MODELING HUMAN PERFORMANCE: IMPACTING SYSTEM DESIGN, PERFORMANCE, AND COST

Laurel Allender
Army Research Laboratory Human Research & Engineering Directorate
Attn: AMSRL-HR-MB
Aberdeen Proving Ground, MD 21005-5425, U.S.A.
lallende@arl.mil

KEYWORDS

Behavioral science, military, system analysis, system design, cost analysis.

ABSTRACT

Human performance must be modeled, and modeled early in order to impact system design, performance, and cost. This notion is consistent with the U.S. Army's push toward simulation-based acquisition and can be implemented through simple equations, stochastic task network modeling, or representation in force-on-force models. The underlying rationale and examples of such modeling are discussed. A key reason that human performance must be modeled is that the human component is probably the "noisiest" component in the system. The examples given here are all based on models developed with the capabilities present in IMPRINT (the Improved Performance Research Integration Tool), developed by the Human Research and Engineering Directorate of the U.S. Army Research Laboratory.

HUMAN PERFORMANCE MODELING

Traditionally, for the military, system acquisition has been about performance effectiveness and cost (acquisition, operations, and support costs). Somewhere along the line, it became clear that attending to design early ultimately benefits both performance and cost. The case will be made here that perhaps the most critical component to be considered--and the earlier the better--is human performance. In this paper, the importance of using human performance modeling in order to impact system design, performance, and cost is addressed. The underlying rationale and examples of such modeling are discussed.

First, it is useful to define human performance modeling. Simply put, it is "modeling of the processes and effects of human behavior" (Pew and Mavor 1998). It can comprise such things as a detailed representation of human memory, 3-D anthropometry, a network of tasks, or the collective actions of a small team or large force. In other words, human performance modeling covers an incredible range.

The mechanisms for human performance modeling are likewise diverse. A simple, written equation or mathematical statement can represent the time to move one's hand to a target location or the limits of short-term memory. An array of measurements of human body dimensions tied to the biomechanics of movement, all running in a complex computerized 3-D graphics simulation, can depict the physical aspect of human reach, fit, and access. A stochastic network simulation of human tasks running on a PC can model human performance from milliseconds to days, from simple task actions to complex decisions under the conditions of stress, from the user-display interface to the impact of the human on overall system performance. Human performance modeling can also refer to a description of aggregated team performance running as a part of a distributed interactive simulation, or human performance modeling can be elements of all these human performance models running in an integrated architecture.

WHY USE HUMAN PERFORMANCE MODELING TO IMPACT SYSTEM DESIGN, PERFORMANCE, AND COST?

The broad use of models and simulations is being encouraged in the military at this time, the catch phrase being "simulation-based acquisition" (SBA). SBA is really the use of any sort of model or simulation as an aid or guide to the decision-making process throughout a system's lifecycle. SBA can be viewed as a cost-cutting approach to the design and procurement of multi-million-dollar systems, as a logical application of the power of computers to design, and as a way to extend our thinking and understanding of a proposed system's concept, design, use, strengths and weaknesses, performance effectiveness, and cost. Currently, much emphasis is on large-scale

Allender, L. (2000). "Modeling Human Performance: Impacting System Design, Performance, and Cost." In M. Chinni (ed.) *Proceedings of the Military, Government and Aerospace Simulation Symposium, 2000 Advanced Simulation Technologies Conference* (Washington, D.C.) 139-144.

distributed interactive simulations; however, as the examples to be presented here will show, PC-based models and simulations can be used to great advantage. SBA has been mentioned here in a military context, but clearly, the general approach has much wider implications.

So, why model human performance to impact system design, performance, and cost? It must be modeled because human performance is probably the noisiest, most variable "component" in the system. With respect to system design, human performance varies not only as a function of the design itself (e.g., the hardware and software interface, the amount of information displayed, the size or weight of the equipment), but also as a function of the context or outside environment in which the system will be used (day, night, urban terrain, extended operations) and as a function of a myriad of unobservable, internal states (e.g., aptitude, cognitive workload, stress, fear, motivation). Thus, understanding human capabilities and limitations is imperative for creating an effective system design.

With respect to overall system performance, human performance must be considered in order to examine the best and worst case scenarios, to examine the full range of system performance. It is the last link in system performance, the last "fail-safe" opportunity, yet the array of possible errors, failures, and mistakes is far greater for the human component than for any other component in the system. If the only models of performance are of the stress tolerance of hardware parts, the execution time of software, or reliability of an engine, or at the other end of the spectrum, the expected probability of kill of a weapons round on impact or the number of messages transmitted and correlated across sources per unit time, the most vital system component (that is, human variability) has been ignored.

The cost of a system is initially viewed as the cost of the acquisition itself, then of the operations and support costs. However, a full understanding of the cost of a system must include a breakout of the human costs. For example, soldier operations and support costs include such things as recruitment, training, medical support, and retirement in addition to salary, and, in times of deployment, basic sustainment costs. These costs increment for each crewmember and required maintenance or support person. The way in which modeling soldier performance impacts acquisition costs is less direct: building, executing, and analyzing any model adds to the acquisition cost; however, this must be compared against other ways to obtain the same information. Some alternatives to computer-based modeling of human performance are to build prototypes or to conduct operational tests. Both are valuable, but both can be quite costly and must necessarily occur later in the acquisition cycle than the earliest modeling efforts. By definition, prototypes and operational testing are too late for the most cost-effective impact. Human performance must be modeled early in the system design process. Early course-correction is less costly.

Confounding the importance of modeling human performance early when it can have the greatest impact is the dilemma of all modeling: that early data are the least mature. Nonetheless, the answer to the question "When should human performance modeling be used?" is "Early, and often in order to ascertain the basic system requirements, the precursor to design, and the role of the human." A classic example from the U.S. Army concerns what is now the Comanche helicopter (Aldrich *et al.* 1989). It was first envisioned as a one-man scout and attack helicopter, but early assessments of expected task performance, mental workload, and reliance on automation (all based on a fundamental understanding of human performance and psychology) proved the basic concept to be unworkable. Today, Comanche is being built as a two-man helicopter, a change which would have been impossible to effect later in the acquisition process without a complete and costly re-design.

Later, as the system concept matures and the design work actually begins, human performance modeling can be more detailed, tied to specific design features of a system. Modeling during design should be iterative, used as a part of the continuous design cycle and tradeoff process. Still later, human performance modeling can be used to both refine and extend testing or even to explore emerging concepts of tactical employment. The refinement or focusing relies on previous modeling efforts to focus on critical issues. The extensions to testing come by testing conditions and extremes that are not affordable or safe. To repeat, human performance modeling should be used early and often during the system acquisition process for maximum impact on design, performance, and cost.

EXAMPLES OF HUMAN PERFORMANCE MODELING IMPACT

In this section, examples are given of the impact on system design, performance, and cost of a specific human performance modeling tool, IMPRINT (the Improved Performance Research Integration Tool). IMPRINT, developed by the Human Research and Engineering Directorate (HRED) of the U.S. Army Research Laboratory (ARL), is a human-system task network modeling tool with specialized analytic capabilities embedded. (See the web site for more information about

Allender, L. (2000). "Modeling Human Performance: Impacting System Design, Performance, and Cost." In M. Chinni (ed.) *Proceedings of the Military, Government and Aerospace Simulation Symposium, 2000 Advanced Simulation Technologies Conference* (Washington, D.C.) 139-144.

IMPRINT, http://www.arl.mil/ARL-Directorates/HRED/imb/imprint/imprint.htm.)

The analytical capabilities in IMPRINT include human versus system function allocation, mission effectiveness modeling, maintenance manpower determination, mental workload estimation, prediction of human performance under extreme conditions, and assessment of performance as a function of varying personnel skills and abilities. While each of these capabilities is important in and of itself, their greatest utility comes from their application to actual systems and the consideration of the outputs in the acquisition and design process. In the examples that follow, the impact on design, performance, and cost is highlighted. Although for any given example the impact on design, performance, or cost might be emphasized, the three are, of course, highly interrelated.

Design Concepts

The first examples are still so early in the design process that we do not yet know how they will turn out. The designs here are not precise physical designs, but concepts of functional designs. In order to model concepts for a future U.S. Navy destroyer bridge, an advanced task network workload modeling capability available in IMPRINT and also available in the commercial product, WinCrew, was used (Archer *et al.* 1996). Four models were built: a baseline and three future design concept models representing a reduction from nine to three personnel on the bridge, sophisticated task workload management strategies, and the addition of specific automated technology to take on some of the previously manual tasks. It was demonstrated that the design concept of a crew size reduction could be successful but only when coupled with care in managing workload and automation. As of this date, the future destroyer is still in development, with exact crew size for the bridge, indeed for the entire ship, still under study. However, this modeling effort will serve as a benchmark or reference point for future analyses.

Another example of the application of IMPRINT's task network modeling capability was the modeling of the Land Warrior system (Adkins *et al.* 1996). (Note: The actual analysis was completed with the DOS-based predecessor to IMPRINT, HARDMAN III (Hardware versus Manpower III). For this model, data from a test of developmental equipment were used as a basis to extrapolate to the future Land Warrior design concept, a proposed clothing and equipment ensemble for the U.S. Army infantry soldier (e.g., cooling devices, new communication technology, global positioning equipment). While the specific design could not be modeled, the expected functionality could be. Fit and comfort of infantry clothing and equipment are evaluated as a matter of routine; however, the ability of the ensemble to support functional performance on the battlefield is often overlooked. This analysis was used to assess the concept of employment (as differentiated from specific design features). Model outputs indicated that while the squad leader could benefit from the addition of command and control and navigation equipment, squad members did not benefit. Today, these results are being folded into the modeling and simulation efforts of the U.S. Army's Future Warrior program, the continuation of the Land Warrior program.

From Concept to Design

Another noteworthy case of human performance modeling was done for the U.S. Army's Crusader system, originally the Advanced Field Artillery System (AFAS). At the inception of the AFAS program, HARDMAN III was used to evaluate issues of crew size, extended operations, and required personnel skill (Lowry 1993). The model predicted that it would take longer to complete a mission and more errors would be made with a two-man compared to the three-man crew for the examined missions. Extended operations were raised as a serious issue requiring further study, and tasks requiring high personnel skill were identified. Using the results of this early modeling analysis, the system concept was translated into a design for a three-man crew. However, even as the first modeling effort was being formally documented, because of other factors such as the overall weight of the system, a reduction to a two-man crew was being re-considered. This sort of re-thinking of a design is not uncommon; it is part of the ongoing tradeoff process that occurs as a design matures. Using the more mature design and scenario information available, the original analysis was updated using the IMPRINT software (Beideman *et al*., in press). This updated model predicted a similar increase in the time to complete a mission with a two-man compared to a three-man crew. Further, excursions from basic two- and three-man models were run to predict the likely decrement in performance expected under continuous operations (i.e., sleep deprivation) conditions. These results indicated that performance with the two-man crew dropped to near zero by the second day, whereas the three-man crew maintained an acceptable level of performance until the third day. In other words, the original design concept was maintained. It can be seen how modeling results can play an important, ongoing role in the evolutionary tradeoff process.

Detailed Design, Performance, and Cost

In contrast to using human performance modeling early in system acquisition to evaluate system concepts and to help formulate an initial design, human performance modeling can be used later, in conjunction with operational test and evaluation. Two examples illustrating using modeling iteratively with testing are the Apache helicopter on-board sideloader and the Fox, a nuclear, biological, chemical (NBC) reconnaissance vehicle.

For the Apache helicopter analysis (Allender *et al.* 1992), a field test was being planned of a proposed integrated loader for the 30-mm ammunition to replace a separate loader that had to be attached and detached each time. The test was designed to assess performance effectiveness and to support a "buy" decision. The integrated loader was expected to reduce the overall turnaround time (i.e., the time to reload ammunition and refuel) in the field by eliminating the "attach" and "detach" steps associated with the ammunition loading process. Two models were built in parallel to the test planning as a demonstration of modeling capability. To populate the model, soldiers were observed in the field using the current equipment and using one system that had already been outfitted with the proposed integrated sideloader. What the model comparisons showed was that, although the integrated loader reduced the time significantly compared to the current equipment, given the other requirements for loading missiles and refueling, additional changes in tactics and an increase in personnel would be required to achieve the overall target time for turnaround of 15 minutes. In other words, despite the predicted performance gain for the loading process, the equipment change was not sufficient to achieve the overall performance requirement. A footnote to this story is that the proposed field test was not held as scheduled and the insights gained from the modeling were no longer just a demonstration but were made available to the decision makers, in essence, substituting for the test.

The Fox NBC chemical reconnaissance vehicle example (also carried out using HARDMAN III) (McMahon *et al.* 1995) takes a system from a failed operational test, through a proposed redesign, and predictive task modeling, to a second operational test. The problem was this: A 4-man crew station was proposed to be manned by only three people--a reduction of personnel. This precipitated workload problems, critical task failures, and safety problems. One person now had to move between two different workstations while performing NBC reconnaissance on the move cross country: monitoring the soil sampling wheels at one station and monitoring and interpreting the soil analysis at the other. The result was that critical tasks were likely to be missed at both workstations. Further, it was dangerous to move between workstations while the vehicle was moving. An equipment re-design was proposed so that the single crewmember could monitor both stations without moving between workstations; however, before the re-design was accepted, task network models representing the baseline and the proposed redesign were built using available data from the operational test. The model comparison predicted a 12% improvement in performance. The next step was a second operational test to validate that the predicted performance improvement was sufficient to judge the system suitable and effective. The footnotes here are several: The second operational test was effective; the Fox model was formally validated and accredited for future use; the system was fielded; and the result was a 25% soldier cost savings over the system lifecycle. In summary, human performance modeling proved itself in its demonstrated impact on design, performance, and cost.

Design for Organizations, Manpower, Personnel, and Training

So far, the examples have covered conceptual design, translating concept to hardware and software design, and also design evaluation in conjunction with operational tests. However, design can also mean organizational design and the impact on manpower numbers, personnel types, and training requirements. In this section, two examples of IMPRINT analyses are given, one a look at unit maintenance manning, and the other an examination of the impact of technology insertion on manning, personnel requirements, and potential training impacts.

The U.S. Army, driven by simultaneous manning reductions and increased mission scope and area of responsibility, is currently evaluating new maintenance concepts. IMPRINT modeling was used to examine organizational and job design rather than the design of hardware and software or even personnel skill requirements (Allender 1999). The current structure in tank and armored vehicle maintenance is for each type of vehicle to have unique sub-system mechanics. One option being considered is to create a "sub-system mechanic" where maintainers would be assigned tasks on similar sub-systems for both tanks and infantry fighting vehicles. Another option is to create a "system mechanic" where one maintainer type would maintain all tank sub-systems and another would maintain all infantry fighting vehicle sub-systems. IMPRINT models were built for all three cases: baseline, sub-system, and system mechanic concepts. The results showed that the sub-system mechanic concept offered very little increase in operational readiness over the baseline but that the system mechanic concept provided substantially increased readiness, when an important assumption was met--that all the necessary tools and maintenance aids were in place. Another critical insight gained from the analysis was that the notion of 100% spare parts availability is of limited utility if the numbers of maintainers are highly constrained; the “wait queues” are simply moved from the parts queue to the maintainer queue. This sort of information, in combination with training and cost effectiveness analyses can provide a robust and valid basis for organizational and job design.

The planned insertion of new technology into an existing system creates the need for an assessment of the performance of the individual items and also of the impact on the overall system. An example of this is the Joint Base Station, an update to the communications center used by the U.S. Army Special Forces (i.e., the “system”). Manual components are being changed out for computer-based ones (i.e., the new technology). This will lead to changes in the procedures in the communications center, the amount and type of operator workload, and the potential for a change in the types of skills, abilities, and training required. The IMPRINT job, workload, and skill analysis (Malkin *et al.* 1997) indicated that a reduction of personnel from eight to seven was supportable, provided that a new job and skill type, appropriately trained, was introduced into the system as well. In short, the model provided quantitative support for the notion that a change in system components requires a change in personnel type to achieve an acceptable level of performance.

Equipment Design and Force Performance

The final two examples of IMPRINT human performance modeling show the link from design to system performance, indeed overall unit performance, and some associated cost implications. The two examples are of modeling done for the Air Warrior program and the Army’s new initiative to acquire a new brigade concept.

The U.S. Army’s Air Warrior program is an initiative to acquire an integrated ensemble for helicopter crews as opposed to piece by piece acquisition of the separate components (e.g., helmet, gloves, armor) as has been done in the past. Additionally, the effort is intended to be non-developmental to the greatest extent possible. Therefore, few specific design changes could be made; rather the emphasis was on equipment selection. IMPRINT modeling was used to help estimate the expected degree of unavoidable performance degradation due to specific equipment items. In other words, the addition of protective gear such as armor will necessarily cause some performance degradation; the idea here was to minimize that degradation. Models of Apache helicopter performance with proposed Air Warrior components were compared against the baseline (Salvi in press). This “stand-alone” information was fed routinely to the Air Warrior design and management teams. These standalone IMPRINT results were also fed into a force-on-force modeling effort (Perry *et al. 1999*). IMPRINT baseline and predicted Air Warrior levels of helicopter crew task performance were provided to the Combined Analysis and Support Task Force Evaluation Model (CASTFOREM). By including crew performance data, the CASTFOREM runs were more sensitive to overall system performance. Although this link was not made model to model, but through spreadsheets of processed data, the advance represented by this work is significant: Human performance modeling was used to help predict high-level force-on-force effectiveness.

The final example is still a work in progress. The U.S. Army has proposed the acquisition of a new fighting brigade concept, a new lighter weight force designed to meet the new missions of urban warfare and peacekeeping missions. Part of the acquisition includes the acquisition of a new vehicle platform intended to serve as the common chassis for multiple variants (e.g., infantry carrier, mobile gun system, scout vehicle). A quick turnaround effort was conducted by the ARL HRED using IMPRINT to demonstrate the effectiveness of human performance modeling and to provide some early insights into system issues. Modeling efforts ranged from detailed modeling of the time to reposition ammunition in a proposed mobile gun concept to the personnel skill demand expected in the reconnaissance, surveillance, and analysis squadron, and importantly to the total soldier cost of the acquisition. Despite some savings in individual crew sizes, the initial cost estimate is actually higher than current brigades of similar structure, largely because of the high cost of personnel required for analysis on the battlefield. In summary, this final example alone shows the link from concept to design, performance, and cost for emerging concepts and technologies.

REFERENCES

Adkins, R.; Murphy, W.; Hemenway, M.; Archer, R.; and Bayless, L. 1996. “HARDMAN III Analysis of the Land Warrior System.” ARL-CR-291. Army Research Laboratory, Aberdeen Proving Ground, MD (February).

Aldrich, T. B.; Szabo S,; and Bierbaum, C. R. 1989. “The Development and Application of Models to Predict Operator Workload during System Design.” In *Applications of Human Performance Models to System Design*. G. R. McMillan, D. Bevis, E. Salas, M.H. Strub, R. Sutton, and L. Van Breda, eds. Defense Research Series Volume 2, Plenum Press, NY, published in cooperation with NATO Defense Research Group.

Allender, L. 1999. “Modeling Maintenance Manning and Battlefield Effectiveness.” *10th ROK-U.S. Defense Analysis Seminar Proceedings*, Seoul, Korea, October 25-28.

Allender, L.; McAnulty, D. M.; and Bierbaum, C. 1992. “Using Simulation to Support Testing: Implications of a HARDMAN III Application." In *Proceedings of the 34th Annual Military Testing Association* (San Diego, CA). Military Testing Association.

Archer, R. D.; Lewis, G. W.; and Lockett, J. 1996. “Human Performance Modeling of Reduced Manning Concepts for Navy Ships.” In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting* (San Diego, CA). Human Factors and Ergonomics Society, 987-991.

Beideman, L. R.; Munro, I.; and Allender, L. (In press.) “IMPRINT Modeling for Selected Crusader Research Issues.” Army Research Laboratory Technical Report, Aberdeen Proving Ground, MD.

Lowry, J. C. 1993. "U.S. Army Research Laboratory Summary Report of the HARDMAN III Assessment of the Advanced Field Artillery System (AFAS) and the Future Armored Resupply Vehicle (FARV)." CAE Link, Falls Church, VA, October.

Malkin, F. J.; Allender, L. E.; Kelley, T. D.; O’Brien, P.; and Graybill, S. 1997. “Joint Base Station Variant 1 MOS-Workload-Skill Requirements Analysis.” ARL-TR-1441. Army Research Laboratory, Aberdeen Proving Ground, MD (November).

McMahon, R.; Spencer, M.; and Thornton, A. 1995. “A Quick Response Approach to Assessing the Operational Performance of the XM983E1 NBCRS through the Use of Modeling and Validation Testing.” In *International Test and Evaluation Association Proceedings*. (Huntsville, AL). 115-119.

Modeling Human and Organizational Behavior Application to Military Simulations. 1998. R. W. Pew and A. S. Mavor, eds. National Academy Press, Washington, D.C.

Perry, C.; Davis, B.; and Fields, J. 1999. “Air Warrior (AW) Analysis.” TRAC-WSMR-SB-99-012. TRADOC Analysis Center, White Sands Missile Range, NM (March).

Salvi, L. (In press.) “The Development of Air Warrior IMPRINT Performance Degradation Factors in IMPRINT.” Army Research Laboratory, Aberdeen Proving Ground, MD.

AUTHOR BIOGRAPHY

LAUREL ALLENDER is a Research Psychologist at the Human Research and Engineering Directorate of the Army Research Laboratory at Aberdeen Proving Ground, MD. Her research is concentrated on the modeling and analysis tool IMPRINT (Improved Performance Research Integration Tool) which has been transitioned to over 120 sites, applied to significant Army programs, and is related to similar Air Force, Navy, and foreign R&D efforts. Dr. Allender received a Ph.D. in psychology from Rice University, Houston, TX in 1987. She is a member of HFES, past chair of the Systems Development Technical Group of the HFES and of the Manned System Modeling Sub-group of the Department of Defense Human Factors Engineering Technical Advisory Group. During the summer of 1994, Dr. Allender was a guest researcher at the French Army Center for Human Factors in Angers, France. In 1997, she was appointed to the industrial engineering department of North Carolina Agricultural & Technology University as adjunct faculty.

Allender, L. (2000). “Modeling Human Performance: Impacting System Design, Performance, and Cost." In M. Chinni (ed.) *Proceedings of the Military, Government and Aerospace Simulation Symposium, 2000 Advanced Simulation Technologies Conference* (Washington, D.C.) 139-144.