the Predator RQ-1A System 4 UAV. This core aerodynamics model has been used to train Air Force Predator pilots at Indian Springs Air Field in Nevada. Built on top of the core Predator model are three synthetic tasks: the *Basic Maneuvering Task*, in which a pilot must make very precise, constant-rate changes in airspeed, altitude and/or heading; the *Landing Task* in which the UAV must be guided through a standard approach and landing; and the *Reconnaissance Task* in which the goal is to obtain simulated video of a ground target through a small break in cloud cover. The design of these synthetic tasks is the result of a unique collaboration between behavioral scientists and expert pilots of the UAV. The aim in developing the tasks was to identify important aspects of the UAV pilot’s overall task—aspects that tax the key cognitive and psychomotor skills required by a UAV pilot. They are tasks that lend themselves to laboratory study, yet do not fall prey to oversimplifications. The design philosophy and methodology are described in [1]. Tests using military and civilian pilots showed that experienced UAV pilots perform better in the STE than pilots who are highly experienced in other aircraft but have no UAV experience, indicating that the STE is realistic enough to tap UAV-specific pilot skill [2]. Figure 1 provides a view of the UAV STE. The UAV STE consists of a two monitor pilot station with attached stick (right hand), throttle (left hand) and rudder (not shown).

Basic maneuvering is the focus of the current modeling effort. The task requires the UAV pilot to fly seven distinct maneuvers while trying to minimize root-mean-squared deviation (RMSD) from ideal performance on altitude, airspeed, and heading. For each maneuver, a trial starts with a 10-second straight and level lead-in period as the pilot prepares to execute the maneuver. At the end of this lead-in period, the timed trial (either 60 or 90 seconds) begins and the pilot is required to maneuver the aircraft at a constant rate of change with regard to one or more of the three flight performance parameters. The initial three maneuvers require the pilot to change one parameter while holding the other two constant. For example, in Maneuver 1 the goal is to reduce airspeed from 67 knots to 62 knots at a constant rate of change, while maintaining altitude and heading, over a 60-second

trial. Subsequent maneuvers increase in complexity by requiring the pilot to fly trials that change in combinations of two parameters. Maneuver 4, for instance, is a constant-rate 180° left turn, while simultaneously increasing airspeed from 62 to 67 knots and holding altitude constant. The final maneuver requires changing all three parameters simultaneously, decreasing altitude from 15300 to 15000 feet, increasing airspeed from 62 to 67 knots, and changing heading left 270° over a 90-second trial.

Table 1: UAV Basic Maneuvers

Maneuver	Airspeed	Heading	Altitude
1	Decrease	Unchanged	Unchanged
2	Unchanged	Right 180°	Unchanged
3	Unchanged	Unchanged	Increase
4	Increase	Left 180°	Unchanged
5	Decrease	Unchanged	Decrease
6	Unchanged	Right 270°	Increase
7	Increase	Left 270°	Decrease

During a maneuver the pilot sees only the Heads-Up Display (HUD) on the left monitor, and the compass rose, bank angle indicator, lead-in and trial clocks on the right monitor. A view of the HUD is shown in Figure 2. In the basic maneuvering task, the camera view out the nose of the UAV over which the HUD is normally superimposed is blacked out to simulate instrument flying. The various digital and analog indicators include (from left to right): Angle of Attack (AOA), Airspeed, Heading (bottom center of display), Vertical Speed Indicator (VSI), RPM (indicating the throttle setting), and Altitude. The cross in the middle of the display is the reticle, which represents the nose of the aircraft and is fixed in the vertical and horizontal center. The hatched line crossing the reticle is the horizon line, which moves up and down relative to the reticle to indicate changes in pitch, and also rotates around its center point to indicate changes in bank.

At the end of a trial, the results for the altitude, airspeed and heading deviations are displayed graphically, with actual and desired values on each performance parameter plotted across time. Quantitative RMSDs provide numerical feedback for tracking performance. A view of the feedback screen following completion of Maneuver 1 is shown in Figure 3.



Figure 1: UAV STE

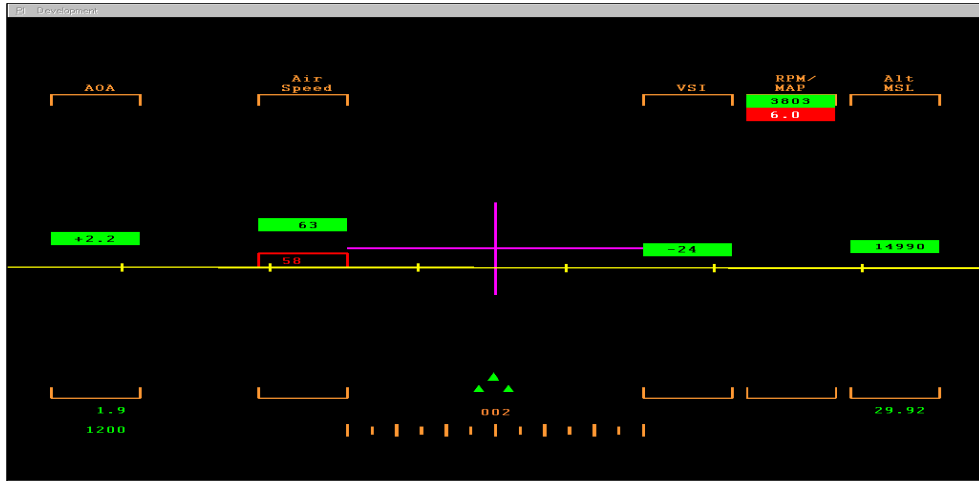


Figure 2: Heads Up Display (left monitor)

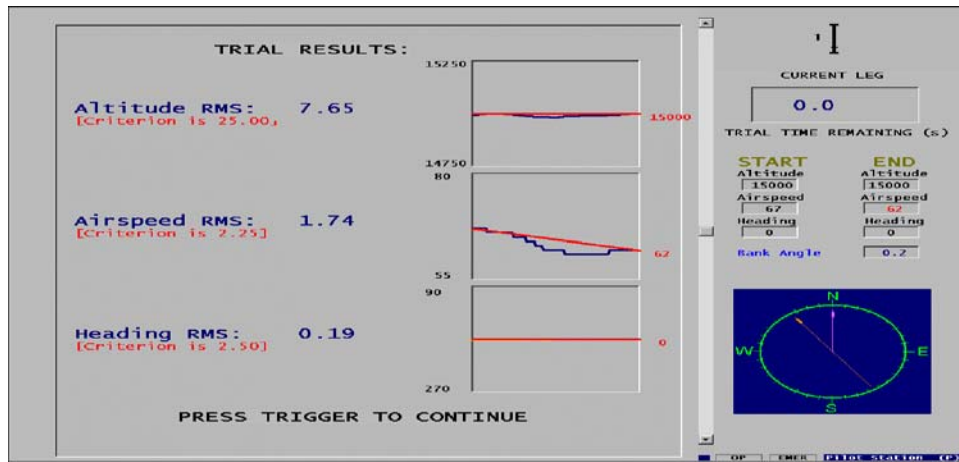


Figure 3 Feedback screen (right monitor)

3. Piloting the UAV.

The instrumentation provided with most aircraft, including the UAV, consists of two different types of indicators: performance indicators and control indicators. Performance indicators reflect the behavior of the aircraft and include airspeed, heading, altitude and vertical speed indicators. Control indicators reflect the settings of the controls which affect the behavior of the aircraft and include pitch, bank and engine speed (RPM) indicators. Adjustments to the controls have a first order effect on the control indicators and a second order effect on the performance indicators. That is, a control adjustment will immediately be reflected by the relevant control indicator. This control adjustment will affect the behavior of the aircraft, which will be reflected in the performance indicators as a second order effect after some delay. According to the Air Force Manual on Instrument Flight [6], a key to expert flight performance is knowledge of the appropriate control settings needed to obtain desired flight performance. For example, a pitch of 3 degrees and an engine RPM of 4300 will maintain straight and level flight of the UAV at 67 knots over a range of altitudes and external conditions. The expert pilot need only set the appropriate pitch and engine RPM to obtain the desired performance, subject to monitoring and adjustment based on variable flight conditions like wind and air pressure.

Lacking knowledge of control settings, novice pilots must rely on the performance indicators to control the aircraft. However, it is quite difficult to control an aircraft with just this knowledge. Often, a novice pilot will overcorrect based on the delay in responsiveness of a performance indicator, requiring a subsequent adjustment in the opposite direction. The result is frequently an oscillation in the performance of the aircraft commonly known as Pilot Induced Oscillation (PIO). There is also a significant interaction between pitch and engine RPM adjustments, and to a lesser extent bank adjustments, which further complicates piloting of the aircraft.

Besides knowledge of control instrument settings, expert pilots are vigilant in maintaining awareness of the status of all indicators. This is typically accomplished by means of a crosscheck across indicators employing either a hub and spoke pattern or a round robin pattern or some mixture of the two. During this crosscheck, it is important not to focus too much attention on performance indicators and to keep control indicators in the crosscheck.

4. Background on ACT-R.

The Atomic Components of Thought – Rational (ACT-R) cognitive modeling architecture and development environment [4,5] is a powerful, yet psychologically constrained, tool for the development of computational cognitive models. ACT-R is being used by researchers around the globe to develop and test cognitive models covering a wide range of behaviors.

ACT-R Version 5 is a Common Lisp based implementation of the ACT-R architecture. It includes a production system integrated with a declarative memory system. The distinction between procedural and declarative memory is a cornerstone of ACT-R and is supported by extensive empirical evidence. ACT-R is a hybrid architecture which provides symbolic productions and declarative memory chunks and subsymbolic mechanisms for production selection and declarative memory chunk activation and retrieval. Production selection and declarative memory chunk activation and selection are implemented as highly parallel processes, however, once selected, production execution is serial—only one production can be executed at a time. These symbolic and subsymbolic components and mechanisms provide the cognitive infrastructure needed to model human behavior at a low enough grain size to model the time course of cognition at the millisecond level. This makes it possible to model real-time human performance in ACT-R. ACT-R Version 5 also provides a perceptual-motor system for interacting with the external world which makes it possible to develop embodied models of cognition. The interface between the production system, declarative memory and perceptual-motor system is coordinated by a collection of buffers. Recent research efforts have focused on mapping the ACT-R architecture to various brain structures and fMRI studies have been conducted to validate that mapping.

The UAV pilot model represents real challenges and opportunities for the use of ACT-R in the development of computational cognitive models. The challenge of integrating ACT-R with an existing real-world simulation environment has not been undertaken by many researchers. The complexity of the task of piloting a UAV is significantly greater than the complexity of the laboratory tasks that provide the historical roots of the ACT-R architecture and its modeling community. The opportunity to take advantage of recent enhancements to ACT-R that support embodied cognition offers the prospect of more ecologically valid models, and in fact makes possible the entire research program of which the work described here is a component.

5. The Three Model Variants

In the interest of exploring the effects of differences in domain knowledge and task strategies on performance, a model of basic maneuvering has been implemented with three distinct variants. The three variants of the model are called Model Variant P (Performance only), Model Variant CP (Control and Performance) and Model Variant CFP (Control Focus and Performance).

Model Variant P lacks the knowledge of control instrument settings typically available to expert pilots and as a result only considers performance indicators and the bank angle indicator in completing the basic maneuvering task. Knowledge of bank angle settings is provided by explicit instruction to participants and is assumed to be available to Model Variant P. Model Variant P centers its crosscheck on the clock and selects from among the following indicators: airspeed, heading, altitude, vertical speed and bank angle. Figure 4a shows the conceptual design of Model Variant P. During an attend-assess-adjust cycle, Model Variant P selects an indicator to attend, locates the indicator on the HUD or right monitor (for bank angle) and encodes its value. Model Variant P then retrieves the desired value for that indicator from memory, compares the desired value to the current value, and sets a qualitative deviation (e.g. small, large, very large) based on the difference between the desired and the current value. The qualitative deviation is then used to assess the need for a control adjustment, and if needed, an adjustment is made. Following this, another indicator is selected and the attend-assess-adjust cycle is repeated.

Unlike Model Variant P, Model Variant CP has knowledge of control instrument settings needed to achieve various performance objectives and uses that knowledge as part of a crosscheck strategy which includes attending to both control and performance indicators. Figure 4b shows the conceptual design for Model Variant CP.

Knowledge of appropriate control instrument settings and how to use them for improved aircraft control distinguishes expert pilots from novice pilots. The first comparison (Model Variant P vs. Model Variant CP) is intended to explore this knowledge difference.

Although Model Variant CP has knowledge of appropriate control settings, it does not focus on a control

indicator until the correct value is achieved. Rather, it makes a control adjustment and continues with the normal crosscheck without checking to see if the adjustment has had the intended effect until the next time the control indicator is attended as part of the normal crosscheck.

Like Model Variant CP, Model Variant CFP has knowledge of appropriate control instrument settings. However, Model Variant CFP uses that knowledge as part of a crosscheck strategy which includes focusing on control instruments until they are properly set, in addition to performing a normal crosscheck across control and performance indicators. Figure 4c shows the conceptual design for Model Variant CFP. In this design there are two distinct attend-assess-adjust cycles: (normal) crosscheck and establish-control (or control focus). The normal crosscheck attend-assess-adjust cycle is the same as that of Model Variant CP. The establish-control attend-assess-adjust cycle is where Model Variant CFP focuses on a single control indicator until the value of that indicator is qualitatively close enough to the desired value (e.g. very small deviation). For Model Variant CFP, the establish-control attend-assess-adjust cycle is entered at the start of a trial in order to establish straight and level flight and whenever the qualitative deviation of a control indicator is assessed to be large (e.g. large deviation). It should be noted that during the lead-in period, Model Variant CP also only looks at the control indicators in order to establish straight and level flight. This is accomplished by restricting the indicators in the normal crosscheck to just the control indicators during the lead-in period and does not involve focusing on a single control indicator until its value is appropriately set, as Model Variant CFP does.

The second comparison (Model Variant CP vs. Model Variant CFP) is intended to explore the effect on performance of using a strategy of focusing on control instrument settings. To some extent, focusing on control instruments disrupts the normal crosscheck across performance and control indicators. Despite this disruption, focusing on control instruments may improve performance since it allows the pilot to complete the correct setting for a control instrument prior to resuming a normal crosscheck.

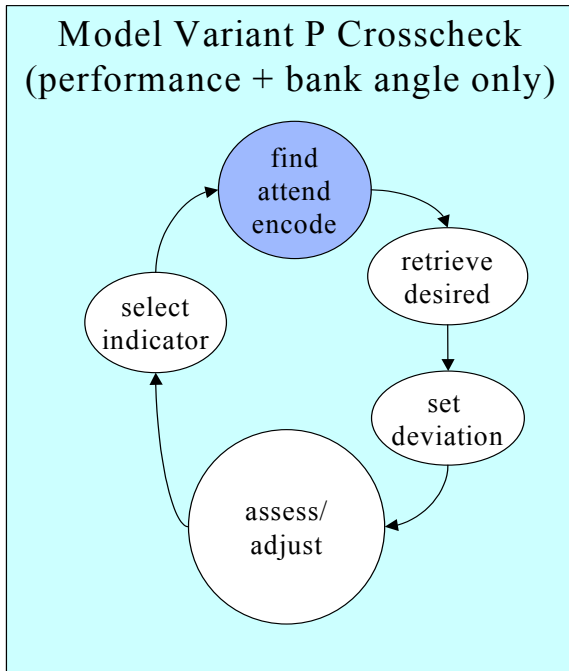


Figure 4a: Model Variant P

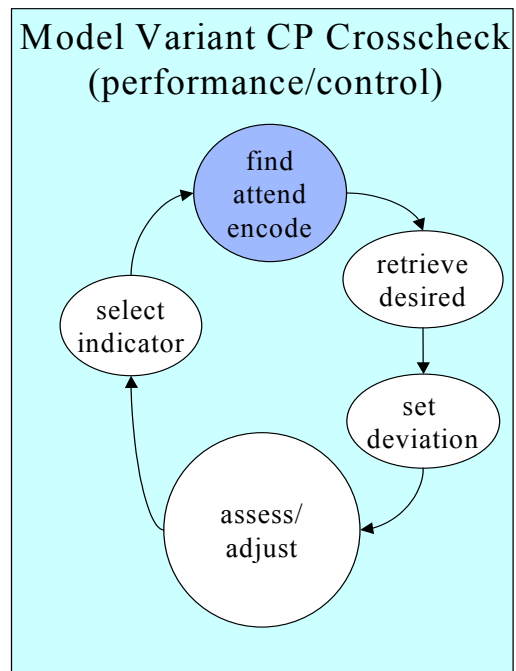


Figure 4b: Model Variant CP

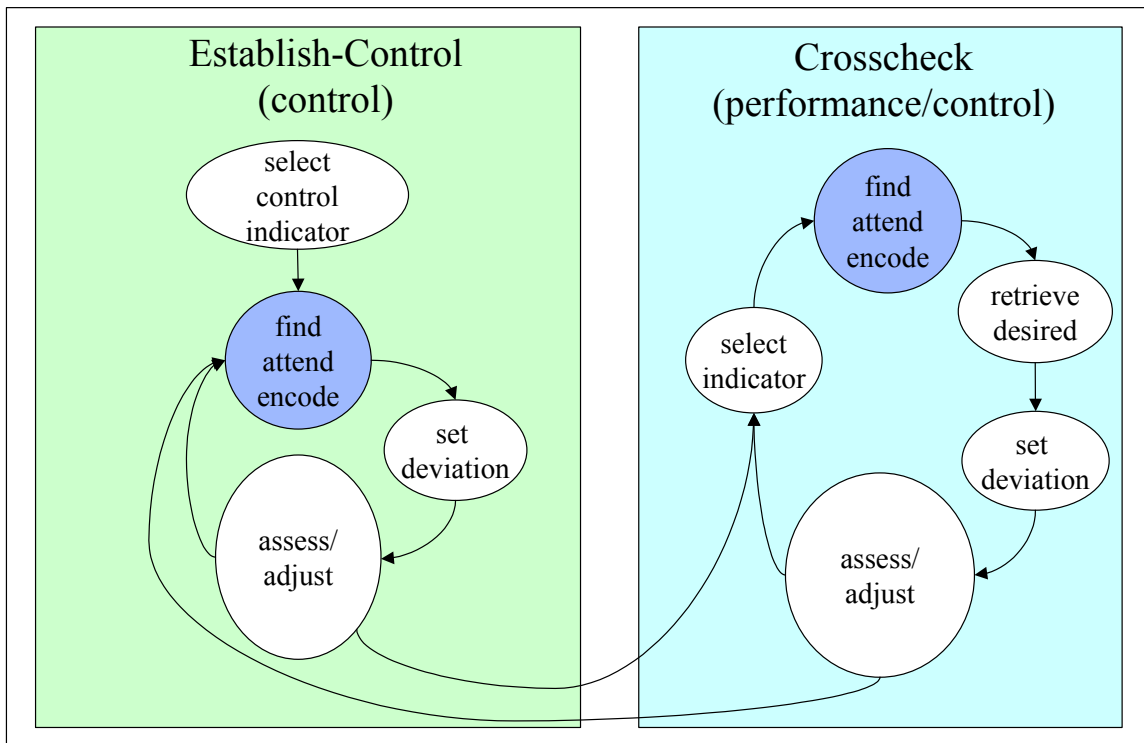


Figure 4c: Model Variant CFP

5.1 Declarative Knowledge

The model variants have declarative knowledge in addition to that described above. The most important of these declarative knowledge chunks are discussed below.

The goal chunk contains the knowledge, or links to the knowledge, needed to fly the UAV. It is organized hierarchically in terms of the top-level knowledge relevant to flying in general, knowledge relevant to managing control instruments, and knowledge relevant to monitoring performance indicators. A common modeling practice in ACT-R models is to restrict the size of declarative memory chunks to 3-5 slots (or pieces) of information. In the case of the goal chunk for the UAV pilot model, however, we found this to be unmanageably restrictive. There is just too much information about the pilot's cognitive state and the aircraft's physical state that needs to be available in the goal chunk for decision making.

On the other hand, having all aircraft state data available to the model variant at all times would be too powerful. We are not creating an autopilot. We are creating model variants intended to replicate actual human performance. Therefore, the productions are designed in such a way that, at any one time, only a few slots in the goal chunk are actually used. For example, if the model variant has just attended to airspeed, then the current-airspeed slot is available to the model variant. Slots with values from previous attend-assess-adjust cycles are not assumed to be available, and new values must be encoded from the indicators or retrieved from memory. Thus, although the goal chunk has a sizeable number of slots, only a few of them have available values at any one time.

During the (normal) crosscheck attend-assess-adjust cycle, each model variant must decide which indicator to attend next. Transitions from one indicator to the next are modeled with crosscheck-intent chunks that map from the current indicator to the next indicator. Retrieval of a crosscheck-intent chunk determines the next indicator to be attended based on the current indicator, the maneuver being performed, and the time-segment. Allowing the maneuver which is being performed to influence this selection results in more frequent attention to indicators whose desired values are changing during the course of the maneuver. However, Model Variant P does not consider the maneuver in selecting the next indicator since it only attends to a few performance indicators and the clock. Based on guidance from a subject matter expert (SME), not all possible transitions are available, but multiple transition chunks are typically matched by the production which selects the next transition and when this

occurs the chunk selection is determined stochastically by a combination of chunk activation and activation noise. Model Variants CP and CFP center their crosscheck on the horizon line/reticle, primarily using a hub and spoke pattern, and the horizon line/reticle is frequently attended. Model Variant P, which is not able to use the control information provided by the horizon/line reticle, centers its crosscheck on the lead-in and trial clocks and frequently visits these indicators.

Model Variants CP and CFP have knowledge of the control settings needed to establish straight and level flight at the start of a trial. They also have knowledge of the appropriate control instrument settings at given points in a maneuver (e.g. 15 seconds, 30 seconds, 45 seconds, 60 seconds). All model variants have knowledge of the altitude, airspeed, heading, and VSI (performance) and bank angle (control) indicator settings at similar points in a maneuver since this information is explicitly provided in the instructions for each maneuver.

5.2 Procedural Knowledge

A crosscheck attend-assess-adjust cycle includes the serial performance of a sequence of productions starting with the selection of an indicator for attention, finding and encoding the value of that indicator, retrieving the desired value, comparing the actual value to a desired value, assessing the deviation, and making a control adjustment, if needed. The typical time to complete an attend-assess-adjust cycle is in the neighborhood of 500-1000 msec. According to Still & Temme (2001) "It is not unusual for even trained pilots to spend as much as 0.5 sec viewing a single instrument and durations of two seconds or more are to be expected even from expert pilots in routine maneuvers." [3]

An establish-control attend-assess-adjust cycle does not need to select a control instrument since this is pre-specified at the start of the cycle. Retrieval of the desired setting for the control instrument is only necessary on the first loop through the cycle. The desired value is available in the goal chunk on subsequent loops so long as the same control instrument is in focus. Otherwise, the establish-control attend-assess-adjust cycle is patterned after the crosscheck attend-assess-adjust cycle.

At the beginning of a maneuver, following the lead-in period, all model variants execute a series of productions to transition from the straight and level lead-in period to the maneuver period. These transitions must be learned since they are maneuver specific and do not reflect general aviator knowledge. However, since we are not modeling learning they are provided to the model

variants. The execution of these productions is triggered by perception of an auditory beep which occurs at the start of a trial following the lead-in period, or by recognition that the lead-in period is nearing completion following attention to the lead-in clock. Perception of the beep relies on ACT-R’s audition module. Learning these transitions is crucial to adequate performance on the basic maneuvers even for expert pilots.

5.3 Parameter Settings

A variety of parameters in ACT-R can be modified to influence the behavior of a model [4, 5]. Most parameters have established defaults or commonly used values. One of the long-term architectural goals in the ACT-R community is to settle on default, or at least “commonly accepted,” values for all parameters, in order to further constrain model implementations. This research has benefited from that goal.

The ACT-R parameters are organized into subsystems. A subsystem is enabled in order to take advantage of the parameters it includes. In the model variants described here, the subsymbolic computation subsystem and the base level learning subsystem are enabled and the parameters within these subsystems have been set to default or commonly accepted values as shown in table 2.

Table 2: UAV Pilot Model Parameters

Parameter	Value
Utility Noise (σ_U)	1 (common)
Goal Weight (W)	1 (default)
Latency Factor (F)	1 (default)
Retrieval Threshold (τ)	1 (common)
Decay Rate (d)	0.5 (default)
Activation Noise (σ_λ)	0.25 (common)

The base level learning component is used to establish the base level activation of declarative memory chunks corresponding to the location of control and performance indicators. These base levels correspond to those expected for expert UAV pilots. Activation noise is used to add stochasticity to the selection of the crosscheck-intent chunks that determine which indicator will be attended to next. The partial matching and other learning subsystems of ACT-R and their associated parameters are not currently being used.

6. Verification of Knowledge and Strategy Differences.

As verification that the three model variants really reflect the knowledge and task strategy differences being considered, Figure 5 compares the attention fixation frequencies of each variant.

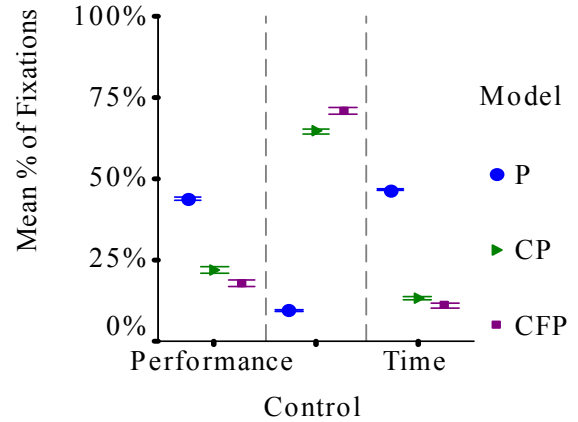


Figure 5: Attention Fixation Mean % for P, CP and CFP

As Figure 5 shows, Model Variants CP and CFP spend significantly more time attending to control indicators and much less time attending to performance indicators, relative to Model Variant P, $t(417) = 58.00, p < .001$ and $t(417) = 38.39, p < .001$ respectively. Although much less dramatic, Model Variant CFP spends significantly more time attending to control indicators than Model Variant CP, $t(417) = 5.49, p < .001$, and significantly less time attending to performance instruments, $t(417) = 4.13, p < .001$. These differences in attention fixation frequency reflect the knowledge and task strategy comparisons of interest.

7. Comparing Knowledge Differences

To compare the effect of knowledge of control instrument settings on performance, Model Variant P is compared to Model Variant CP. To evaluate overall performance on a trial, a composite measure of deviation from desired altitude, airspeed, and heading was computed. On each trial, RMSDs were computed for altitude, airspeed, and heading. Because these performance measures are on different scales, RMSDs were converted to z-scores within class. The resulting standardized RMSDs were added together for each trial, resulting in a standardized sum RMSD. Figure 6 presents the Mean and 95% Confidence Interval for overall performance of the model variants on 20 trials in each of the seven basic maneuvers, averaged across trials within a maneuver.

As Figure 6 shows, a significant main effect of model variant on performance is observed, $F(1,266) = 94.01, p$

< .001, with performance of Model Variant CP significantly better than Model Variant P on 6 of the 7 maneuvers. On maneuver 7, the mean performance of Model Variant P is slightly better, although within the 95% Confidence Interval of Model Variant CP.

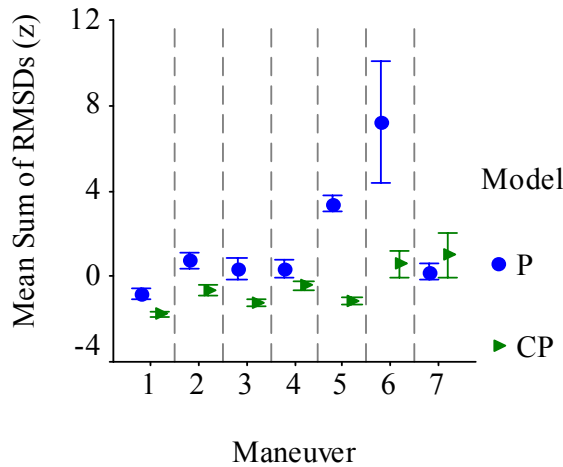


Figure 6: Model P vs. Model CP

Knowledge of appropriate control settings is crucial to pilot performance on 6 of the 7 maneuvers. The performance of Model Variant P was poor enough on many trials that the aircraft was at risk of entering an unstable state. However, Model Variant P performed much better relative to Model Variant CP on maneuver 7. An explanation for this is under investigation. Except for maneuver 7, the performance of Model Variant P is quite poor relative to Model Variant CP. Part of the problem for Model Variant P is that it does not succeed in establishing straight and level flight during the lead-in period since it does not know the appropriate control settings for straight and level flight. As a result, the initial adjustments to the control instruments which occur at the start of a trial following the lead-in period and which are based on achieving straight and level flight, are not appropriate, except perhaps for maneuver 7. Further, the size of the adjustments that are made by Model Variant P to correct performance during the course of the trial are relatively small. This is necessary to avoid the oscillation that can occur from over-correcting and chasing the performance indicators—a common mistake among novice pilots. But when a sizeable adjustment is needed, it means that several separate adjustments must be made and it may take many seconds to accomplish the needed adjustment in conjunction with the normal crosscheck.

Knowledge of control instrument settings provides a counterbalance against adjustments based solely on performance indicators. Model CP often corrects its own

adjustments, backing out (part of) an adjustment based on a performance indicator if that adjustment moves the control indicator away from the desired setting. This counterbalancing helps Model CP avoid over-correcting and destabilizing the aircraft. However, it is important that the control settings be correct in order to maximize performance. One possible explanation for the relatively good performance of Model Variant P compared to Model Variant CP on maneuver 7 is that Model Variant CP was not using optimal control settings. However, as Figure 6 shows, it is the relatively good performance of Model Variant P on maneuver 7 and not the poor performance of Model Variant CP that contributes most to this result.

8. Comparing Strategy Differences

To compare the effect of the control focus vs. normal crosscheck task strategy on performance, Model Variant CP is compared to Model Variant CFP. Figure 7 presents the performance results of these model variants across all seven basic maneuvers averaged across 20 trials within each maneuver.

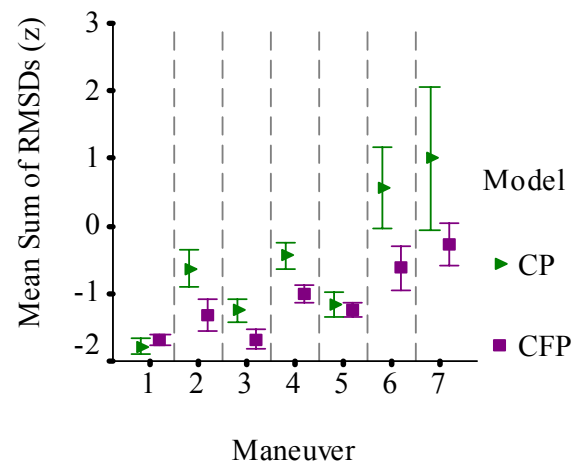


Figure 7: Model CP vs. Model CFP

As Figure 7 shows, a significant main effect is found between Model Variants CP and CFP, $F(1,266) = 37.65$, $p < .001$, with performance of Model Variant CFP significantly better than Model Variant CP on maneuvers 2, 3, 4, 6, and 7, but not on maneuvers 1 and 5. When performing a normal crosscheck—which Model Variant CP is limited to—if an adjustment is made, the effect of that adjustment is not considered until the next time an indicator which reflects the adjustment is attended. It is possible that several seconds may pass before that occurs. If a sizeable adjustment is needed, the full adjustment is not accomplished on a single attend-assess-adjust cycle.

(As noted above, this is needed to avoid overcorrecting and putting the aircraft into oscillation.). This makes it difficult for Model Variant CP to recover from large deviations and the performance of Model Variant CP degrades significantly on the more difficult maneuvers (especially those involving heading changes).

9. Comparison with Human Data

The comparison of model variants is useful to the extent that the model variants accurately model the aspects of human performance under consideration. Is Model Variant P an accurate model of novice pilot behavior at least with respect to the lack of knowledge of control instrument settings? Does the comparison of Model Variant P with Model Variant CP reflect this knowledge? Is Model Variant CP or CFP a better model of expert pilot behavior with respect to the control focus task strategy?

A comparison of the attention fixation frequencies of Model Variant CFP with the eye fixation frequencies collected from two SMEs in preliminary research on maneuvers 1, 2 and 3 suggests that Model Variant CFP focuses more attention on control indicators than is evident in the eye fixation data from the two SMEs, although there is considerable variability between the SMEs. Model Variant CP, which does not focus on control instruments, comes closer to modeling the fixation frequencies of these two SMEs. However, large variability between the two SMEs on the frequency of fixations among indicators motivated a study to collect data from seven additional SMEs on all seven maneuvers. Performance data from this larger sample of SMEs are compared to the model variants below, although eye fixation data from these SMEs are yet to be analyzed.

In the interest of keeping the three model variants comparable, all three variants were implemented with minimal variation other than the variation relevant to the comparisons, namely, knowledge of control instrument settings and control focus task strategy. This meant that the performance of Model Variant P could not be improved by changing the initial control adjustments at the start of a trial to compensate for the failure of Model Variant P to achieve straight and level flight during the lead-in period—although novice pilots may learn such compensatory settings. Moreover, even novice pilots attain reasonable performance on the basic maneuvers given enough trials, however, there is a huge learning curve, with performance on early maneuvers quite abysmal. Given the wide range of performance variability from trial to trial and choices made to keep the model

variants comparable, it is difficult to determine what human data could be compared to that of Model Variant P. As a result, Model Variant P will not be directly compared to human data.

To validate the behavior of Model Variants CP and CFP, we compare the performance of the model variants against the performance of the seven SMEs in the larger study mentioned above. Figure 8 compares the performance of the SMEs against Model Variants CP and CFP on all seven maneuvers. A composite measure of performance was computed to compare overall performance among SMEs and Model Variants. For each SME on each trial, RMSDs for altitude, airspeed, and heading were screened for outliers, converted to z-scores, and summed to generate standardized sum RMSDs. Next, to make model and SME data comparable, through a simple linear transformation, model data for each performance measure were converted to z-scores using the same means and standard deviations that were used to compute z-scores for SME data. The resulting z-scores on RMSDs for altitude, airspeed, and heading were added together. Finally, aggregate means were computed on standardized sum RMSDs for all passed trials of each SME on each maneuver, and for all trials of each model on each maneuver. Thus, Figure 8 depicts means and 95% confidence intervals of standardized sum RMSDs for a sample of 7 SMEs on passed trials on each maneuver. Depicted for Model Variants CP and CFP on each maneuver are means of standardized sum RMSDs computed from the 20 trials completed. Mean performance of model variants CP and CFP can be considered point predictions of performance of highly competent UAV pilots implementing different control focus strategies.

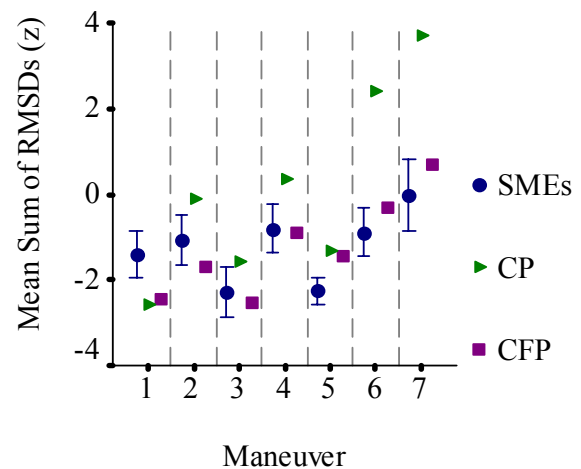


Figure 8: Model Variants CP and CFP vs. SMEs

As Figure 8 shows, the behavior of Model Variant CFP falls within the 95% confidence interval of the SMEs on trials 2, 3, 4, 6 and 7. Model Variant CFP is significantly better than the SMEs on maneuver 1 and significantly worse than the SMEs on maneuver 5. There appears to be a learning effect not captured in the model in that Model Variant CFP does better than the SMEs on maneuvers 1, 2 and 3 and worse than the SMEs on maneuvers 5, 6 and 7. Model Variant CP only comes close to SME performance on maneuver 3 and is significantly worse than SME performance on all maneuvers except maneuver 1. It should be noted that the SME data are for passed trials only—which further avoids problems with outliers and minimizes learning effects—whereas the model data is for all trials. If we aggregate across maneuvers as well as trials within maneuver, we see that Model Variant CFP falls well within the confidence interval of the SMEs, whereas Model Variant CP does not (see Figure 9).

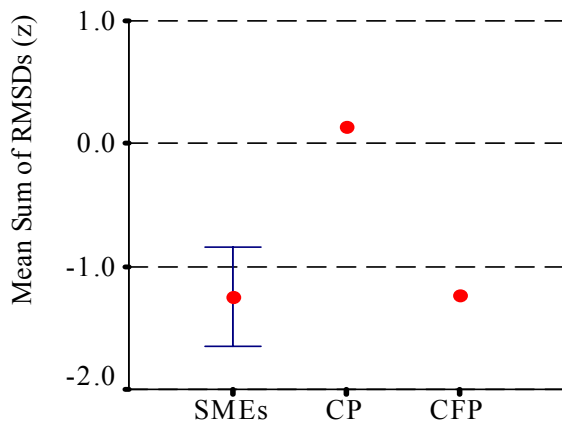


Figure 9: Performance data aggregated across maneuvers

Overall, Model Variant CFP compares much better to SME performance than does Model Variant CP. How then to explain the observation in our preliminary research that Model Variant CFP attends to control indicators more than our two SMEs on maneuvers 1, 2, and 3? One of those SMEs almost never attended to the RPM indicator. How then did this SME control the throttle setting? This SME may have been relying on feedback from the throttle itself, rather than the RPM indicator, to control the throttle. Expert pilots have knowledge of the position of the arm in setting the throttle and stick and receive proprioceptive feedback from these instruments. ACT-R does not currently support the modeling of proprioceptive feedback. As a result, Model Variant CFP (and CP) must rely exclusively on the information provided by the control

indicators to manage control instrument settings. Nonetheless, focusing on control instrument settings—whether via control indicators or via proprioceptive feedback—is likely to be important to overall performance.

10. Relevance to Pilot Training

The analysis of the performance of the model variants suggests that knowledge of control instrument settings combined with a control focus strategy leads to the best performance on the basic maneuvers. Whether or not a control focus strategy is used, knowledge of control instrument settings is crucial to pilot performance. Without this knowledge, Model Variant P is unable to maintain stable aircraft behavior on many trials.

The comparison of the three model variants demonstrates the utility of using computational process models to assess the effects of knowledge and task strategy on performance. Computational process models can be used to conduct very precisely controlled studies of the effects of knowledge and strategy differences on human performance in complex environments. Without such models, those studies require human participants that must be carefully trained to have the desired knowledge or use the desired strategy. The use of human participants is considerably more time consuming and expensive, and considerably less well controlled. The experimenter has less confidence that the human participants are using the desired knowledge/strategy. Once the relative merits of different knowledge and task strategies are assessed via computational cognitive models, the results can be used to make changes to training programs in order to improve warfighter readiness.

11. Acknowledgments

We thank the Air Force Office of Scientific Research for their support. This research is funded by AFOSR grant 02HE01COR.

12. References

- [1] Martin, E., Lyon, D. R. & Schreiber, B. T. (1998). *Designing Synthetic Tasks for Human Factors Research: An Application to Uninhabited Air Vehicles*. Proceedings of the Human Factors and Ergonomic Society.
- [2] Schreiber, B. T., Lyon, D. R., Martin, E. L., & Confer, H. A. (2002). *Impact of prior flight experience on learning Predator UAV operator skills*

(AFRL-HE-AZ-TR-2002-0026). Mesa, AZ: Air Force Research Laboratory, Warfighter Training Research Division

- [3] Still, D. & Temme, L. (2001). "OZ: a human-centered computing cockpit display." In *Interservice/Industry Training, Simulation, and Education Conference*, Orlando, FL.
- [4] Anderson, J. & Lebiere, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: LEA.
- [5] Anderson, J. R., Bothell, D., Byrne, M. D., & Lebiere, C. (2002). An integrated theory of the mind. Submitted to *Psychological Review*.
- [6] Air Force Manual 11-217, Volume 1 (2000). Instrument Flight Procedures.

non-linear problems; and the design and implementation of autonomous, adaptive machines.

Author Biographies

JERRY BALL is a Senior Principal Software Engineer working for L-3 Communications at the Air Force Research Laboratory's Warfighter Training Research Division in Mesa, AZ. Dr. Ball earned a PhD in Cognitive Psychology from New Mexico State University. His research interests include computational psycholinguistics and text comprehension, and the modeling of human behavior in complex environments.

KEVIN GLUCK is a Research Psychologist at the Air Force Research Laboratory's Warfighter Training Research Division in Mesa, AZ. Dr. Gluck earned a PhD in Cognitive Psychology from Carnegie Mellon University in 1999. He is the Director of AFRL/HEA's Performance and Learning Models Research Program and his research is in the area of basic and applied computational cognitive process modeling.

MICHAEL KRUSMARK is a Research Psychologist working for L-3 Communications at the Air Force Research Laboratory's Warfighter Training Research Division in Mesa, AZ. He earned a Masters degree in Cognitive Psychology from Arizona State University. His research interests include qualitative and quantitative methods for validating human behavior models.

STU RODGERS is a Lieutenant Colonel in the United States Air Force and is the Deputy Commander of the Air Force Research Laboratory's Warfighter Training Research Division (HEA) in Mesa, AZ. He earned a Master of Science degree in Computer Science from the University of West Florida. As an experimental test pilot, he has over 3,000 flight hours in 31 different types of aircraft. His research interests include cognitive modeling; heuristic algorithms to solve highly-complex,