Infotech@Aerospace
26 - 29 September 2005, Arlington, Virginia

AIAA 2005-6970

A Prospective Look at a Synthetic Teammate for UAV Applications

Kevin A. Gluck,[*] Jerry T. Ball[#], Glenn Gunzelmann[†]
Air Force Research Laboratory, Mesa, AZ, 85212

Michael A. Krusmark[∞], Don R. Lyon[‡]
L3 Communications, Air Force Research Laboratory, Mesa, AZ, 85212

and

Nancy J. Cooke[§]
Cognitive Engineering Research Institute and Arizona State University Polytechnic, Mesa, AZ, 85212

This paper describes current progress and future plans for research and development in synthetic teammates for applications in training, analysis, and system design for UAV operations. The development of these teammates involves the eventual integration of several distinct, yet related, basic and applied research lines, including navigation and orientation in virtual environments, computational cognitive process modeling of aircraft maneuvering and reconnaissance missions, verbal interaction between human operators and synthetic entities, and the formal analysis of team skill. The use of the ACT-R cognitive modeling architecture to create computational cognitive process models serves as a common thread that will be helpful in integrating the products of these research lines into a functional system. The paper provides a summary of the current status of our research, as well as a description of externally developed technologies we plan to leverage in order to achieve our goal of a high-fidelity cognitive model that is able to operate as a member of a team performing UAV reconnaissance missions.

I. Introduction

Newell, Shaw, and Simon[1] established the research agenda for several generations of computational cognitive scientists in their seminal paper on the formal analysis, representation, and simulation of human problem solving. In that paper they proposed that formal explanations of observable human behaviors could be created through the use of digital computers to generate the sequence of information processing activities required to produce those behaviors. In other words, they proposed that we can use computers to simulate human cognition.

Growth within that research community was slow at first because, among other things, computers were relatively hard to come by until the widespread adoption of personal computing in the early 1980's. Nevertheless, a small, dedicated group of cognitive scientists trained themselves in the necessary methods and technologies, and began developing computational theories and cognitive architectures[2] that accounted for the processes and phenomena in which they were interested.

By the late 1980s a sufficiently large number of these computational accounts were available in adequate breadth and depth that Newell felt motivated to write a book[3] proposing that the time was right to begin pulling these disparate computational accounts together into unified theories of cognition. Shortly thereafter, the emphasis shifted to embodying cognitive models within realistic perceptual-motor constraints.[4] This has culminated in the current emphasis on integrated cognitive systems.[5] Today there are no fewer than two dozen such systems available to

---

[*] Senior Research Psychologist, Air Force Research Laboratory, 6030 S. Kent St., Mesa, AZ, Non-member.
[#] Senior Research Psychologist, Air Force Research Laboratory, 6030 S. Kent St., Mesa, AZ, Non-member.
[†] Research Psychologist, Air Force Research Laboratory, 6030 S. Kent St., Mesa, AZ, Non-member.
[∞] Research Scientist, L3 Communications at Air Force Research Laboratory, 6030 S. Kent St., Mesa, AZ, Non-member.
[‡] Senior Research Scientist, L3 Communications at Air Force Research Laboratory, 6030 S. Kent St., Mesa, AZ, Non-member.
[§] Professor, Applied Psychology, 7001 E. Williams Field Rd., Bldg. 140, Mesa, AZ 85212, Non-member.

those interested in basic and applied research in cognitive modeling. They exist in varying levels of maturity and integration. A series of summaries, overviews, and comparisons of different subsets of these systems has been published recently.[6-9]

The field of cognitive modeling has made significant progress in the last half century. However, as we begin to think of the possible applications for cognitive models, in areas like training, analysis, and design, we realize there is still plenty of room for improvement. For instance, even among the architectures which have a learning capability, their capacity for acquiring entirely new knowledge is modest at best. The result of this is that large investments in knowledge engineering and hand tailoring are required to get the models to behave as desired. This knowledge engineering requirement, along with various degrees and combinations of time pressure, publication pressure, and funding limitations, leads to models that do a good job accounting for specific datasets or empirical phenomena, but tend to be small scale, scripted, and brittle. Our application interests, however, require large-scale, generative, robust models. Thus, in the field of cognitive modeling there are gaps to bridge between models developed for scientific purposes and models developed for applications. These gaps exist along continua associated with scale, generativity, and robustness.

The amount of knowledge required by a model is one way to think of scale issues. Another concern regarding scale is the timescale on which the modeled behaviors are taking place. Anderson[10] described the challenges associated with bridging the timescale gap between typical cognitive phenomena (e.g., the fan effect), which occur in approximately the 10 ms timescale, and typical educational and training applications, which may require hundreds of hours. He referred to success in bridging this gap with computational cognitive models as "… an accomplishment for cognitive science on the order of the Human Genome Project" (p. 106). Thus, there exists an assortment of gaps between the desired goal state for cognitive modeling and the current state of the science, and bridging those gaps is an ambitious undertaking.

Our research approach is a collection of methods selected because we feel they are the best way to make the fastest progress possible in bridging those gaps without adopting an AI approach that sacrifices cognitive plausibility. We use the ACT-R cognitive architecture[11] to develop formal models of human performance and learning, in both simple laboratory tasks and complex synthetic environments, and compare data from the models to data from human participants doing the same tasks. It is worth taking the time to comment briefly on the benefits associated with each component of this comprehensive research strategy.

The cognitive architecture serves an integrating role across our research efforts, both within our research team and between our team and other laboratories who also are using the architecture. It facilitates the sharing of methods and the understanding of model implementations. The simultaneous use of both simple laboratory tasks and complex synthetic environments is an attempt to bridge the domain gap mentioned earlier, through the careful selection of tasks that isolate cognitive phenomena relevant to performance in the complex environment. Finally, the use of human data to assess the validity of model implementations is critical for establishing the utility of the models, either as psychological theories or as tools for applying cognitive science to improve Air Force operations.

The portion of our current research portfolio that is the focus of this paper is a collection of computational modeling efforts selected on the basis that we feel they are on the critical path for achieving our desired goal of a cognitively realistic synthetic teammate. One line of research is focused on the demands placed on spatial cognition when navigating and orienting in virtual environments. The second line of research is the development of a Predator pilot model capable of maneuvering the aircraft and flying reconnaissance missions in a synthetic task environment. The third line involves language understanding and generation to support verbal communication between humans and synthetic entities. The final line of research involves team skill acquisition and retention. The next several sections of the paper describe each of these research lines in more detail.

II. Navigation and Orientation in Virtual Environments

Despite many decades of research, our understanding of how humans encode, store, process, and use spatial information remains limited. There is an extensive literature documenting a variety of phenomena that relate to spatial information processing,[12-19] however an integrated theory that can account for a large subset of those findings is lacking. Some basic principles have been proposed for particular areas of competence. For instance, for large-sized spaces, such as those traversed in complex navigation, principles include hierarchical encoding,[18,20,21] encoding based upon landmarks,[17,22] and the regularization of angle estimates to be nearer to 90 degree intervals.[13,23] For skills like mental rotation, the emphasis has been on the representation and manipulation of visual images.[24,25] Finally, researchers focusing on vision have investigated a variety of phenomena, including perceptual grouping[26,27] and 3-D object recognition.[28] However, these noteworthy empirical and theoretical contributions have not been integrated together to produce a comprehensive understanding of human spatial ability.

Our research on orientation and navigation in virtual environments is targeted at developing such a comprehensive theory. There are three critical aspects of this research. First, we have a series of experiments under way that are aimed at understanding the fundamental capacities and limitations of visual-spatial working memory (VSWM). We have developed a new experimental task that allows for detailed investigation of how people represent complex, 3D spatial information, while limiting the opportunity to use non-spatial strategies to facilitate performance. Second, we are conducting a series of experiments investigating how individuals perform orientation tasks using maps. These experiments are providing an additional level of understanding, beyond the research on VSWM, by uncovering how individuals use their VSWM in a naturalistic context. Finally, we are constructing computational cognitive models in ACT-R to develop a formal understanding of the processes involved in these two tasks. We are using these models to identify the kinds of representations and processes that are needed to accurately capture human spatial competence. All of this research will be brought together to develop an implementation of spatial competence in ACT-R. Subsequently, we will be able to use those mechanisms to facilitate the development of a high-cognitive fidelity computational model that is able to fly UAV reconnaissance missions, which will provide a challenging test of those mechanisms. Each of these components of our research in this area is discussed briefly next.

A. Visuospatial Working Memory (VSWM)

Visuospatial working memory (VSWM) is the set of cognitive processes people use to visualize temporary spatial arrangements of things. VSWM is sometimes called the *visuospatial sketchpad*,[29] a term that captures the purpose and character of this system. VSWM is ubiquitous in everyday life (for example, imagining different furniture arrangements), and is critical for many occupations (engineers, architects, pilots, etc.). However this nonverbal, ephemeral process is difficult to measure objectively. We have developed a technique, called *path visualization*, which allows us to load VSWM and obtain detailed measures of the accuracy and speed with which information can be retrieved from it. Path visualization is similar to some existing techniques,[30-34] but these techniques require people to report a single visualized location, whereas path visualization requires holding a complex path in visual memory. In path visualization, people are given a sequential list of directions to visualize as a path (forward 1 step, left 1 step….). Each time a new segment of the path is described, a decision is required regarding whether or not the new segment intersects with any previous part of the path. Data consist of accuracy and response time for each intersection/no-intersection decision. Additional details of the method are described in Lyon's tech memo.[35]

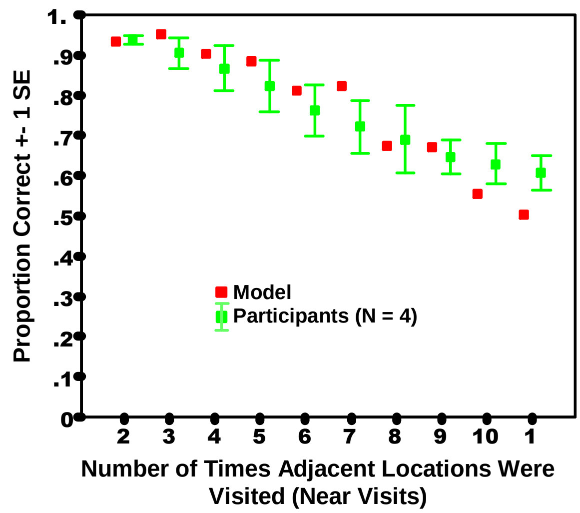


Figure 1. Spatial interference effect in visuospatial working memory – human data and model predictions

Path visualization data reveal a new spatial interference process in VSWM not previously identified. When parts of a path wind around in a small area of (imaginary) space, the parts interfere with each other, degrading memory for them all. So proximity has measurable consequences in imaginary space, just as in real space. Two other characteristics of VSWM are the same as in verbal memory – the likelihood of accessing a part of the path drops over time, and repetition increases the stability of a representation.

We developed a model of path visualization performance in ACT-R, using standard parameter values for the effects of decay and repetition. We modeled the spatial interference process by emulating a 3D spatial field, in which interference varied with Euclidean distance between locations. To test this model, we generated predictions of accuracy as a function of the number of 'near visits', by which we mean the number of previous segments of the path that visit locations adjacent to the most recently presented path segment. This is a measure of the amount of spatial clutter that is near the decision point. Figure 1 shows the model's predictions and the human data.[36] The model's combination of decay, repetition effects and spatial interference successfully capture the data (r^2 = .88; RMSD = 0.045).

B. Spatial Orientation with Maps

The path visualization task provides an opportunity to investigate spatial ability while minimizing the impact of cleverly devised strategies that bypass the need to use spatial abilities. However, real-world tasks provide a rich context that frequently offers opportunities to use a variety of strategies,[14,37-39] which may exercise an individual's spatial ability in a variety of ways. The orientation with maps task is designed to investigate that kind of situation. The task presents participants with two views of a space, an egocentric-based visual scene and a map (Figure 2). There are several variations on the task, but the task always requires that the two views be brought into correspondence to answer correctly. Participants may be asked to identify a highlighted object in one view on the other view, determine the viewer's location or orientation, or perform a more complex task involving route planning or navigation. We are using this task to examine the sorts of strategies that people use, and evaluate how they are using their spatial abilities to solve this kind of problem.

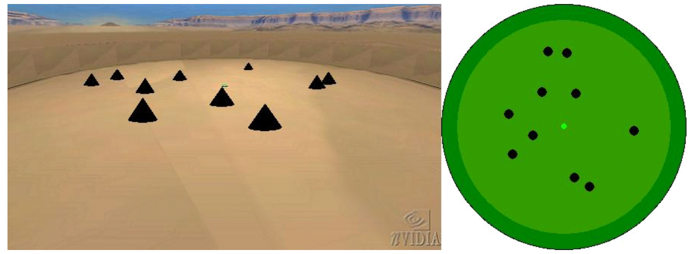


Figure 2. Orientation with maps task. The task is to identify the location of viewer on the perimeter of the map, based upon the view of the space shown in the visual scene on the left.

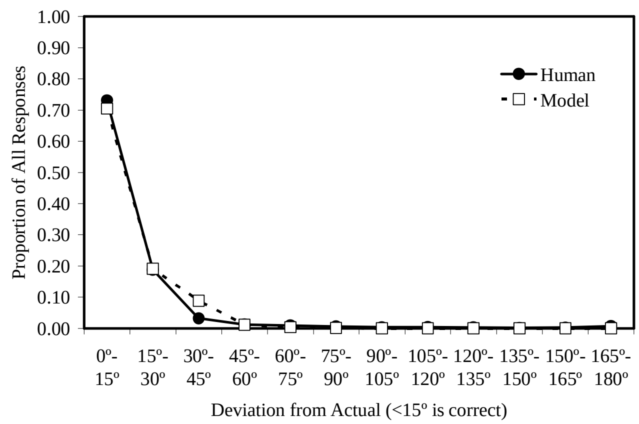


Figure 3. Proportion of responses as a function of angular distance from the correct response. Responses less than 15 degrees from the correct response were scored as correct.

To monitor performance on this task in as much detail as possible, we are collecting a variety of dependent measures. Of course, we obtain response times and accuracy data on a trial by trial basis. These data, by themselves, are informative about how participants solve the problems, and the kinds of strategies they use. For instance, Figure 3 shows the response proportions for the locate-viewer task shown in Figure 2 as a function of how far the response was from the actual correct answer. Responses were scored as correct if they were within 15 degrees of the correct answer. This figure shows that when participants were wrong, their responses were close to being correct in the majority of cases. This suggests that participants are good at developing a qualitative sense of the relationship between the two views, but that they stumble a bit on the quantitative estimates. This result, along with others, provides useful information about the problem solving strategies that participants are using. These strategies, then, form the basis for the computational cognitive model that we are developing to perform the same task. The performance of the model we have developed is similar to human performance on the measures we have tested so far (e.g., Figure 3, $r^2 = .96$; RMSD = .017).

For better resolution on the problem solving process, we are also gathering eye and mouse movement data from participants in these studies. These data provide very fine-grained detail about how each trial is solved, and give us a moment-by-moment indication of what each participant is doing on each trial. We have not yet analyzed these data, as we hope to use the computational cognitive models to make predictions about what the trends in the eye data should be. This is another step toward using computational cognitive models as predictive and prescriptive tools for applications like training.

The orientation task is relevant for helping us understand the spatial demands placed on Predator pilots. Maintaining awareness of the Predator's location and orientation in space are important for making appropriate navigational decisions. Thus, this task provides a good assessment of how spatial competence may be brought to bear in the context of piloting a UAV. For instance, in Figure 4, it is challenging to reason about which way the opening in the cloud layer would move on the left view if the scenario depicted were set in motion. Of course, there are other spatially demanding aspects of the task, including reasoning about wind speed and direction and how that impacts the plane, as well as determining how to maneuver the plane to maximize the amount of surveillance footage that is obtained. These tasks, however, depend fundamentally on an ability to relate the information about the two views of the space, which is the focus of the orientation task.

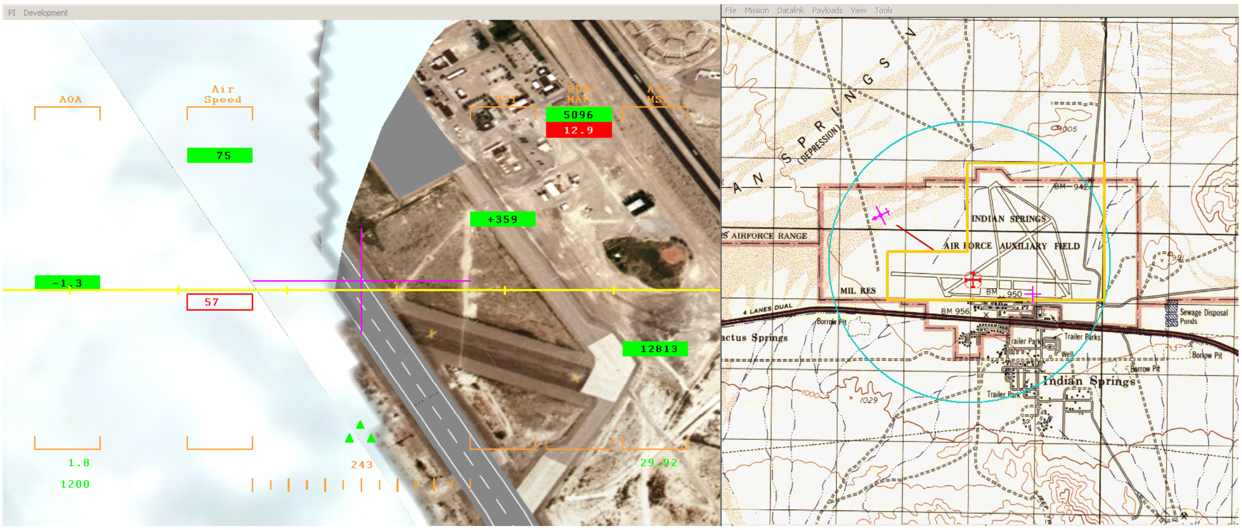


Figure 4. UAV Reconnaissance task (described below). Critical information is depicted on the map. The left view presents an image from a surveillance camera mounted on the bottom of the UAV, which is directed toward the target.

C. A Computational Account of Spatial Competence

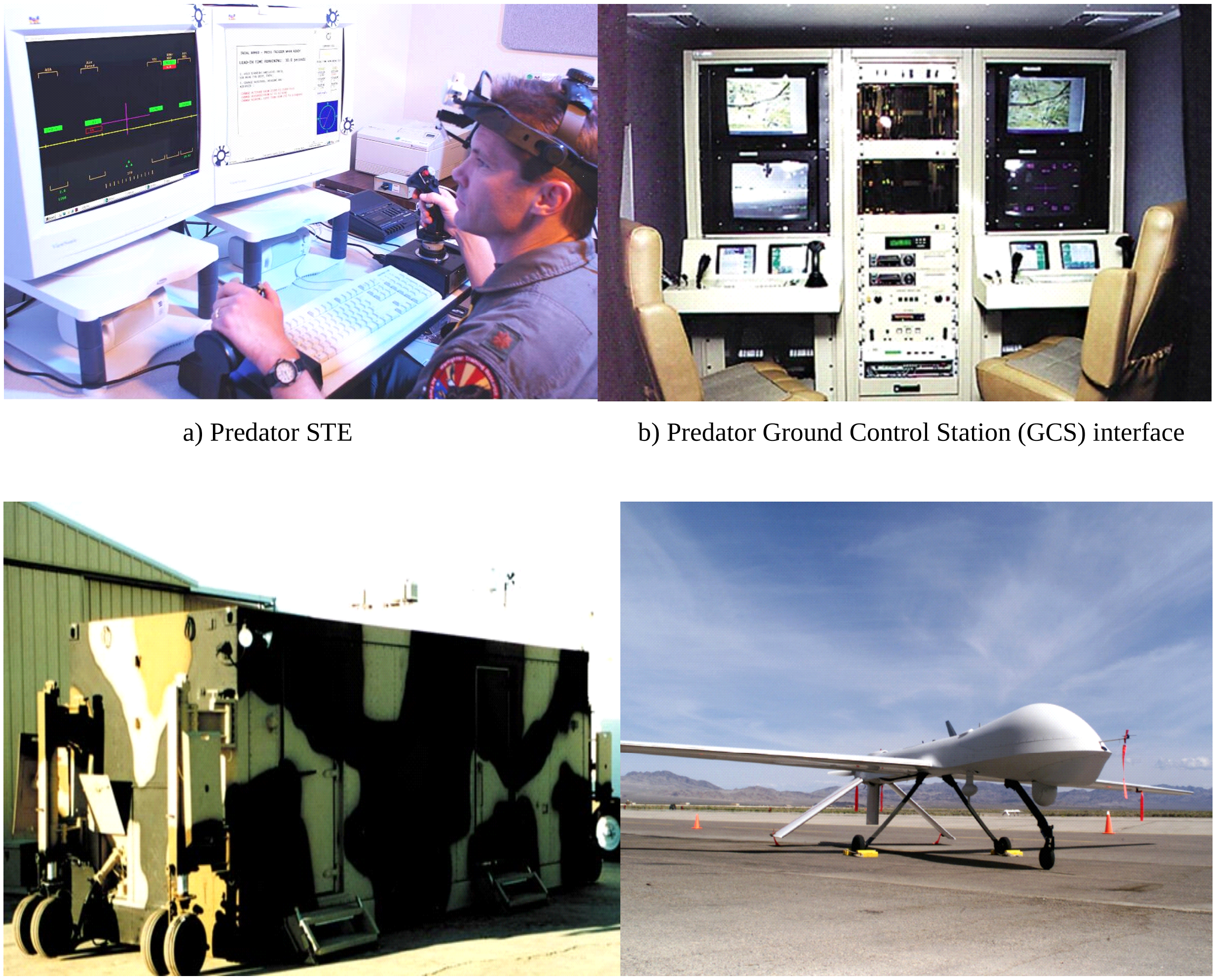
In addition to modeling human performance in the tasks described above, we are developing a detailed theory of spatial competence, which we intend to implement and integrate into the ACT-R architecture. These efforts may appear largely independent on the surface. However, we are using the models we are developing to guide the development and implementation of a broad computational theory of spatial cognition, and the use of a common architecture allows us to draw connections between the VSWM and Orientation tasks.

The models we have developed do a good job of accounting for the data for the individual tasks (see Figures 1 and 3), which lends support to our conceptualization of the spatial processing used by participants to do them. We are integrating many of the concepts identified at the beginning of this section, including hierarchical encoding, a focus on reference features, mental imagery, and regularization of angular estimates (we use qualitative encoding like *left* and *right*, which produces this effect). Thus, we are drawing upon the existing literature to motivate the implementations of these models. The next step is to generalize across the models we have developed for these tasks, to create an integrated account of human spatial competence that can serve as the basis for models of both tasks, and which can also scale up to account for the kinds of spatial problem solving performed by Predator pilots, which is described next.

III.   Basic Maneuvering and Reconnaissance Tasks

A. Predator STE

The primary application context for our current cognitive modeling research is Predator operations. We are using a Predator Synthetic Task Environment (STE) developed at the Air Force Research Laboratory in Mesa to facilitate bridging the gap between basic research and applications of that research that create value for the Air Force.[40] The Predator STE (Figure 5a) is a laboratory version of the system interface available in the Predator Ground Control Station (GCS; Figure 5b), which is housed in a trailer (Figure 5c). The STE includes a high fidelity simulation of the flight dynamics of the Predator RQ-1A (Figure 5d). Wrapped around this core flight model are three synthetic tasks with data collection capabilities: (a) the Basic Maneuvering Task wherein operators make very precise, constant-rate changes to the aircraft's airspeed, altitude, and/or heading; (b) the Landing Task wherein operators fly a standard approach and landing; and (c) the Reconnaissance Task wherein the operator must maneuver the Predator to obtain simulated video of a ground target through a small break in the cloud layer. It has been found that experienced Predator pilots perform better in the STE than highly experienced pilots that have no Predator experience, suggesting that the STE taps Predator-specific pilot skill.[41] Our strategy is that through the use of this realistic, validated STE for cognitive model development, we will increase the transition potential of our basic and applied research.



a) Predator STE

b) Predator Ground Control Station (GCS) interface

c) GCS Trailer

d) Predator RQ-1A

Figure 5. The Predator, the GCS, and the STE

The synthetic tasks that comprise the Predator STE fit well within the larger context of our overall research program. The reconnaissance task in particular places spatial demands on the pilot that directly relate to research questions that are being addressed in our navigation and orientation research. Thus, a major advantage of using the Predator STE in our research is that it provides a relevant environment in which to explore the implications of the models we've developed to account for fundamental cognitive processes in simpler tasks that abstract away from much of the domain knowledge that complicates performance in the real-world. Moreover, because the STE has many of the same complex, dynamic characteristics of real Predator operations, it provides us with an opportunity to push forward the science of cognitive modeling into contexts that align well with the needs of the warfighter. While there has been much progress in computational cognitive process modeling over the last 20 years, the majority of the research has focused on expanding and enriching our understanding of basic cognitive science using simple, controlled, static tasks. Only in recent years has computational cognitive modeling moved into more complex, dynamic[42-44] and real-world[45,46] domains.

B. Basic Maneuvering

For a Predator pilot, the knowledge and skills necessary to effectively maneuver are essential to success. A natural place to begin a research program aimed at developing a fine-grained cognitive process model of a Predator pilot/teammate is the basic maneuvering task. This task was inspired by an instrument flight task originally designed by Wickens and colleagues at the University of Illinois at Urbana-Champaign.[47] The task requires the pilot to fly seven distinct instrument flight maneuvers. Preceding each maneuver is a 10 second lead-in during which time the pilot is asked to stabilize the aircraft in straight and level flight. Following the lead-in is a timed maneuver of 60 or 90 seconds during which time the pilot maneuvers the aircraft by making constant rate changes to altitude, airspeed, and/or heading, depending on the maneuver, as specified in Table 1.

Table 1. Maneuvering requirements in the Predator STE basic maneuvering task.

Maneuver	Airspeed	Heading	Altitude
1	Decrease 67–62 knots	maintain 0°	maintain 15,000 feet
2	maintain 62 knots	Turn Right 0-180°	maintain 15,000 feet
3	maintain 62 knots	maintain 180°	Increase 15,000-15,200 feet
4	Increase 62–67 knots	Turn Left 180-0°	maintain 15,200 feet
5	Decrease 67–62 knots	maintain 0°	Decrease 15,200-15,000 feet
6	maintain 62 knots	Turn Right 0-270°	Increase 15,000-15,300 feet
7	Increase 62–67 knots	Turn Left 270-0°	Decrease 15,300-15,000 feet

During the basic maneuvering task the pilot sees only the Heads-Up Display (HUD), which is presented on two computer monitors (see Figure 6). Instruments displayed from left to right on the first monitor are Angle of Attack (AOA), Airspeed, Heading (bottom center), Vertical Speed, RPM's (indicating throttle setting), and Altitude. The digital display of each instrument moves up and down in analog fashion as values change. Depicted at the center of the HUD are the reticle and horizon line, which together indicate the pitch and bank of the aircraft. On the far right of the second monitor are a trial clock, bank angle indicator, and compass. During a trial, the left side of the second monitor is blank.

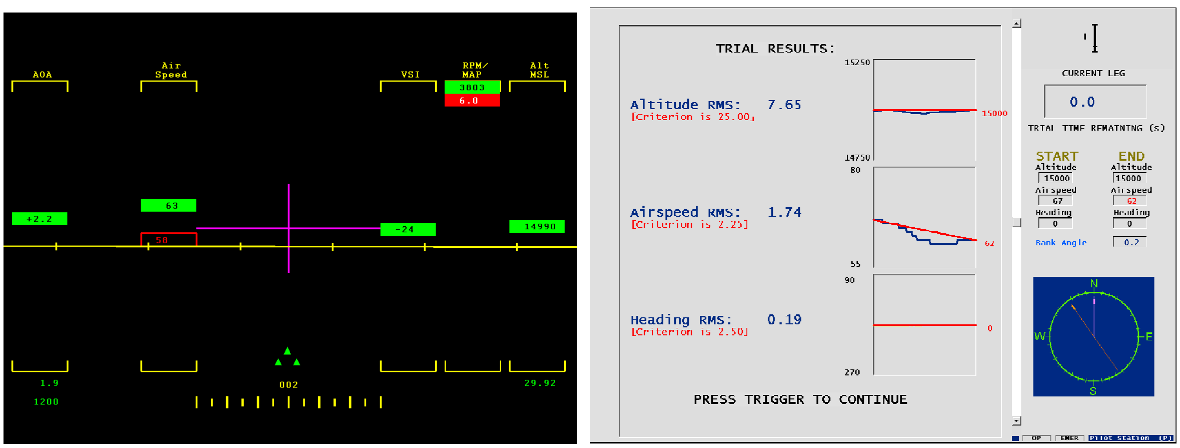


Figure 6. Heads-Up Display (HUD) and Feedback Screen for the Predator STE Basic Maneuvering Task.

At the end of a trial, a feedback screen appears on the left side of the second monitor. The feedback depicts deviations between actual and desired performance on altitude, airspeed, and heading plotted across time, as well as quantitative feedback in the form of root mean squared deviations (RMSDs). The pilot's goal for each trial is to minimize the deviation between actual and desired performance on airspeed, altitude, and heading.

We have developed an expert model of basic maneuvering that is based on an instrument flight strategy called the 'Control and Performance Concept'.[48] The strategy involves first establishing appropriate control settings (pitch, bank, power) for desired aircraft performance, and then crosschecking instruments to determine whether desired performance is actually being achieved. This is an effective flight strategy because control actions with the stick and throttle have immediate, first-order effects on pitch, bank, and power, which then result in lagged, second-order effects on performance parameters like airspeed, altitude, and heading. Controlling a dynamic system on the basis of first-order effects is more efficient and effective than controlling a dynamic system on the basis of second-order effects, so an effective way (and the recommended way) to maneuver an airplane is to adjust the controls until the control instruments show the desired readings, and then simply let the aircraft's performance change as a result of the control surfaces (along with proper crosschecking of all instruments, of course).

Validation of the model comes from both performance and process data that were collected from the model and seven aviation experts – highly experienced pilots located at the Air Force Research Laboratory in Mesa. The model compares well with experts on overall performance, and performance by maneuver, as assessed through a composite performance measure that considers deviation between actual and desired airspeed, altitude, and heading.[49]

Several specific results[49] are worth highlighting. First, the model captures an effect of maneuver complexity even though it was not intentionally designed to do so, wherein for both the model and expert pilots, performance was best on one-axis maneuvers, followed by two-axis maneuvers, and then the three-axis maneuver. Second, goodness of fit estimates computed from model and expert performance data compared well with average fit estimates computed from each expert's performance compared to the rest of the experts. In fact, the fit of the model to the experts' data is better than the fit of one particular expert's data to the rest of the experts' data. Both of these results, in addition to results from other analyses, suggest that the model is a good approximation of expert performance on this task.

Verbal protocol results suggest that not only does the performance level of the model compare well to that of experts, but also the processes that underlie that performance compare well to those used by experts.[50] While experts were performing trials of the basic maneuvering task, we collected fine-grained process measures: retrospective and concurrent verbal reports, and eye-tracking data. Retrospective verbal reports from the experts suggest that they were indeed using the control and performance strategy when performing the maneuvering task. Concurrent verbal reports suggest that maneuver goals influence how experts perform the task, as one would expect. Participants verbalized attention to heading much more frequently on maneuvers that required a heading change (maneuvers 2, 4, 6, & 7) relative to those that did not (1, 3, & 5). This result is consistent with the way the model is implemented. In more recent analyses of these data, we have found that model fixations, expert eye-fixations, and expert verbalizations on instruments displaying information about the lateral axis (bank, heading, & compass) were more frequent on heading change maneuvers relative to non-heading change maneuvers.[51]

C. Reconnaissance

Currently we are in the process of extending our Predator pilot model to the reconnaissance task. Recall that during the reconnaissance task the operator must maneuver the aircraft to obtain simulated video of a target through a small hole in a cloud layer (sometimes referred to as the cloudbreak). During the reconnaissance task the pilot sees the HUD on the left monitor. The HUD is superimposed over a simulated video feed from either the Predator's nose or sensor camera. On the second monitor is a map that tracks the location of the Predator relative to the ground target (see Figure 4).

The reconnaissance task is challenging in several respects. Not only must the pilot maneuver the Predator so that the aircraft, target, and cloud hole are all aligned, but this must be done while accounting for an unpredictable cloud hole location, effects of wind on the UAV, no-fly zones, altitude and time restrictions, and maneuverability constraints of the Predator itself. The goal of the task is to maximize time on target while minimizing flight violations.

We are presently collecting data from aviation experts that will be used to validate the model that is under development. The protocol requires participants to spend one day completing basic maneuvering trials until they reach a set performance level on each of the seven basic maneuvers. Then, on day two, the experts fly eight reconnaissance missions that are designed to stress dynamic spatial reasoning through proximal (and variable) placement of the ground target, cloud hole, and no-fly zone, as well as wind speed and direction. Data collected during these reconnaissance missions include various performance and process measures including time on target, time in violation of flight constraints, flight path, eye-tracking data, concurrent verbal reports, and retrospective verbal reports.

D. Interfacing ACT-R to the Predator STE

Computational cognitive models "see" their visual environment by moving visual attention around within a digital representation of that environment. This is fairly trivial with simple, static tasks that are implemented in the same software language as the cognitive model, but it is more complicated when the architecture must interface with an external simulation. The approach we adopted in interfacing our models to the Predator STE was to re-implement the visual displays of the STE in Lisp, the programming language in which ACT-R is written. The focus of the reimplementation was on matching the information provided by the visual display without necessarily reverse engineering the full graphics display of the STE. This was facilitated by the use of digital readouts for the flight instruments (other than the horizon line and reticle) in the STE, such that the model was not required to process an analog device in order to determine the value of the flight instrument. In the case of the horizon line and reticle, ACT-R returns a digital value for pitch and bank to the model (as reflected in the orientation of the horizon line with respect to the reticle), even though a graphic depiction of the horizon line and reticle is displayed. Other than the visual displays, the Predator STE provides a Variable Information Table (VIT) data structure that contains data on most of the flight parameters of the UAV.

Although the Predator STE models both a nose camera (looking forward) and a sensor camera (looking downward), there is no nose or sensor camera view in the basic maneuvering task, since the goal of the task is to require instrument flight. However, for the reconnaissance task, those views had to be represented. This turned out to be a significant challenge which required the support of an aeronautical engineer with a background in 3-D simulation. In addition, not all the data we needed was available in the VIT. A separate cloudbreak data structure provides this information.

It was also necessary to develop a server on the Predator STE computer to trap virtual keystrokes coming from the cognitive model (which runs on a separate computer and sends keystrokes via the Microsoft Windows API) and send them to the Predator STE. These keystrokes are used to change from the nose camera to the sensor camera and back.

In the reconnaissance task, there are many additional visual features that were required in the Lisp representation of the task. For each screen object, we create a virtual object that the cognitive model can access as well as a graphical object for visual display purposes. A decision was made not to fully model the graphics of the tracker map (the right monitor in Figure 4), including contour lines, longitude and latitude lines, terrain features, the runway and surrounding buildings, etc. Instead only the objects that are directly relevant to the reconnaissance mission are modeled: target, ground control station, no fly zone, ring indicating the limit of where the cloud hole can appear, UAV icon. This simplifies the representational requirements, but it is something we will reconsider if data suggest we are somehow sacrificing model validity.

IV. Verbalization Between Operators and Synthetic Entities

The VERBOSE (VERbalization Between Operators and Synthetic Entities) project is an applied research effort aimed at the development of language-enabled synthetic entities for use in training simulation environments. The plan is to merge the Reconnaissance task model (discussed above) with an extended version of a language comprehension model, called Double R Model,[52] which is also under development. The combined model will be integrated into the CERTT Testbed (discussed below) and will perform the role of the Predator pilot as part of a three-person Predator team performing a reconnaissance mission.

As with the other models described in this paper, the VERBOSE cognitive model is being implemented within ACT-R.[11] The language model is unique in attempting to model human language capabilities within a cognitive architecture (distinguishing it from most AI and computational linguistic systems) as part of a large-scale, functional language comprehension system (distinguishing it from most models of language processing in cognitive science). The construction of language-enabled synthetic entities is a complex research endeavor and the VERBOSE research is proceeding in several different directions. To the maximum extent practical, we plan to use existing knowledge bases and linguistic and cognitive resources in the construction of a functional system. Besides our commitment to using ACT-R, we are working on the integration of WordNet[53] — a large lexical database motivated on psycholinguistic principles—to provide a full lexicon. We are also investigating the use of FrameNet[54,55] and/or VerbNet[56] for the representation of verb centered constructions (e.g., transitive vs. intransitive verb) — a capability not provided by WordNet. We are extending Double R Model to support the recognition and processing of multi-word expressions and constructions (currently Double R Model processes one word at a time). An earlier effort[57] looked at integrating CYC,[58] a massive knowledge base of commonsense knowledge, with Double R Model. Integrating these resources without sacrificing cognitive plausibility is a key research objective.

Research in the development of a Situation Model[59] to ground the referring expressions in the linguistic input is also ongoing. The situation model is a spatial-imaginal representation that will make use of the visuo-spatial module being developed for ACT-R as part of the Navigation and Orientation research effort (discussed above). The situation model will contain a representation of the objects and entities and their relative orientation (and other relations) as described in the linguistic input and perceived in the environment. The situation model replaces the use of abstract "concepts" in many other approaches to the representation of meaning. In Double R Model terms, the concept PILOT is viewed as just an alternative linguistic form for "pilot" and claims that uppercase words are somehow representative of non-linguistic concepts is eschewed in favor of their grounding in a spatial-imaginal representation of objects and relations among objects. This spatial-imaginal grounding is not yet specified in a computational implementation, but it is the direction in which the research is headed.

An Historically Black Colleges and Universities (HBCU) research contract has just been awarded to the City College of New York (CCNY) to investigate the use of Latent Semantic Analysis (LSA) for determining word sense frequencies. LSA is a statistical technique based on Singular Value Decomposition (SVD) of matrices reflecting word to text associations extracted from large text corpora. SVD can be used to reduce the number of dimensions of association between words and texts (initially the number of words times the number of texts) leading to the extraction of the latent (i.e., non-explicit) semantic similarity between the words and texts (and indirectly between words and words). The goal of this project is to determine the base word sense frequencies of the various senses of words for use in the VERBOSE system as part of the word sense disambiguation (WSD) component of the system.

Some additional requirements of a functional language-enabled synthetic entity include the integration of speech recognition and generation capabilities, some mechanism for inferencing over linguistic[60] and/or spatial-imaginal representations,[61] representation of discourse-level knowledge (provided in part by the situation model) in addition to word-, phrase-, and clause-level knowledge, and a mechanism for tying linguistic representations to behavior (e.g., motor actions, shifts of attention, verbal responses). These represent future areas of research for eventual integration into VERBOSE.

The underlying linguistic theory adopted in the VERBOSE effort is motivated by Cognitive Linguistic approaches to meaning and the basic claim that the meaning of words and expressions is grounded in embodied experience and not in some purely abstract, disembodied conceptual realm.[62-64] Further, linguistic structure and meaning go hand-in-hand, whether at the level of words, fixed expressions, or larger constructions. This is in contrast to the predominant Generative Linguistic approach[65] which advocates an autonomous syntax that can be studied in isolation from meaning. In terms of language processing, the VERBOSE system is highly interactive, with words and expressions in the input activating representations in memory that are dynamically integrated into a coherent representation of meaning (assuming the input text is itself coherent). Many of the representations activated in memory correspond to linguistic constructions—larger linguistic units with variable elements—that have been acquired over a lifetime of experience with language (e.g. the transitive construction "Subject kicked Object" is activated by "kicked" in "the man kicked the ball"). The basic language comprehension process involves construction activation (based on the linguistic input and context), selection and integration.[66] Given the focus on the development of a computational implementation of a language comprehension system founded on principles of Cognitive Linguistics, VERBOSE can be described as a Computational Cognitive Linguistic system, a term that is not yet in currency.

The creation of language-enabled synthetic entities entails integrating VERBOSE into a software agent that is capable of interacting in a simulation environment. The simulation environment we have chosen is the CERTT UAV testbed that was designed to study team training and which will provide a useful testbed for studying communication between the synthetic entity and human teammates. Our cognitive model of a Predator pilot flying a reconnaissance mission will provide the basis for creation of the software agent.

V. Measurement and Modeling of Team Skill

Although there are platform-to-platform variations, operation of the Predator system requires multiple individuals on the ground functioning as a command-and-control team. The CERTT (Cognitive Engineering Research on Team Tasks) Laboratory hosts a three-person simulation of UAV ground control based on Predator operations.[67] This synthetic environment provides an ideal testbed for understanding and measuring team performance and cognition in a command-and-control setting. The simulated version of this UAV ground control task requires participation of the pilot or Air Vehicle Operator (AVO) who flies the UAV, the Payload Operator (PLO) who controls the camera systems to take pictures, and Data Exploitation, Mission Planning, and Communications (DEMPC) operator who determines the route and is a source of information. The three team members work interdependently, each at a console, in order to position a UAV at target waypoints to take photographs. The synthetic teammate described in this paper will replace a human pilot in this setting.

Although individual performance is measured in this task (e.g., a score based on course deviation for the AVO), the performance of the team is of most interest. The team performance measure is tied to team goals and is a composite score based largely on number of targets photographed and amount of resources used. Because the UAV team, like command-and-control teams in general, has members with heterogeneous backgrounds, averaging of individual data does not necessarily reflect team-level performance that is represented in the team score. Further, team performance data collected in the context of the CERTT UAV task[68] shows improvement over trials and this improvement is not associated with changes in individual or team knowledge. Instead, it seems that team skill is attributed to team coordination or the timely and adaptive sharing of information among team members.

Current research in the CERTT Lab is investigating the development of team coordination skill and its retention over time. In addition, modeling efforts are underway to apply dynamical system techniques to these data. Team coordination is measured by extent of deviation from an optimal model (passing of information in timely manner at each target waypoint).

The synthetic teammate needs to interact with the two human teammates in order to seamlessly integrate into this coordinating system. The synthetic task is so structured that much of the required interaction can be scripted or rule-based with some flexibility engendered by natural language understanding (so that non-synthetic teammates can pass and ask for information in a variety of ways). However, the challenge arises when there are unexpected events or changes in the plan. For example, equipment may break down or targets of opportunity may appear on the scene. This will require not only natural language understanding but also a deeper understanding on the part of the synthetic teammate of information needs of others and its own capabilities. The synthetic pilot will need to understand team members' roles and task-related goals and subgoals in order to adapt to these novel situations. Also, there are some subtle timing constraints in information sharing that are exhibited by experienced team members. The synthetic teammate will also have to be able to respond or request information of the right person at the right time.

VI. Concluding Thoughts

In this paper we have provided an overview of our past, current, and future computational cognitive modeling research and a description of how that research is intended to come together in support of the applied goal of creating a synthetic teammate for training, analysis, and system design. This has been a prospective look at some of the key cognitive capabilities and constraints on this synthetic teammate because the research is in progress and the integration of the various research lines has not happened yet. Each of the research lines described here (orientation and navigation in virtual environments, Predator pilot modeling, natural language, and team skill) could stand alone as a justifiable research investment area unto itself, but we find it helpful to think of them as each supporting a common application goal state.

Acknowledgments

The research described in this paper was sponsored partly by the Air Force Research Laboratory's Warfighter Readiness Research Division and partly by grant numbers 02HE01COR, F49620-01-1-0261, and F49620-03-1-0024 from the Air Force Office of Scientific Research.

References

[1]Newell, A., Shaw, J. C., and Simon, H. A. "Elements of a Theory of Human Problem Solving." *Psychological Review, 65*, 1958, pp. 151-166.

[2]Anderson, J. R. *The Architecture of Cognition*. Lawrence Erlbaum Associates, Mahwah, NJ. 1983.

[3]Newell, A. *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA. 1990.

[4]Kieras, D. E., and Meyer, D. E. "An Overview of the EPIC Architecture for Cognition and Performance with Application to Human-Computer Interaction." *Human-Computer Interaction, 12*, 1997, pp. 391-438.

[5]Gray, W. D. (Ed.) *Integrated Models of Cognitive Systems*. Oxford University Press, New York. In press.

[6]Pew, R. W., and Mavor, A. S. (Eds.) *Modeling human and organizational behavior: Applications to military simulations*. Washington, DC: National Academy Press. 1998.

[7]Ritter, F. E., Shadbolt, N. R., Elliman, D., Young, R. M., Gobet, F., and Baxter, G. D. *Techniques for modeling human performance in synthetic environments: A supplementary review (HSIAC-SOAR-2003-01)*. Wright-Patterson Air Force Base, OH: Human Systems Information Analysis Center. 2003.

[8]Morrison, J. E. *A review of computer-based human behavior representations and their relation to military simulations* (IDA Paper P-3845). Alexandria, VA: Institute for Defense Analyses. 2004.

[9]Gluck, K. A. and Pew, R. W. (Eds.) *Modeling Human Behavior with Integrated Cognitive Architectures: Comparison, Evaluation, and Validation*. Lawrence Erlbaum Associates, Mahwah, NJ.

[10]Anderson, J. R. "Spanning Seven Orders of Magnitude: A Challenge for Cognitive Modeling". *Cognitive Science, 26*, 2002, pp. 85-112.

[11]Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. "An Integrated Theory of the Mind." *Psychological Review, 111*, 2004, pp. 1036-1060.

[12]Franklin, N. and Tversky, B., "Searching Imagined Environments." *Journal of Experimental Psychology: General, 119*, 1990, pp. 63-76.

[13]Glicksohn, J. "Rotation, Orientation, and Cognitive Mapping." *American Journal of Psychology*, Vol. 107, 1994, pp. 39-51.

[14]Gunzelmann, G., Anderson, J. R., and Douglass, S., "Orientation Tasks Involving Multiple Views of Space: Strategies and Performance." *Spatial Cognition and Computation*, Vol. 4, No. 3, 2004, pp. 209-256.

[15]Hintzman, D., O'Dell, C. and Arndt, D., "Orientation in Cognitive Maps." *Cognitive Psychology*, Vol. 13, 1981, pp. 149-206.

[16]Rieser, J., "Access to knowledge of spatial structure at novel points of observation." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 15, 1989, pp. 1157-1165.

[17]Siegel, A. and White, S., "The Development of Spatial Representations of Large-Scale Environments." *Advances in Child Development and Behavior, Vol. 10*, edited by H. Reese, Academic Press, New York, NY, 1975, pp. 9-55.

[18]Stevens, A. and Coupe, P., "Distortions in Judged Spatial Relations." *Cognitive Psychology*, Vol. 10, 1978 pp. 422-437.

[19]Thorndyke, P. and Hayes-Roth, B., "Differences in Spatial Knowledge from Maps and Navigation." *Cognitive Psychology*, Vol. 14, 1982, pp. 560-589.

[20]Hirtle, S. and Jonides, J., "Evidence of Hierarchies in Cognitive Maps." *Memory and Cognition*, Vol. 13, 1985, pp. 208-217.

[21]McNamara, T., "Mental Representations of Spatial Relations." *Cognitive Psychology*, Vol. 18, 1986, pp. 87-121.

[22]McNamara, T. and Diwadkar, V., "Symmetry and Asymmetry of Human Spatial Memory." Cognitive Psychology, Vol. 34, 1997, pp. 160-190.

[23]Moar, I. and Bower, G., "Inconsistency in Spatial Knowledge." *Memory and Cognition*, Vol. 11, 1983, pp. 107-113.

[24]Kosslyn, S. M., *Image and Brain*, Harvard University Press, Cambridge, MA, 1994.

[25]Shepard, R. and Metzler, J., "Mental Rotation of Three Dimensional Objects." *Science*, Vol. 171, 1971, pp. 701-703.

[26]Koffka, K., Principles of Gestalt Psychology, Harcourt, Brace, New York, 1935.

[27]Van Oefelen, M. P., and Vos, P. G., "Configurational Effects on the Enumeration of Dots: Counting by Groups." *Memory and Cognition*, Vol. 10, 1982, pp. 396-404.

[28]Hummel, J. E., and Biederman, I., "Dynamic Binding in a Neural Network for Shape Recognition." *Psychological Review*, Vol. 99, No. 3, 1992, pp. 480-517.

[29]Logie, R. H., *Visuospatial Working Memory*, Lawrence Erlbaum Associates, Hove, U. K., 1994.

[30]Attneave, F., and Curlee, T. E., "Locational Representation in Imagery: A Moving Spot Task," *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 9, 1983, pp. 20-30.

[31]Brooks, L. R., "Spatial and Verbal Components in the Act of Recall," *Canadian Journal of Psychology*, Vol. 22, 1968, pp. 349-368.

[32]Kerr, N. H., "Locational Representation in Imagery: The Third Dimension," *Memory and Cognition*, Vol. 15, No. 6, 1987, pp. 521-530.

[33]Kerr, N. H., "Rate of Imagery Processing in Two Versus Three Dimensions," *Memory and Cognition*, Vol. 21, No. 4, 1993, pp. 467-476.

[34]Vecchi, T., and Girelli, L., "Gender Differences in Visuo-Spatial Processing: The Importance of Distinguishing Between Passive Storage and Active Manipulation," *Acta Psychologica*, Vol. 99, 1998, pp.1-16.

[35]Lyon, D. R., "Measuring Visuospatial Working Memory Using Path Visualization," L3 Communications Link Division Technical Memorandum 23-1, Mesa, Arizona, 2004.

[36]Lyon, D. R., Gunzelmann, G., and Gluck, K. A., "Emulating a Visuospatial Memory Field Using ACT-R." *Sixth International Conference on Cognitive Modeling*, Lawrence Erlbaum Associates, Mahwah, NJ, 2004, pp. 368-369.

[37]Aginsky, V., Harris, C., Rensink, R., & Beusmans, J., "Two Strategies for Learning a Route in a Driving Simulator." *Journal of Environmental Psychology*, Vol. 17, 1997, pp. 317-331.

[38]Gunzelmann, G., & Anderson, J. R., "Strategic Differences in the Coordination of Different Views of Space." *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum Associates, Mahwah, NJ, 2002, pp. 387-392.

[39]Murakoshi, S., and Kawai, M., "Use of Knowledge and Heuristics for Wayfinding in an Artificial Environment." *Environment and Behavior*, Vol. 32, 2000, pp. 756-774.

[40]Martin, E., Lyon, D. R., and Schreiber, B. T., "Designing synthetic tasks for human factors research: An application to uninhabited air vehicles," *Proceedings of the Human Factors and Ergonomic Society 42^nd^ Annual Meeting*, Human Factors and Ergonomic Society, Chicago, IL, 1998, pp.123-127.

[41]Schreiber, B. T., Lyon, D. R., Martin, E. L., and Confer, H. A., "Impact of prior flight experience on learning Predator UAV operator skills," Air Force Research Laboratory, Warfighter Training Research Division. AFRL-HE-AZ-TR-2002-0026, Mesa, AZ, 2002.

[42]Lee, F. J., and Anderson, J. R., “Modeling eye-movements of skilled performance in a dynamic task,” *Proceedings of the 3rd International Conference on Cognitive Modeling*, edited by N. Taatgen and J. Aasman, Universal Press, Veenendaal, Netherlands, 2000, pp. 194-201.

[43]Lee, F. J., and Anderson, J. R., “Does learning a complex task have to be complex? A study in learning decomposition,” *Cognitive Psychology*, Vol. 42, 2001, pp. 267-316.

[44]Schoelles, M., and Gray, W. D**.,** “Argus Prime: Modeling emergent microstrategies in a complex simulated task environment,” *Proceedings of the 3rd International Conference on Cognitive Modeling*, edited by N. Taatgen and J. Aasman, Universal Press, Veenendaal, Netherlands, 2000, pp. 260-270.

[45]Salvucci, D. D., “Predicting the effects of in-car interface use on driver performance: An integrated model approach,” *International Journal of Human-Computer Studies*, Vol. 55, 2001, pp. 85-107.

[46]Salvucci, D. D., Boer, E. R., and Liu, A., “Toward an integrated model of driver behavior in a cognitive architecture,” *Transportation Research Record*, No. 1779, 2001, pp. 9-16.

[47]Bellenkes, A. H., Wickens, C. D., and Kramer, A. F., “Visual scanning and pilot expertise: The role of attentional flexibility and mental model development,” *Aviation, Space, and Environmental Medicine*, Vol. 68, No. 7, 1997, pp. 569-579.

[48]Air Force Manual 11-217, “Instrument Flight Procedures,” Vol. 1, 2000

[49]Gluck, K. A., Ball, J. T., Krusmark, M. A., Rodgers, S. M., and Purtee, M. D., “A Computational Process Model of Basic Aircraft Maneuvering,” *Proceedings of the 5th International Conference on Cognitive Modeling*, edited by F. Detje, D. Doerner and H. Schaub, Universitats-Verlag Bamberg, Germany, 2003, pp. 117-122.

[50]Purtee, M. D., Krusmark, M. A., Gluck, K. A., Kotte, S. A., and Lefebvre, A. T., “Verbal protocol analysis for validation of UAV operator model,” *Proceedings of the 25th Interservice/Industry Training, Simulation, and Education Conference*, National Defense Industrial Association, Orlando, FL, 2003, pp. 1741-1750.

[51]Gluck, K. A., Ball, J. T., and Krusmark, M. A., “Cognitive control in a computational model of the predator pilot,” *Integrated Models of Cognitive Systems*, edited by W. D. Gray, (to be published).

[52]Ball, J. T., “A Cognitively Plausible Model of Language Comprehension,” *Proceedings of the Thirteenth Conference on Behavior Representation in Modeling and Simulation*, Arlington, VA, 2004, ISBN: 1-930638-35-3, pp. 305-316.

[53]Fellbaum, C. (ed.), *WordNet, an Electronic Lexical Database*, MIT Press, Cambridge, MA, 1999.

[54]Ruppenhofer, J., Ellsworth, M., Petruck, M., and Johnson, C., “FrameNet: Theory and Practice”, [online publication], URL: http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=126 [cited 19 September 2005].

[55]Fillmore, C. J., Johnson, C. R., and Petruck M. R. L., “Background to FrameNet,” *International Journal of Lexicography*, Vol. 16, No. 3, 2003, pp. 235-250.

[56]Palmer, M., Gildea, D., and Kingsbury, P., “The proposition bank: An annotated corpus of semantic roles,” *Computational Linguistics*, in press.

[57]Ball, J., Rodgers, S., & Gluck, K., “Integrating ACT-R and Cyc in a Large-Scale Model of Language Comprehension for Use in Intelligent Agents,” *Papers from the AAAI Workshop, Technical Report WS-04-07*, AAAI Press, Menlo Park, CA, 2004, pp. 19-25.

[58]Lenat, D., “Cyc: A Large-Scale Investment in Knowledge Infrastructure,” *Communications of the ACM* , Volume 38, No. 11, 1995.

[59]Zwann, R., and Radvansky, G., “Situation models in language comprehension and memory,” *Psychological Bulletin*, Vol. 123, 1998, pp. 162-185.

[60]Zadeh, L., “Precisiated Natural Language,” *AI Magazine*, Fall 2004.

[61]Johnson-Laird, P., *Mental Models*, Harvard University Press, Cambridge, MA, 1983.

[62]Langacker, R., *Foundations of Cognitive Grammar, Volume 1, Theoretical Prerequisites*, Stanford University Press, Stanford, CA, 1987.

[63]Langacker, R., *Foundations of Cognitive Grammar, Volume 2, Descriptive Applications*, Stanford University Press, Stanford, CA, 1991.

[64]Lakoff, G., *Women, Fire and Dangerous Things*, University of Chicago Press, Chicago, 1987.

[65]Chomsky, N., *Lectures on Government and Binding*, Foris, Dordrecht-Holland, 1981.

[66]Ball, J., “Construction Driven Language Processing,” *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, edited by B. Bara, L. Barsalou and M. Bucciarelli, 2005.

[67]Cooke, N. J., and Shope, S. M., “Designing a Synthetic Task Environment.” In S. G. Schiflett, L. R. Elliott, E. Salas, & M. D. Coovert, *Scaled Worlds: Development, Validation, and Application*, Surrey, England: Ashgate, 2004, pp. 263-278.

[68]Cooke, N. J., DeJoode, J. A., Pedersen, H. K., Gorman, J. C., Connor, O. O., and Kiekel, P. A.. *The Role of Individual and Team Cognition in Uninhabited Air Vehicle Command-and-Control*. Technical Report for AFOSR Grant Nos. F49620-01-1-0261and F49620-03-1-0024, 2004.