Modeling the Complexities of Human Performance

Brian F. Gore
San Jose State University/
NASA Ames Research Center
MS 262-12
Moffett Field, CA, USA
bgore@mail.arc.nasa.gov

Peter Jarvis
QSS Inc./
NASA Ames Research Center
MS 262-12
Moffett Field, CA, U.S.A.
pjarvis@mail.arc.nasa.gov

Abstract – *The Man-machine Integration Design and Analysis System (MIDAS) is an integrated human performance modeling tool that represents many mechanisms that underlie and cause human behavior. It integrates computational representations of human perceptual, cognitive and motor systems to enable high-level behavioral representations characteristic of actual human behavior. MIDAS has been augmented in a number of significant ways to simulate even more realistic simulations of human behavior in various aerospace operational contexts. The effort undertaken in the current project culminated in an agent-based sub model (e.g. slots) that can be used by a variety of models to predict and recreate short- and long-term effects of stressors (fatigue, stress, time pressure, inadequate situation awareness,* etc*) on performance in aerospace accidents/incidents causation. A computational simulation demonstrated performance influences brought to task performance by fatigue that is incurred while undertaking activities required to complete a goal behavior, and the impact of performance-influencing factors on human performance output by combining this with a primitive based action error vulnerability.*

Keywords: Human performance modeling, MIDAS, performance influencing factors, fatigue modeling, error modeling.

1 Introduction

Human performance models (HPMs) and the human performance modeling process have attempted to integrate operator characteristics (cognitive, attentional, and physical) with environmental characteristics to more accurately represent vulnerabilities brought to the system's performance by the human and enable emergent outputs. Some dynamic, integrated HPMs utilize the cognitive task analysis (CTA) process to create their procedural models. The CTA is an ideal way to begin uncovering the cognitive and performance demands associated with aspects of complex human behavior. The integrated representation and association among the various static and dynamic HPMs, includes theoretical, pragmatic, physical and task network models [1]. The output of HPMs has traditionally taken the form of workload predictions, situation awareness predictions and procedural and task timelines.[1] A key to the HPM simulation methodology is that the human operator is simulated and is not physically present in the simulation environment. Only the characteristics associated with a representative human operator's performance are contained within the environment. Workload predictions, timing considerations and system state information are produced as output from the dynamic integrated models. Since the human operator responsible for interacting in these systems is not present in the actual system evaluation, the risks to the human operator and the costs associated with system experimentation are greatly reduced: no experimenters, no subjects, and no testing time. The framework integrates many aspects of human performance allowing each micro model component to behave as designed, the integration of which replicates a human.

[1] For a review of the types of modeling tools that operate in this manner [1,2].

2 The Man-machine Integration Design and Analysis System (MIDAS)

The Man-machine Integration Design and Analysis System (MIDAS), one of the more comprehensive of these models, has been used for procedural analysis and design since 1986[2]. MIDAS has proven useful for identifying general human-system vulnerabilities and cross-domain error classes and for recommending mitigation strategies and job re-designs to account for the vulnerable areas, or risks, in system design [1,3]. MIDAS also possesses a complex visualization environment that can demonstrate the integration of human/system elements. MIDAS represents a "first principles" approach, based on computational models of the mechanisms that underlie and cause human behavior within the context of human-system performance.

MIDAS' agent architecture is made up of physical component agents and human operator agents. Physical component agents use commercially available computer-aided design (CAD) databases to graphically represent physical entities in an environment. Physical component agents are the external environmental influences such as terrain and aeronautical equipment. Human operator agents represent models of cognitive, perceptual and motor operations of a task that describe within their limits of accuracy the responses that can be expected of the human operator for safe operation of advanced automated

[2] NASA Ames Research Center, and the US Army co-developed MIDAS for military-related applications, while NASA and SJSU augmented MIDAS within the complex, multi-crew aviation-related environment.

technologies. Attention demands are represented by Wickens' Multiple Resource Principle and incorporate a task-loading index initially created by McCracken and Aldrich for quantifying attention [4,5]. This scale was modified to include a six-channel representation of task load. Combining attention demands along the input (visual, auditory), central cognitive processing (spatial, verbal), and output (psychomotor, visual) resources accomplish the goal of developing a measure of attention demands. In addition, MIDAS incorporates functions that simulate the effects of stressors on skilled performance through workload and timing exceedances. When the cumulative demands of concurrent tasks exceed an arbitrary threshold of seven, the operator is assumed to be at greater risk for shedding tasks or reduced performance levels, thereby leaving the operator vulnerable to error.

Internal (e.g., intelligence, expertise, personality, emotion, attitudes) and external (e.g., physiological stressors such as fatigue and time stress) moderators of behavior impact human performance in a variety of ways that are very important for the aerospace community. For example, fatigue, stress and other attention decrements have been found to be precursors to operator errors [6]. In aviation, it has been estimated that flight crews' alertness levels are degraded approximately 15% of the time that they are on duty, possibly because they are approaching the boundaries or limitations of their performance capabilities. Excessive time on task has been found to negatively impact a human operator's vigilance, accuracy, grammatical reasoning ability and simple sensory experiences [6]. An inverse relationship has been found between hours of wakefulness and performance on a critical task [6]. Accurately representing human behavior computationally, therefore, requires that many aspects internal to an operator that might impact his performance capability be accurately represented. This fostered the need for models of erroneous performance and performance influencing factors (PIFs) in MIDAS. This report will summarize the recent efforts undertaken to generate a primitive-based error model that becomes triggered by a fatigue-related PIF.

2.1 Primitive Action Error Model

The primitive action error model is one that simulates the variance in the quality and error probability of an operator's primitive actions (reach, grasp, etc) as workload increases. The primitive-level fatigue model is based on a performance degradation model from aviation performance literature in which erroneous performance occurred [7]. The MIDAS model displays error vulnerability information based on the likelihood of error given an operator's predicted workloads. It was determined that designing a simulation environment in this manner would permit creating procedures that could be iteratively refined through examination of resource loaded, simulated operators.

The initial boundary of performance capabilities used the Yerkes-Dodson arousal model of operator performance as the theoretical underpinning for MIDAS behavioral predictions [8]. The Yerkes-Dodson characterization of arousal predicts an inverted U-shaped relationship between arousal and performance. A certain amount of arousal can be a motivator toward change (e.g., learning). Too much or too little change will certainly work against the performer (learner). The desired state of arousal is likely at some mid-level to provide the motivation to change (learn). Too little arousal has an inerting effect while too much has a hyperactive effect. There are optimal levels of arousal for each task. These are:

- Lower for more difficult or intellectually (cognitive) challenging tasks (the learners need to concentrate on the material)
- Higher for tasks requiring endurance and persistence (the learners need more motivation).

This relationship is consistent with the MIDAS representation of operator loads, resource availability, and performance decrements, and serves to augment the predictive capabilities of MIDAS. The decrement is set to occur whenever any of MIDAS primitives are "performed" by the simulated operator (a subset of the MIDAS primitive error table is identified in Table 1) AND when concurrent workload levels exceed the boundary level. Table 1 demonstrates the mapping that was developed to link the primitive actions, the likelihood of error (as identified by the Information Processing Model's probability of failure prediction [9] and the Human Factors Analysis Classification System (HFACS) error classification) [7]), the additional time taken if an error occurs [10], and the quality of the operator's performance as a function of the operator's performance/reliability level [11]. This linkage is a necessary component in generating realistic human behaviors in MIDAS.

Table 1: *Fragment* of the primitive error table

MIDAS Primitive Action	Performance Range[1]	Error Percentage[2]	Time Penalty[3]	Quality Penalty[4]
`fixate-alphanum-object`	0 – 20	2.15	150	0
`fixate-alphanum-object`	21 – 45	.72	150	25
`fixate-alphanum-object`	46 – 80	.72	150	50
`fixate-alphanum-object`	81 – 100	2.15	150	0
`fixate-spatial-object`	0 – 20	2.15	150	0
`fixate-spatial-object`	21 – 45	.72	150	50
`fixate-spatial-object`	46 – 80	.72	150	80
`fixate-spatial-object`	81 – 100	2.15	150	0

It is important to note that an operator's performance in MIDAS is set to 100% and is degraded when the load on any workload channel exceeds the theoretical limit of seven (based on [5]). The error classification was broken into three categories - low, medium, and high. This division was created so that the modeling tool would have a trigger range of behaviors that parallels theoretical boundaries of operator performance that are associated with workload levels [12]. These boundaries were determined from the error potential for the behavior primitives as outlined by HFACS summary of error research [7]. For example, the HFACS classification system outlines that of the total error potential attributable to internal human performance was 91%, only 2.84% of these were classified as the specific errors type within the MIDAS primitive structure (sensory errors), and therefore resulted in 2.58% error rate

(0.91*0.0284*100) in terms of the MIDAS primitive based task performance level. There were three performance ranges based on the Yerkes-Dodson characterization of arousal and performance that were implemented based on HFACS data. Range 1, considered the *low* range, will call the low end of the error scale (0.0082 or 0.82%). Range 2, considered the *medium* range, will call a mid level error rate (0.82%*2 = 1.64%). Range 3, considered the h*igh* range, will call a high level of error rate (0.82% * 3 = 2.47%).

This primitive action error model links the operator's attention loads to the likelihood that an error will occur and becomes activated by the performance influencing models of the simulated operator. One model that incorporates a two-dimensional potential for impacting the performance of the simulated operator is the MIDAS fatigue model.

2.2 Developing the MIDAS Fatigue Behavioral Model

Fatigue is a multidimensional construct that requires appropriate consideration of the fatigue that an operator brings to an activity (pre-task fatigue) and the fatigue that develops as time-on-task extends (within-task fatigue). Two paths were taken to develop the fatigue models in MIDAS – one fatigue model that is triggered at a global, or state, level and one that is triggered at a primitive level. The global level is termed Pre-task fatigue and the primitive level is termed Within-task fatigue.

2.3 Pretask fatigue

Fatigue that an operator brings to a task based on his/her previous activities is modeled as a penalty that is applied to the MIDAS *performance level* for the operator. This feature also serves as a placeholder for future linking to other fatigue models (e.g., the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) circadian rhythm model of sleep deprivation [13] or the two-process model of sleep regulation that interacts the homeostatic regulation process with the circadian process to generate the timing of sleep-wake cycles [14]). The performance level approach dovetails cleanly with the primitive action error. The more fatigued an operator, the lower his/her performance level and the greater the likelihood of errors occurring.

2.4 Within Task Fatigue

The fatigue accumulated during the execution of a task, or within task fatigue, is modeled as a function of the number of times a step is repeated. Table 2 shows a fragment of the table used to compute the effects of in task fatigue. This table demonstrates the implementation of the fatigue-based model of degradation in MIDAS. This table outlines the 'from count', the 'to count', the error probability, the time penalty and the quality penalty.

Table 2: Fragment of the Within-Task Fatigue Table.

Action	From Count	To Count	Error Probability	Time Penalty	Quality Penalty
Fixate-alphanum-object	0	99	1	150	0
Fixate-alphanum-object	100	∞	20	150	25
Fixate-spatial-object	0	99	1	150	0
Fixate-spatial-object	100	∞	20	150	25
Fixate-symbol-object	0	99	1	150	0
Fixate-symbol-object	100	∞	20	150	25

When a new action is instantiated, the number of previous instances executed is retrieved and the error probability, time penalty, and quality penalty values are looked up from the appropriate row in the fatigue table. The within-task values listed in the table are placeholders for more accurate numbers grounded in research. The "from count" and "to count" in Table 2 is a hypothetical point in the model that outlines the time in the simulation when the MIDAS will perform according to a specific error probability (percentage). In the first row of the table, it can be seen that from 0 to 99 ticks in the simulation, the operator will engage in behavior with a 1% likelihood of erroneous performance. When the error occurs, the time penalty associated with this is 150% (i.e., the task will need to be re-started). The *hypothetical* quality construct indicates that during periods of low error probability, there will be a very low likelihood of a quality penalty, while during periods of higher error probability, there will be a higher likelihood of a quality penalty.

The process of integrating the PIFs into MIDAS has also fostered the need to collect data on error likelihood. To collect data an *Error Summation and Presentation Function* was created that aggregates the total error predicted by the *Pre-task* and *Within-Task fatigue* models for a given simulation clock tick.

2.5 Error Summation and Presentation Function

In calculating the error summation and presentation functions, MIDAS calculates two sets of performance values each time an action is executed; one set for Pre-task fatigue (1) that operates as a function of the operator's performance level and a second set for the *Within-task fatigue* (2) that operates as a function of the number of times that actions of this class have been repeated. These numbers are defined mathematically in MIDAS as:

$$\Pr obabilityError(task)_{primitivetask,} TimePenalty(task)_{primitivetask,} \quad (1)$$
$$QualityPenalty(task)_{primitivetask}$$

$$\Pr obabilityError(task)_{fatigue}, \Pr obabilityPenalty(task)_{fatigue},$$
$$QualityPenalty(task)_{fatigue} \quad (2)$$

The total error values for an action are calculated in MIDAS by taking the maximum value of the numbers of the probability of error, the time penalty and the quality penalty. The formulas that were encoded in MIDAS can be found in equations 3, 4, and 5.

$$\Pr obabilityError(task)_{total} = \max(\Pr obabilityError_{primitivetask,} \Pr obabilityError(task)_{fatigue}) \quad (3)$$

$$TimePenalty(task)_{total} = \max(TimePenalty(task)_{primitivetask}, TimePenalty(task)_{fatigue}) \quad (4)$$

$$QualityPenalty(task)_{total} = \max(QualityPenalty(task)_{primitivetask}, Quality(task)_{fatigue}) \quad (5)$$

Several actions are typically executing at any given clock tick in a simulation. The overall error value for a clock tick is calculated in MIDAS by taking the maximum of the total for each task. Given a simulation with *N* tasks running at a clock tick *T,* MIDAS will calculate and collect the data according to the formulae in equations 6 and 7.

$$\Pr obabilityError(T) = \underset{i=1}{\overset{i=N}{Max}}(\Pr obabilityError(i)_{total}) \quad (6)$$

$$\Pr obabilityPenatly(T) = \underset{i=1}{\overset{i=N}{Max}}(TimePenalty(i)_{total}) \quad (7)$$

The top portion of Figure 1 shows the MIDAS runtime display graph (visualization) of the probability of error, time penalty, and quality penalty for each tick of the simulation clock while the lower portion of Figure 1 demonstrates the runtime prediction of operator workload for each of the six attention channels in MIDAS. These predictions are generated from the cumulative summation of the loads associated with the activity that is called by the MIDAS software in response to the environment.

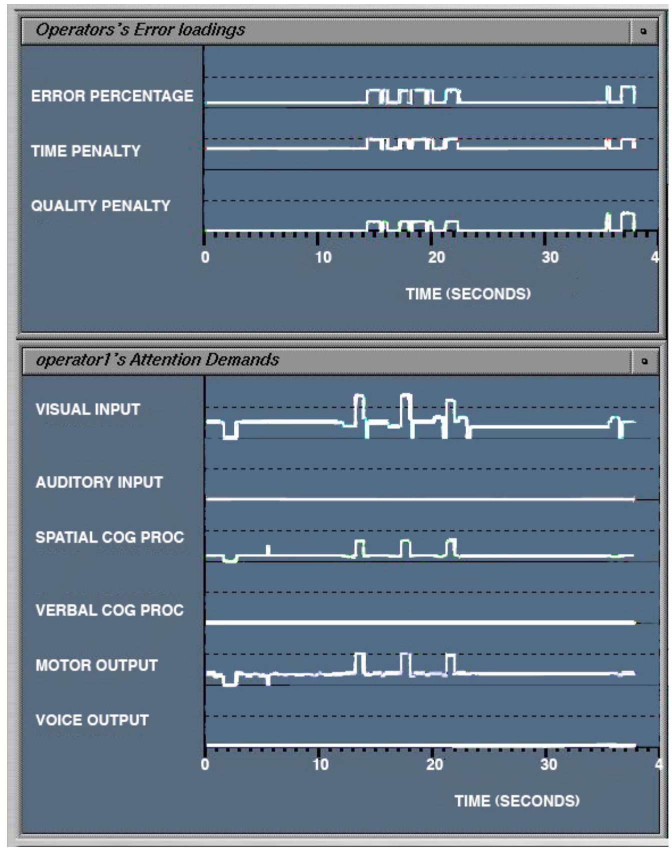

Figure 1. Timeline output (notional) produced by MIDAS integrating the new error and fatigue models.

2.6 Discussion

It is important that integrated HPMs accurately account for the impact of relevant human conditions on human-system performance. Developing an error predictive capability becomes particularly important in the aviation community when attempting to predict operator performance in the face of advanced display designs or new rules of operation and procedural specification associated with passenger travel. HPMs that do not include appropriate PIFs may produce different results than human in the loop performance. Accepting the data output from human performance models that do not account for the effects of various performance influences increases the risks of selecting inappropriate technologies or developing unrealistic procedures.

Creating PIFs within the context of existing, validated HPM software tools enables predictions of performance when humans are operating near their maximum capacity as well as when they are more likely to miss the onset of critical events (e.g., missing flight path deviation as time on task increases). Understanding when the human operator is most vulnerable permits the development and evaluation of mitigation strategies. Testing such advanced system concepts in the relative safety of a HPM is both cost- and time-efficient and, when used in concert with empirical research, is a system design concept that is likely to achieve maximum human performance. Creating PIFs also enables developing technologies and methods to augment the human operator's vigilance and attention to critical events during vulnerable periods. Adopting a system perspective enables the development of mitigation strategies based on performance predictions from a human performance modeling environment.

The Yerkes-Dodson theoretical threshold-model that was implemented in the current modeling effort is consistent with the MIDAS representation of operator loads. It augments the capabilities of MIDAS with a validated model of the impact of arousal (as estimated by the overall level of workload computed within MIDAS) on operator performance. The MIDAS model displays error vulnerability information based on error likelihood given an operator's predicted workloads. This capability allows a procedure designer to create and refine procedures based on predictions generated by the performance of a simulated operator in conditions where performance is suboptimal or where performance vulnerabilities are predicted.

It is important to highlight that the PIF model that has been developed as part of the current effort has not been rigorously validated against empirical data; the models function as designed but may need fine-tuning. The data programmed into the MIDAS primitive-based model of error are valid human performance data that were empirically collected and tied to the primitive behavior level within MIDAS. At this time, the fatigue model is still very simple, but offers "hooks" to insert a more refined model.

2.7 Limitations and Directions for Future Research on PIF Implementation

The development of the fatigue conceptualization as completed in the current effort is limited as it considers fatigue within each action type separately. It does not model fatigue across action types. It is anticipated that a reservoir approach, or a two- or three-stage model of the impact of fatigue on operator performance is warranted given that the current "proof of concept" has shown that the model is responsive to PIFs such as fatigue and human error.

3 Conclusions

The effort undertaken in the current project culminated in an agent-based model (e.g. slots) that can be used by a variety of models to predict and recreate short- and long-term effects of stressors (fatigue, stress, time pressure, inadequate situation awareness, workload exceedances, *etc*) on performance in creating a situation in which incidents or accidents would be more likely. The PIF model demonstrated performance influences brought to task performance by fatigue (in a pre-task fatigue algorithm), by fatigue that is incurred while undertaking activities required to complete a goal behavior, and the impact of these PIFs on human performance output by combining this with a primitive based action error vulnerability. Further research is required to validate the hooks that have been implemented in the MIDAS PIF representation.

References

[1] B.F. Gore, "The study of distributed cognition in free flight: A human performance modeling tool structural comparison," Proceedings of the Third Annual SAE International Conference and Exposition - Digital Human Modeling for Design and Engineering, Dearborn, Michigan, Paper # 2000-01-2181, pp. 1-6, June 2000.

[2] B.F. Gore and J.D. Smith, "Risk assessment and human performance modeling: the need for an integrated approach," in K.A. Malek (ed.) *International journal of human factors of modeling and simulation*, Vol. 1(1), ISSN (Online): 1742-5557 - ISSN (Paper): 1742-5549, (in press).

[3] B.F. Gore, "Human performance cognitive-behavioral modeling: A benefit for occupational safety," in B. Chase & W. Karwowski (Eds.), *International journal of occupational safety and ergonomics* (JOSE), Vol. 8, No. 3, pp. 339-351, 2002.

[4] C.D. Wickens, "The multiple resource model of human performance: implications for display design" In Proceedings of AGARD/NATO Report, Williamsburg, VA, AGARD/NATO, 1984.

[5] J.H. McCracken and T.B Aldrich, "Analysis of selected LHX mission functions: Implications for operator workload and system automation goals," Technical note ASI 479-024-84(b), Anacapa Sciences, Inc., Fort Rucker, AL, 1984.

[6] M.M. Mallis, S. Mejdal, T.T. Nguyen[3], and D.T. Dinges, "Summary of the key features of seven biomathematical models of human fatigue and performance," *Aviation, space, and environmental medicine*, Vol. 75, No. 3, Sect. 2, pp. A4-A14, March 2004.

[7] D.A. Wiegmann and S.A. Shappell, "Applying the human factors analysis and classification system (hfacs) to the analysis of commercial aviation accident data," 11th International Symposium on Aviation Psychology, Columbus, OH, The Ohio State University, 2001.

[8] R.M. Yerkes and J.D. Dodson, "The relation of strength of stimuli to rapidity of habit-information," *Journal of comparative neurology and psychology,* Vol 18, pp. 459-482, 1908.

[9] C.D. Wickens and J.M. Flach, "Information processing", in E.L. Wiener & D.C. Nagel (Eds.), *Human*

[3] Posthumously.

factors in aviation, pp. 111-155, Academic Press, San Diego, 1988.

[10] K.R. Boff and J.E. Lincoln, *Engineering data compendium: human perception and performance*, John Wiley & Sons, NY: NY, 1988.

[11] E. Hollnagel, *Cognitive reliability and error analysis method*, Elsevier Science, NY, 1998.

[12] S.G. Hart, "Pilot workload coping strategies", AIAA/NASA/FAA/HFS Conference on Challenges in Aviation Human Factors: The National Plan, Tyson's Corner, VA., 1991.

[13] S.R Hursh, D.P. Redmond, M.L. Johnson, D.R. Thorne, G. Belenky, T.J. Balkin, W.F. Storm, J.C. Miller, and D.R. Eddy, "Fatigue model for applied research in warfighting," *Aviation, space, and environmental medicine*, Vol. 75, No. 3, Sect. 2, pp. A44-A54, March 2004.

[14] A.A. Borbély, "A two process model of sleep regulation," *Human neurobiology*, Vol. 1, pp. 195-204, 1982.

4 Acknowledgements

The research presented herein is the result of NASA Office of Biological and Physical Research funding source UPN 107-07-07. The technical monitor of NASA Ames Research Center Grant # NCC2-1302 is Dr. Jessica Nowinski. The authors would like to thank K. Michael Dalal from QSS Group, for his programming assistance on this task and for all reviewers for their insightful feedback. Sandra Hart and Faith Chandler were instrumental in identifying various opportunities within the Space application domain for the current MIDAS applications and in promoting the refinement of the MIDAS tool.