# Effect of Decreasing Accuracy in the Temporal Processor for Attention Switches in a Complex Dual Task


*Brian McClimens*
*Derek Brock*
Naval Research Laboratory
4555 Overlook Ave., SW
Washington, DC 20375
202-404-3798
brian.mcclimens@nrl.navy.mil, derek.brock@nrl.navy.mil



Keywords:
Attention, Human factors and Human-computer interaction, Human movement modeling Model comparison



**ABSTRACT**: *Research at the Naval Research Laboratory (NRL) has shown that the use of auditory cueing can dramatically improve operator performance in dual-task environments for which efficient task-switching plays a crucial role (Ballas, 1992, Brock, 2004, Brock, 2006). In order to better exploit the benefits of auditory cueing for the purpose of attention management in multitasking environments, the Navy desires a model-based understanding of the mechanisms driving human performance in these scenarios. Empirical studies utilizing a complex dual task and related cognitive modeling work developed with the EPIC cognitive architecture [5], have focused on understanding the methods subjects employ to effectively time their transitions between tasks. These models support the notion that time spent on the primary, relatively stateless, tracking task is regulated by state information retained from the secondary, radar task. However, the models do not sufficiently capture the benefits observed in conditions utilizing auditory cuing to assist in attention management. A minor modification to these models results in a dramatic change in model performance, provides insight into when and how auditory cues provide benefit, and raises questions about the methods used by the models to time attention switches between tasks.*


## 1. Introduction

A series of studies conducted at NRL, utilizing the Ballas dual task presented on widely separated monitors, have repeatedly shown a robust improvement in performance on both tasks when auditory cues are used to assist in attention management between the tasks *(Ballas, 1992, Brock, 2004, Brock, 2006).* In order to better understand the factors contributing to improved performance, cognitive modeling has been employed to examine viable strategies for directing attention in conditions with and without the presence of auditory cues.

In this dual task environment, proactive attention management is critical because the two tasks are presented on screens that are separated by a ninety-degree arc. Figure 1 shows a workstation similar to those used in empirical studies at NRL. The tasks utilize the far left and far right screens, while the center screen is left blank. As a result, subjects attending to either task cannot receive visual information from the other. In conditions for which no auditory cues are present, all attention switches between tasks must be self-directed, and cannot rely on external cues.

Cognitive modeling work at NRL has supported the notion that state information from one task could be used to determine the duration of an attendance to the other task (McClimens, 2011). Such information could be used to set an internal timer prior to switching tasks, and performance benefits could be realized without requiring a simulation of progression on one task while attending to the other. These models did not sufficiently replicate the benefits observed from the use of auditory cues to aid in attention management. This paper examines an adjustment to these models intended to address that deficiency.

**Figure 1.** Three-screen console configuration of the Common Display System, the new information workstation being acquired for the U.S. Navy's modernization program and next- generation surface ships. The described dual task utilizes the far left and far right monitors in a similar workstation.

## 2. Background

### 2.1 EPIC

The EPIC cognitive architecture (Kieras, 1997) has been used to build several models of this dual task in the past (Kieras, 2001, Brock, 2006, Hornof, 2010). The models in this paper are an extension of previous modeling work at NRL, and again use the EPIC architecture. These models also make use of a custom-designed encoder for the hostility property of blips on the radar screen, and a timing mechanism that regulates the amount of time spent on the tracking task between attendances to the radar task.

### 2.2 Ballas Dual Task

The Ballas dual task consists of a simple yet demanding, continuous tracking task in which subjects are required to follow the motion of a target object onscreen using a joystick, and an intermittent decision task loosely based on a radar display, in which subjects must classify incoming objects of three types as either hostile or neutral based on rule sets unique to each of the three object types. These two tasks are presented on monitors separated by a ninety-degree arc such that subjects focusing on one of the two tasks cannot receive visual information in their periphery for the other task.

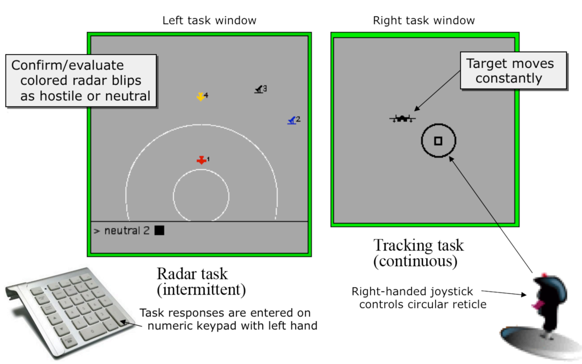


**Figure 2.** A depiction of the radar and tracking tasks.

The tracking task is presented to subjects as their primary task. Although there are neither complex decisions to be made nor critical events in the tracking task, it is a task that demands constant attention in order to perform well. Subjects control a reticle on the screen via joystick, and the performance criteria is simply the average distance between the reticle and a target object over the duration of the experiment. The target moves around the screen with quick, irregular movements that are difficult to predict. These movements are small enough that most subjects are able to track the target relatively well while attending to the task, but even quick glances to the radar screen incur a rapid drop in performance. As a result, the percentage of time spent attending to the tracking screen is a very good predictor of performance. As a general rule, subjects spend between seventy and eighty percent of their time focused on the tracking task depending upon the experimental condition.

The radar task consists of a series of classification judgments based on the motion of three types of objects, collectively referred to as blips. Over the course of a thirteen minute scenario, subjects are required to make sixty five classifications (one every 12s on average). The pace of activity changes throughout the scenarios, but in the studies this model is based upon, there are never more than five blips on the screen at any given time. Blips appear near the top of the screen, and move down towards the bottom of the screen over the course of approximately twenty seconds. When a blip initially appears on the screen, it is black and subjects are not permitted to enter a classification for that blip. When a blip has moved about halfway down the radar screen, it changes color to signify that the subject should enter a response. In experiment conditions that utilize auditory cues, the blip's color change is accompanied by an alert sound so that subjects are made aware of blips ready for classification even when they are attending to the tracking task.

### 2.3 Performance Measures

Recent research at NRL regarding the Ballas dual task has focused on the role that attention management plays in the performance of the radar and tracking tasks. The benefits of auditory cues have been measured primarily using three key performance measures: reaction times for blip assessments, the percentage of time spent on the tracking task, and the number of attention switches between the radar and tracking tasks.

The reaction times on blips in the radar task are defined as the amount of time that passes between a blip changing color and the completion of a response by the subject. These reaction times are more than a simple measure of performance on the radar task. Because blips often change color while a subject is attending to the tracking task, the reaction times increase if subjects fail to effectively manage their attention.

The number of attention switches between tasks, and the percentage of time spent on each task are measures of how much effort a subject is putting in to staying aware of blips on the radar task. In this setup, with the two screens set ninety degrees apart, attention switches are costly, and directly result in poorer performance on the tracking task. Although additional attendances to the radar screen should improve performance on that task, previous research has shown that when auditory cues are present to assist in attention management, subjects are able to improve performance on the radar task while making fewer attention switches and spending less time on the radar task.

# 3. Modeling

### 3.1 2009 data

Early models of the Ballas dual task at NRL suffered from a lack of precision in the data collected regarding the head movements of subjects. Data was recorded during an experiment in 2002 by hand. To record a head turn, the experimenter tapped buttons on a PDA, and time stamps were recorded to the nearest second. Though this process was sufficient for showing significant differences between the sound and no-sound conditions, a greater degree of resolution would be required for more sophisticated modeling efforts.

In 2009, a study was conducted using the Ballas dual task to evaluate the use of new presentation method for auditory cues. This provided the opportunity to collect more detailed empirical data regarding the allocation of attention between the radar and tracking tasks. Previous studies had reported the number of head turns subjects made during the experiment, but collection of data via a head-mounted tracking device was added to allow for more detailed analyses of individual attention switches between tasks. The head-tracking data allowed for the measurement of the durations of each attendance to a task, and allowed for an association between individual task attendances and the states of each task at that time. It was predicted that subjects would maintain an awareness of the radar task's state, and as a result would spend less time in episodes of tracking when there was more activity on the radar task. In other words it was thought that when a subject left the radar task to perform the tracking task, they made note of the current state of the radar task, and used that information to determine when they should return to the radar task. A subject who saw that there were no blips on the radar screen before attending to the tracking screen, would be likely to track for a longer period than they would if the radar screen had a large number of blips on it when they looked away to begin tracking. Data collected in a pilot study supported this prediction, and a model was created to test the impact of allowing state information from the radar task to guide the timing of attention switches between tasks.

### 3.2 Modeling Radar-Driven Task Switches

In order to test the effectiveness of using the radar task's state to guide attention switches, a model was created and run in three modes on two conditions. The two conditions were a "sound" condition, in which an auditory alert was presented whenever a blip was ready to be classified, and a "no-sound" condition in which no alert was used. In both conditions, blips would change color when they were ready for classification, so that a subject attending to the radar task at the time would be aware of the state change. In the sound condition, the audio alert allowed participants to be aware of this state change while attending to the tracking task as well. In each of these conditions the model was run in three modes that used progressively more information from the radar task to inform the temporal processor how much time should be spend on any given attendance to the tracking task. The first mode, referred to as 'mono', was a baseline in which the temporal processor was simply used to ensure that the model spent the same average time on the tracking task as was observed across all tracking attendances in the human subject study. The second 'dual' mode divided the tracking episodes into two categories: those in which tracking began while there were zero blips on the radar screen, and those in which tracking began with one or more blips on the radar screen. When blips were present on the radar screen, the model would spend less time on the tracking screen, in accordance with observed human behavior. In the third, 'multi' mode, four divisions of tracking attendances were made: one group for instances in which there were zero blips on the radar task as tracking began, one for instances with one blip on the radar screen, a third for instances with two blips, and a final group for instances with three or more blips on the radar screen.

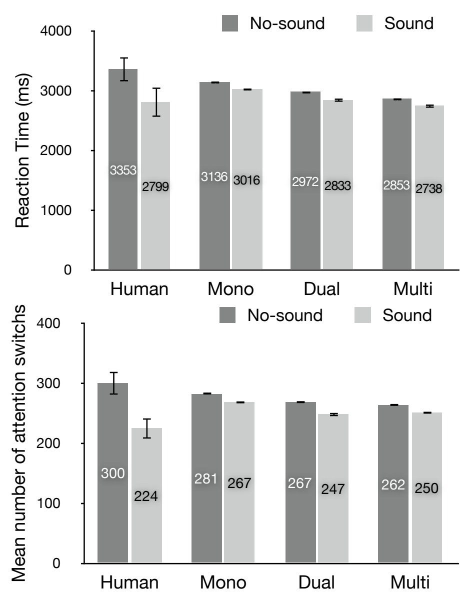


**Figure 3.** These graphs show improved reaction times for classification events on the radar task and a reduction in the number of attention switches between tasks as the model makes more use of state information from the radar task to regulate time spent on the tracking task. This data is reported in (McClimens and Brock, 2011).

As seen in figure 3, the models that made greater use of the radar task's state information to govern attention switches were able to perform better on the radar task, most notably with reduced reaction times for radar task classifications, and also showed a slight decrease in multitasking overhead, reflected by the decreased number of attention switches (McClimens and Brock, 2011). Unfortunately, while there was a small improvement for reaction times in the sound conditions, the differences between sound and no-sound conditions were not nearly as pronounced in the model as was observed in human data. This was also reflected in the number of attention switches made by the model. At the time, the reason for this discrepancy between the model and human data was unknown, but it appears that the reason can be found within the temporal processor used to determine the amount of time spent on each attendance to the tracking task. The model was designed to approximate the average amount of time spent on each attendance to the radar task. In the no-sound condition, this approximation was relatively close (3518ms observed, 3888 modeled), but the addition of auditory alerts did not affect the model as anticipated, and the disparity between the model and observed data increased (4620ms observed, 3986ms modeled). An examination of the distribution of the attendance times reveals an even more striking contrast between the model and observed data, as is shown in figure 4.

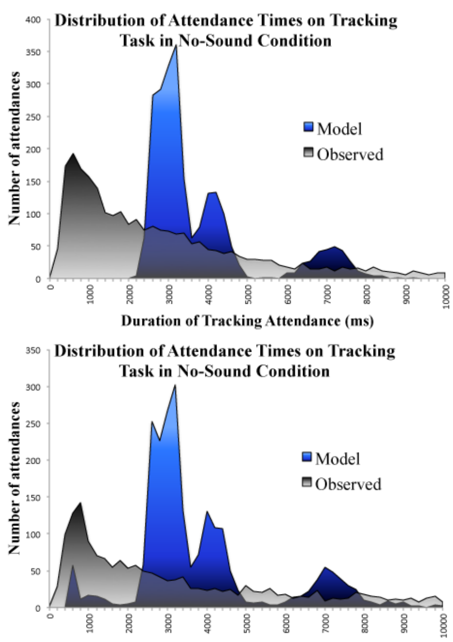


**Figure 4.** Compared to the empirical data, the model showed little variance in the amount of time spent on individual tracking attendances.

### 3.3 Performance Effect of Attendance Distribution

The model's variance in time spent on the tracking task was much narrower than the observed behavior. It was hypothesized that this characteristic of the model might be a key factor in the lack of distinction between the model's performance in the sound condition as opposed to the no-sound condition.

To see how the shape of this distribution can affect performance, consider the effect of an auditory cue on a subject attending to the tracking task. When a subject leaves the radar task to begin tracking, they make an internal estimate of when they need to return to the radar task. Our model works under the assumption that the desired return time coincides with a blip activation, when a blip changes color and is ready for a response. In the sound condition, this event is accompanied by an auditory cue. If a subject is accurate in their estimation, they will arrive on the radar task as the auditory cue sounds. If the subject returns early, the auditory cue will not have been presented. In these two cases, there should be no difference in performance between the sound and no-sound conditions. If the subject makes a poor time estimate, and would be late returning to the radar screen a performance difference results from the two conditions. In the sound condition, the auditory alert will interrupt the long tracking attendance, and prompt the subject to return to the radar task. In the no-sound condition, the subject will continue tracking, unaware that an active blip is on the radar screen, and their reaction time for classification events will suffer as a result. A distribution of attendance times on the tracking screen with less variance provides fewer opportunities for an auditory alert to be beneficial.

### 3.4 Adjusting Noise in the Temporal Processor

The internal clock used in EPIC's temporal processor is an implementation of a timing mechanism developed by Taatgen (Taatgen, 2007). This timer is based on a pacemaker-accumulator model, in which pulses are generated, and an accumulator keeps track of how many pulses have passed. These pulses are not evenly spaced, but rather grow gradually more distant as time passes. Three parameters govern the production of these pulses. The first is an initial pulse length. The second parameter, $\alpha$, determines how quickly the pulse lengths grow. Each pulse length is on average $\alpha$ times the previous pulse length. The final parameter, $\beta$, determines variability within the internal clock. When each pulse's timing is calculated, noise from a logistic function determined by the current pulse length times the third parameter is added. The Ratkin et al. (1998) experiment was used as a benchmark task to find approximate values for these parameters, and values were estimated at 11ms for the initial pulse length, 1.1 for $\alpha$, and 0.015 for $\beta$.

$$t_{n+1} = \alpha t_n + noise(M = 0, SD = \beta * \alpha t_n)$$

The initial model used these same values for the three parameters. As seen in figure 3.2, this results in a narrow distribution around the desired tracking durations.

In order to remedy the differences between the variance of tracking attendance times observed and those produced by the model, the β parameter was heavily modified. Assuming the new model would have a distribution of target attendance times similar to the old model, analysis showed that increasing the value of β from .015 to .2 would increase the variance enough to result in a much closer approximation to the distribution of tracking attendance times in observed data.

## 4. Results

The 'multi' mode of the old model was rerun with a β parameter value of .2 in both the sound and no-sound conditions. The resulting distribution of attendance durations is shown below in figure 5 alongside the distributions from figure 4.

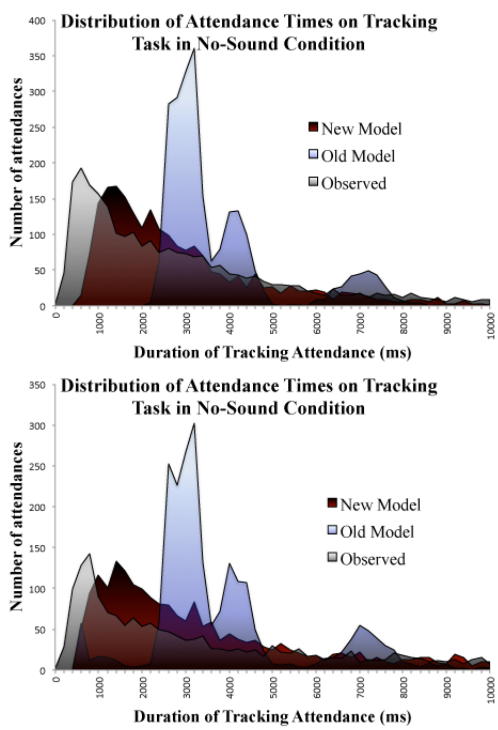


**Figure 5.** Adjusting the β parameter results in a distribution of tracking attendance times that better fits empirical data. Note that despite the apparent differences between the distributions generated by the two models, they use the same formula, with just one parameter modified.

The distributions from the new model peak slightly later than observed data, but increasing the β parameter does result in a much closer approximation of human performance in tracking attendance times. Note too in figure 7, that the mean in the no-sound condition, and the standard deviations are a better fit in the new model. Although the mean tracking time in the sound condition is further from the observed data, note that the difference in means between the no-sound and sound conditions is greater. The original model spent sixty-five percent of its time on the tracking task in both no-sound and sound conditions. The model with increased β spent sixty-six and sixty-seven percent of its time in the no-sound and sound conditions respectively. Empirical data shows subjects spending seventy-five and eighty percent of their time on the tracking task in no-sound and sound conditions.

| | No Sound | | Sound | |
|---|---|---|---|---|
| | Mean | St. Dev. | Mean | St. Dev. |
| Observed | 3518 | 94.3 | 4620 | 140.2 |
| Old Model (β 0.015) | 3888 | 28.8 | 3986 | 39.1 |
| New Model (β 0.2) | 3534 | 109.5 | 3850 | 114.0 |

**Figure 6.** The mean and standard deviation of tracking attendance durations for observed data, original and increased β parameter, in no-sound and sound conditions.

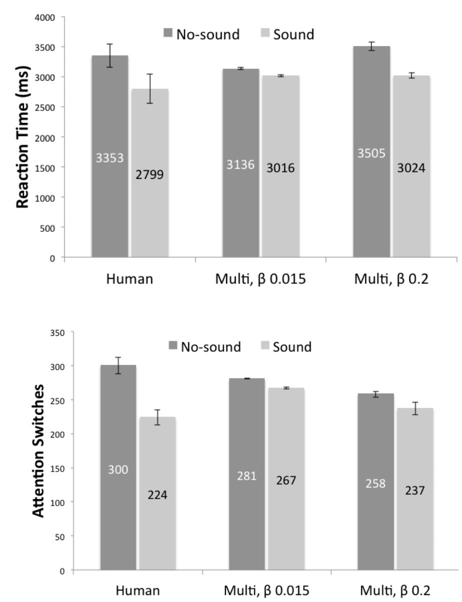


**Figure 7.** As predicted, the new model (the rightmost columns) shows an increased distinction between the no-sound and sound conditions as compared to the old model (middle columns).

## 5. Discussion

Increasing the β parameter for the timer in EPICs temporal processor from 0.015 to 0.2 resulted in a model that was a better approximation of the observed human performance data. Reaction times for classification events on the radar task showed a distinction between the sound and no-sound conditions, and the distribution of tracking attendance durations was more representative of the empirical data. Despite these results, one can question the validity of adjusting a parameter in the temporal processor's formula to fit a single case. In Taatgen's work (Taatgen 2007), it is noted that the behavior of the timing module is assumed to be task-independent, and as such the parameters should be determined by fitting performance to a single benchmark task and then be left alone.

Alternative methods should be explored to determine whether it is possible to fit the empirical data for tracking attendance durations without adjusting the β parameter. Recall that this model uses simplified state data from the radar task to bin tracking attendances into one of four categories (0,1,2 and 3+ blips on radar). For each category, there is a single target tracking duration. A narrow distribution of estimates of these times results in the sharp peaks shown in figures 4 and 5. A model may produce a more realistic distribution by using a continuous function to determine the desired tracking duration rather than discrete categories.

The presented model demonstrates that strategies utilizing state information from the radar task to regulate time spent on the tracking task, can benefit from the effects of auditory cuing. Previous such models that had shown a lack of distinction between no-sound and sound conditions can instead attribute this quality to a lack of fidelity to empirical data in the distribution of tracking attendance durations. Though a simple solution can be attained by increasing the β parameter in EPICs temporal processor, a more sophisticated solution should first replace the discrete function used in the model to determine the desired tracking attendance duration with a more nuanced continuous function.

## References


Ballas, J. A., Heitmeyer, C. L., & Perez, M. A. (1992). Evaluating two aspects of direct manipulation in advanced cockpits. *Proceedings of ACM CHI '92: Conference on Human Factors in Computing Systems*, 127-134.

Brock, D., Ballas, J. A., Stroup, J. L., and McClimens, B. (2004b). The design of mixed-use, virtual auditory displays: Recent findings with a dual-task paradigm. *Proceedings of the 10th International Conference on Auditory Display*. Sydney, Australia.

Brock, D., McClimens, B, Hornof, A, and Halvorson, T. (2006). Cognitive models of the effect of audio cueing on attentional shifts in a complex multimodal dual-display dual-task. *28th Annual Meeting of the Cognitive Science Society*.

Hornof, A.J., Yunfeng, Z. (2010). Task-Constrained Interleaving of Perceptual and Motor Processes in a Time-Critical Dual Task as Revealed Through Eye Tracking. *Proceedings of ICCM 2010: The 10th International Conference on Cognitive Modeling.*

Kieras, D. and Meyer, D. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human Computer Interaction, 12, 391-438*.

Kieras, D. E., Ballas, J., & Meyer, D. E. (2001). Computational Models for the Effects of Localized Sound Cuing in a Complex Dual Task (*EPIC Report No. 13*). Ann Arbor, Michigan: University of Michigan, Department of Electrical Engineering and Computer Science.Hearn, J. and Mills, J. H. (2005). Finding the knowledge edge. CHIPS Magazine, U.S. Navy, vol. 13, no. 3, pp. 17-20.

McClimens, B. and Brock, D. (2011) Modeling Attention Allocation in a Complex Dual Task with and without Auditory Cues. *HCI International 2011 – Poster's Extended Abstracts, Part 1, 317-312*.

Rakitin, B.C., Gibbon, J., Penney, T.B., Malapani, C., Hinton, S.C., & Meck, W.H. (1998) Scalar expectancy theory and peak-interval timing in humans. *Journal of Experimental Psychology: Animal Behavior Processes, 24,* 15-33.

Taatgen, N. A., van Rijn, H., & Anderson, J. R. (2007). An integrated theory of prospective time interval estimation: The role of cognition, attention, and learning. *Psychological Review, 114(3), 577-598.*