# Simulating UAS: How Much Fidelity and Why?

Gerald P. Chubb

# SIMULATING UAS: HOW MUCH FIDELITY AND WHY?

Gerald P. Chubb
OSU Dept. of Aviation
Columbus, OH 43210-1110

The physical, functional, and operational fidelity of a simulation can impact design assessments, training device effectiveness, and the validity of research results done with desk-top simulators (or synthetic tasks). While difficult to quantify, fidelity issues need to be considered in each of these contexts as attempts are made to improve human system integration. This paper reviews some of the implications of fidelity, discusses current efforts to model the impacts training can have on performance, and outlines the kinds of empirical testing that could be done to compare Improved Performance Research Integration Tool (IMPRINT) model predictions of training impacts to actual performance.

## Introduction

Simulations of Uninhabited Aerial Systems (UAS) are needed for: a) design evaluation to support human system integration, b) design of training aids in concert with Instructional Systems Design / Development (ISD), and c) research on issues associated with team performance in this system context. Computational models such as IMPRINT are commonly used as part of MANPRINT and SEAPRINT (and in the future, perhaps AIRPRINT). Simulations are often later embedded in simulators: devices that allow human interaction with the simulation or dynamic model of some UAS (e.g., Schreiber, et al., 2002).

Simulators are commonly used as training aids, more often for operators but sometimes for maintenance training as well. Research related to system design and human-system integration issues rarely use simulators much less real systems. Some form of scaled down synthetic task environment or desk top simulation is developed to address fundamental human performance relationships. Recommendations based on such data are sometimes rejected by operational / maintenance stakeholders because they perceive the lack of fidelity in the synthetic task environments (STEs) used for the performance research (e.g., Gluck, 2005).

It is the author's belief that acquisition tradeoff studies require fidelity to a different degree than training device design or empirical research studies aimed at understanding the fundamentals of crew behavior in a systems context. Tradeoff studies during design can often be limited in scope and use some form of part-task simulation. Training devices typically require a richer task context, depending on the training objective being supported. STEs created to study human behavior in a laboratory setting with more realistic workload and team interactions (than typical academic research) often intentionally simplify the task to be performed.

Research on actual systems, using operational personnel has seldom been accomplished for a number of reasons. First, any such research is intrusive by nature, impinging on the operational unit's ability to conduct its mission. Second, systems are not typically instrumented to collect data, making it difficult to obtain useful measures of performance. Third, scenarios used in peacetime typically include restrictions that are not imposed in any actual wartime operating environment (e.g., low altitude restrictions or safe separation distances for safety of flight; limited emulation of enemy systems, etc.)

While issues raised here may apply to many other systems, the ideas expressed arose in the context of looking at what is being done in the Predator (MQ-1), an operational UAS that began life as an Advanced Concept Technology Development (ACTD) program, which allowed development to occur without the imposition of formal analysis and design requirements imposed on typical system acquisition programs.

Early Predator operation had to be done without any supporting training device. A Multi-Task Trainer (MTT) was later developed, largely by reverse engineering of the actual system. Only recently has a Predator Mission Aircrew Training System (PMATS) been developed to better support operator training. AFRL/HEA was tasked to enhance the capabilities of IMPRINT to treat the impact of training, in order to support trade studies conducted early in system acquisition. This paper reflects some of the lessons learned as that development has progressed.

## Background

Simulation has been defined by Law and Kelton (1991) as "… using computers to imitate, or *simulate* operations of various kinds of real-world facilities or processes." It entails modeling some real system, capturing selected aspects of its form and function. Computer modeling and simulation often occurs early in the design and development of aviation systems (particularly aircraft) in order to perform analyses that support trade studies to select the preferred alternative from a set of implementation options.

Simulation fidelity is typically assessed in terms of physical and functional fidelity. Physical fidelity requires the simulator to have (some of) the same controls, displays, and layout as the actual system. However, a simulator is of little use without some degree of functional fidelity. Functional fidelity refers to whether the simulation behaves like the real system. It only requires that the simulation has stimuli and responses that are functionally equivalent to those of the actual system, allowing that the physical features and layout may vary to some degree, so long as that does not adversely affect performance with respect to the purpose for doing the simulation.

Both kinds of fidelity can vary (independently) in the degree to which they are realized in any particular simulator, training device, or STEs. In research situations (and many training applications), it is believed that some degree of physical fidelity may be sacrificed, so long as an appropriate degree of functional fidelity is retained.

However, there are risks in doing that. Face validity is often compromised: operational personnel are not likely to see that the behavioral aspects of the task are retained even with the loss of physical or functional fidelity. They will simply recognize the layout is nothing like the real system, and they may not trust the experimenter to give valid design recommendations based on what they see as a simplistic task environment.

Operational fidelity is another useful concept, but one not typically mentioned in the literature. Since behaviors are largely event driven, constructing a scenario of mission events representative of operational situations can influence the quality and validity of trade studies, operator / maintainer training, or research study results. If the purpose of such simulations is to generalize to real operational contexts, then some consideration should be given to getting the number, type, pacing, and intensity of simulated events to match what operators / maintainers will face in real situations.

However, during training (prior to trade studies, for real operator / maintainer instruction, or for research studies), the level of operational fidelity will need to be systematically altered. The more complex and difficult the tasks to be performed, the longer the period of required preparatory training as the content of simulated scenarios is incrementally increased. Realistic levels of workload or other imposed stressors cannot be introduced before adequate levels of skill and proficiency are achieved.

Compromising any of these three kinds of fidelity is dangerous for two different but sometimes related reasons: a) selected parts of the stakeholder community may reject results or recommendations if fidelity is too low, and b) the validity of any empirical study results may be challenged when expected levels of fidelity are lacking.

User acceptance as well as predictive validity typically require some acceptable (though difficult to define) level of physical, functional, and operational fidelity. Problems that can arise out of not providing adequate fidelity include the inability to: 1) extrapolate research study results to real world situations, 2) get design recommendation buy-in from stakeholders, and 3) achieve legitimate training objectives (achieving positive & effective transfer).

Later, real-time dynamic mockups (simulators) are used to verify and validate implemented designs, sometimes performing experiments with surrogates of the system operators / maintainers (e.g., engineers, test pilots, or even human factors specialists serving as pseudo-operators). Finally, simulators may be built as training aids, most often for operational crews, but sometimes for maintenance personnel as well.

**IMPRINT and Its Enhancement**

IMPRINT (Archer, 1998) is a network-oriented, discrete event simulation package used by the Army as part of its MANPRINT program (Booher, 1990). It is used during system acquisition to model operator and maintenance crew performance, including assessments of workload, the impact of selected stressors, and the impact of training (and skill proficiency retention). To develop such models of crew activities, data are needed on the task times (and optionally, the accuracy of task performance). These data can come from a variety of sources (past experience, estimates based on micro-models of the tasks, expert opinion of task duration, etc.). Predictions based on such data can, at least in theory, be validated later when either a simulator or the actual system is available for empirical studies with real operators / maintainers (or suitable surrogates).

Two questions arise in any modeling or simulation effort: a) how detailed should the model be, and b) what fidelity is needed in a simulator? No one has a completely satisfactory answer to those questions. The variables involved are difficult to define much less measure. So our discussion here is simply limited to some of the perceived issues that one needs to address.

**Computer Models and Simulation in Support of Human-System Integration**

IMPRINT is supposed to help decide how to achieve human-system integration (Booher, 2003), across the life cycle. That may mean developing more than one IMPRINT simulation model. Early in design, crew sizing is an issue, and the model needed to address that issue may be quite different from the model needed to design the human-system interface(s) or the impact of training. The number and kinds of manpower specialties needed to operate and maintain a system affect how well it can be supported in the field during sustained military operations.

Manning questions that affect system design have to be answered during concept formulation, well before the details of interface design and task performance can even be addressed. That affects the level of detail such models can have. They must intentionally focus on larger, more global issues and leave the details of individual execution of specific tasks for subsequent design decisions. IMPRINT models at this level are not particularly detailed in their representation of tasks. They focus on multiple job categories / specialists in a variety of activities which are at best only modestly well-defined.

During the development phase, preliminary design focuses on which functions get allocated to hardware, software, and 'liveware' – people internal and external to the system itself. Modeling at this stage becomes more detailed and human activities, even specific tasks, better defined. Even then, only the nominal performance of such tasks is easily represented.

Until detailed design is complete, one cannot define malfunction modes, so the workload associated with reacting to those conditions is next to impossible to describe. Once such details do become available, the cost to redesign a system is often prohibitive: something of the proverbial 'Catch 22' where the information the designer needs to do interface design (considering realistic levels of workload under malfunction conditions) is not available until it is too late to change that design.

ISD rests upon the ability to not only define the tasks operators / maintainers are expected to perform, but determining which of those requires formal instruction, considering the conditions under which performance will likely occur. That, in part, assumes one can determine how often certain tasks will occur and what kinds of information will be available to support task performance.

Check lists and job guides support operators and maintainers in executing required procedures. In complex systems, some tasks are infrequently performed, so proficiency may decline after initial learning occurs. Some form of refresher or proficiency training is sometimes required to assure adequate proficiency is maintained.

**Treating Training and Its Impacts**

Models of training and its impacts could be developed several ways. IMPRINT's initial treatment of training estimated proficiency decay in task times and accuracy as a function of intervening periods of non-practice. If a task is not practiced / performed regularly, then one would expect subsequent performance to take longer and be less accurate. The longer the period of non-sue, the greater the expected proficiency decrement.

The literature suggests that the greater the degree of over-learning, the more robust proficiency becomes to this decay with non-use (Chubb, 2004). Archer (2006) describes empirical studies, algorithm development, and initial modeling of the Predator STE task (Gluck, 2005) in IMPRINT. The algorithms adjust task time and accuracy based on two levels of over-learning and two different retention intervals. To date, nobody has used IMPRINT to model such things as learning strategies or changes that occur with learning, such as different network branching variations that occur as people learn to do their tasks different, often more efficient or effective ways.

**Using IMPRINT to Consider the Impact of Training Investments**

However training impacts might be modeled in future versions of IMPRINT, two uses can be envisioned for examining the implications of improved training on operator / maintainer performance. The first is addressed in the context of system acquisition. The second is in the context of ISD or more specifically, curriculum design. In theory, ISD should proceed in parallel to, perhaps lagging, system design. However, there is little reason for it to follow system development and subsequent deployment, which is all too often the case now.

*Acquisition Applications.* During system design, IMPRINT could be applied either (or both) of two ways: 1) to determine how more / better human engineering could impact training requirements, and 2) how various training devices might facilitate effectiveness / efficiency of training. No one ever advocates poor human engineering, but the quality of any design enterprise to some degree rests upon how much funding is allocated to that part of a design team's effort.

In many cases, a proposed crew station gets accepted only later to be found deficient, leading to its redesign. Each version of that crew station interface has different training requirements / implications. Comparing the cost of one design to the other should consider not only how well they function in operation, but what the costs are to train for each design / configuration of that crew station.

The training community must consider the identified training requirements and then design an instructional system that adequately meets those defined needs. Today, that often occurs only after design is completed. Following Initial Operational Capability (IOC), the only training that often occurs is 'on system' (or on the 'one-of-a kind' simulators that may have been developed to support design / development activities) – since any proposed training aids or mission simulators are still 'in development.'

Use of the 'one-of-a-kind' trade study simulators for actual training may serve initial needs (e.g. flight test crews and the initial instructor cadre indoctrination), but these simulators typically would not be suitable for crew training applications unless these devices are replicated. Being unique, the trade study simulators typically do not get system design or software updates unless mandated by the customer or user community.

During system acquisition, primary focus is placed on the design of the system itself. Little if any attention gets focused on how design impacts training requirements or how simulators for training operators and maintainers could lower overall life cycle costs. Modeling could be used to examine the impact of lessened training requirements or providing better (more efficient / effective) training aids and strategies. Those trade studies could have a potentially dramatic effect, lowering the overall cost of ownership.

Part-task static or dynamic mockups can be less expensive than full mission simulators, but simpler devices also tend to be less effective, since they are more limited in what they can do: how they can be used to train. However, a mix of different kinds and types of simulators might prove more effective than trying to design a single device that meets multiple needs. That seems to be seldom considered

Too often, one and only one level of physical and functional fidelity is considered, and that may constrain the range of operational fidelity that can be attained. That one 'optimal' simulator is then built as the only training device, rather than considering some mix of appropriately configured training aids (mockups, part-task simulators, and full mission simulators). IMPRINT could be used to assess whether some mix of devices could be more cost effective than simply buying one device that 'fits all needs' – typically satisfying none very well.

It is commonly believed that until a system is fully designed, any attempt to build a training simulator is futile, since the details of how the system operates cannot be fully known. That is a myth that needs to be dispelled. During Preliminary Design, software development documents (SDDs) are generated that specify the functions that must be implemented to get the crew interface to interact with other embedded hardware functions (flight management as well as offensive and defensive avionics).

The SDDs (typically available at or shortly after Preliminary Design Review) have (or should have) sufficient detail to allow simulating a system well enough to have a dynamic mockup for operator / maintainer training analyses, prior to first flight of the aircraft. As design progresses, the mockup and its software can be upgraded to mimic approved engineering change proposals (ECPs).

Commercial Off the Shelf (COTS) software and non-certifiable mockups of crew station panels are sufficient for this purpose. They may not have complete physical fidelity, but they have nearly complete functional capability and can be used with a broad spectrum of scenarios with varied operational fidelity.

Such design simulators (or dynamic mockups) can serve at least two roles: 1) provide for independent verification and validation of the system and human engineering design / integration (and validate any models used to assess such designs), and 2) provide a prototypical simulator for assessing training aid /device utility (supporting ISD). IMPRINT at this stage can be used to support curriculum design trade studies: what methods should prove relatively more effective than others in meeting the training requirements imposed by system design, both for initial training and for any subsequent upgrade or refresher training.

A hypothetical acquisition test matrix for training impacts is presented in Table 1. The rows suggest two levels of human engineering, one more complete / satisfactory than the other. The columns represent levels of training aid or device design. The simplest level ignores provision of any training aids. The second level is some form of part-task trainer, which may not be complete representations. The third level provides simulators for all crew (operational and / or maintainer) positions: a full mission simulator.

The rows then represent two hypothetical levels of human factors engineering, one giving minimal attention to human system integration and the other representing more complete analysis and design. The cells in the matrix are then combinations of these two independent variables (level of design and training device complexity).

***Table 1.*** IMPRINT Training Impact Analyses During Acquisition.

| Human Engineering: | Training Device Complexity: None | Part-Task | Full Mission |
|---|---|---|---|
| Level 1: | Baseline System Design | Multi-Task Static Mockup | Crew Station Simulation System |
| Level 2: | Improved Interface Design | Multi-Task Dynamic Mockup | Complete Training System |

The upper left cell in Table 1 represents the baseline system design, with minimal (if any) human engineering, and no training aid: all training would have to be 'on system' once deployed. The middle cell in the top row combines minimal human engineering with only a static mockup for operator / maintainer training. This is adequate for familiarization and for procedures rehearsal, but not for time-paced procedures practice. The right most cell in the first row represents a complete mission-crew / maintenance simulator, but without any corresponding ISD / curriculum design.

In the second row of Table 1, greater emphasis is placed on design: first in the system itself and then for the training aids that support that system. It is assumed that improved interface design will reduce the training requirements, irrespective of subsequent training system or training device design. Dynamic mockups provide the ability to support individual and perhaps team rehearsal of event-paced procedures, something static mockups cannot do. This is perhaps more important for operational crew training than for maintenance. Finally, complete training systems design, to include full-mission simulation provides maximum support for the operational and maintenance crew training.

The question during system acquisition is which level of funding is justified in order to minimize life cycle ownership costs. Complex training devices can equal or even exceed aircraft costs. So, justifying the cost-effectiveness of such proposed training solutions becomes important. IMPRINT modeling might very well assist analysts in justifying such expenditures, or preclude investing in them, if unwarranted.

*Curricular Design Applications.* The other use of an enhanced version of IMPRINT could be applied during ISD in order to assess the relative merits of curriculum design, training strategies, and presentation methods. There are at least three areas of concern: 1) initial knowledge and skill acquisition, 2) declines in knowledge and proficiency with non-use, and 3) recovery with practice / re-use. IMPRINT is being modified to address such issues. This is a complement to and parallels its use during system acquisition.

The principal problem in such model development is the lack of supporting data. The literature does describe the rate of skill acquisition and variables that affect proficiency and retention. Less is known about rates of decay with non-use, and even less about recovery rates with subsequent re-use / practice / rehearsal. However, this problem is of growing interest, especially as the military uses joint forces where the combat teams are some mix of regular, reserve, and guard personnel, each of whom may have different levels or kinds of training and varied opportunities to practice learned skills under combat-like conditions.

Current attempts to collect empirical data have been based on synthetic tasks in a laboratory setting with surrogates for test subjects rather than subject matter experts using actual system operating or maintenance procedures. This presents a number of problems yet to be resolved. First, the synthetic tasks lack face validity to operational personnel: the data and predictions based on such data are suspect. Second, the synthetic tasks do not provide context validity: they do not include the full range of knowledge and skills operational personnel are expected to have. Third, there have been no studies of predictive validity. So, how well these data can be used to predict training impacts for real situations is unclear.

However, with all those deficiencies duly noted, some analysis is believed to be better than no analysis at all. Systems will be designed with or without the help of human systems integration specialists, and with or without any analyses of operator / maintainer performance. It is believed that the present attempts to enhance IMPRINT to treat training impacts is needed and useful. Validation of model predictions and extension of the database to support such modeling efforts is obviously necessary future research.

Two case studies seem warranted as the next step in validating IMPRINT training impacts enhancement efforts. First, the Ground Control Station for Predator (MQ-1) may be redesigned, improving the human engineering of the Air Vehicle Operator and Sensor Operator interfaces, which will change the training requirements. Concurrently, the current Multi-Task Trainer is being replaced by the Predator Mission Aircrew Training System. Studies of each of these efforts could fill two, possibly three cells of the test matrix presented in Table 1. Such case studies would go a long way toward convincing the acquisition and training communities that IMPRINT modeling can be useful.

## References


Archer, Rick (2006), "IMPRINT TRAINING DATA", DRAFT Technical Report, Alion Inc., Boulder, CO, 15 November.

Archer, S. (1998), Improved Performance Research Integration Tool (IMPRINT) Analysis Guide 4.0, US Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD.

Booher, Harold R. (1990), editor, MANPRINT: An Approach to Systems Integration, Van Nostrand Reinhold, New York, NY.

Booher, Harold R. (2003), editor, Handbook of Human Systems Integration, Wiley, New York, NY.

Chubb, Gerald P. (2004), "Extending IMPRINT to Treat Retention of Trained Performance", Air Force Summer Faculty Fellowship Program Final Report.

Gluck, Kevin (2005). Predator UAV Synthetic Task Environment. http://www.mesa.afmc.af.mil/UAV%20STE/Predator%20UAV%20STE%20Fact%20Sheet.doc

Law, Averill M. and W. David Kelton (1991), Simulation Modeling and Analysis, McGraw-Hill, Inc., New York, NY.

Schreiber, Brian T. Don R. Lyon, Elizabeth L. Martin, and Herk A. Confer, (2002), Impact of Prior Flight Experience on learning Predator UAV Operator Skills, AFRL-HE-AZ-TR-2002-0026, Warfighter Training Research Division, AFRL Human Effectiveness Directorate, Mesa, AZ.


## Acknowledgment


This paper was supported by the Air Force Research Laboratory's Human Effectiveness Directorate, Warfighter Readiness Research Division, Mesa, AZ under an Intergovernmental Personnel Agreement (award number GRT00000948). The ideas and opinions expressed in this paper are those of the author and do not necessarily represent the official position of the Department of Defense.