*Proceedings of the 1994 Winter Simulation Conference*
*ed. J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila*

# THE EFFECTIVE USE OF ANIMATION IN SIMULATION MODEL VALIDATION

Christopher L. Swider

Headquarters AFOTEC
Logistics Studies & Analysis Division
Kirtland AFB, New Mexico 87117, U.S.A.

Kenneth W. Bauer Jr.
Thomas F. Schuppe

Department of Operational Sciences
Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433, U.S.A.

## ABSTRACT

This paper documents a continuing effort to determine how animation is most effectively used for both model validation and model communication. This research examined how to effectively use animation to display invalid model behavior. The invalid behavior displayed by each animation was a violation of an explicit model assumption. Two animation displays (moving icons and bar graphs) were evaluated individually and in combination at two different presentation speeds. The communication ability of each animated presentation was measured both subjectively and objectively. This paper presents the objective performance of each animated presentation measured across eight different violations of model assumptions. Objective measurements included the subject's identification accuracy for each violation and their response times. The results showed that the moving icon animation was superior to the other two displays and that subjects identified violations sooner when the animation was shown at the slower of the two presentation speeds.

## 1 INTRODUCTION

In this paper, we examine different animations' ability to communicate invalid model behavior. Our investigation examined two factors (animation display and presentation speed) to determine the effect that each factor or combination of these factors had in communicating violations of explicit model assumptions.

This paper is organized in three sections. First, we discuss the use of animation in the validation of simulation models and relate one relevant display theory to the design of animation displays. Next, we describe the design and analysis of the animation experiment and the results obtained. Finally, we conclude with a summary of the results and a few guidelines for effective animation.

## 2 BACKGROUND

First, we present some background information on the use of animation in simulation model validation. Then, we describe the relationship between display theory and the design of effective animations of simulation models.

### 2.1 Animation Use in Model Validation

Schlesinger et al. (1979) define simulation model validation as "a substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model". The computer model is judged as valid if its users are confident that the model and its output are representative of the real system. Unfortunately, this judgment is not easily made. A variety of validation methods are often necessary to validate the conceptual model, the input data, and the operational validity of the model output (Sargent, 1992).

Computer animation is an emerging validation technique which has become increasingly popular with the proliferation of high quality, low cost animation software. Sargent (1992) defines animation (operational graphics) as "a graphical display of the operational behavior of the model over time". Carson (1989) feels that observing an animation reveals key model assumptions and helps establish their accuracy. Law and Kelton (1991) also feel that animation can increase model credibility, but they caution not to rely too heavily on a small sample of model output. Although animation cannot replace a statistical analysis of model output, it can enhance model credibility and provide convincing evidence that model behavior is representative of the system under study.

Animation vendors are its most enthusiastic proponents. They advocate the use of animation throughout the design process including verification, validation, and presentation of results (Henriksen and Earle, 1992; Haigh, 1992; Kalaskey and Davis, 1991).

They feel that animations highlight dynamic interactions between model entities and increase user understanding of model behavior.

Simulation consultants also value animation in their modeling efforts. Cyr (1992) states several advantages using animation including the ability to "demonstrate problems with the model itself which would otherwise be difficult to detect". Carson and Atala (1990) feel that animation is most useful in the validation of models with interacting entities. Zhao and Pirasteh (1991) find animation to be useful throughout the entire life cycle of their modeling efforts.

Animation vendors and consultants share an appreciation for the utility of animation. The key to the effectiveness of any animation is its ability to clearly portray model behavior. The clear depiction of model behavior requires an effectively designed animation display.

### 2.2 Relevant Display Research

Empirical evidence on effective animation designs is limited. Carpenter (1993) found the movement of icons to be more important than their detail or color in communicating the behavior of a simulation model with moving entities. Subjects identified displayed problems more accurately in less time when viewing moving icons, irrespective of the icon detail or color. Unfortunately, significant empirical results are limited to this research.

Although specific animation guidance is limited, other relevant research exists in the area of display design. One display guideline with implications for computer animations is the proximity compatibility principle.

The proximity compatibility principle relates information processing requirements to display design. According to this principle, information integration tasks require users to combine the information from more than one display. These tasks benefit from displays in which elements requiring integration are closely grouped. Focused attention tasks require users to concentrate on a single display. These tasks improve when displays are separate and distinct (Wickens and Andre, 1990).

For animation displays, the guidance is clear. To effectively compare/contrast the behavior of animated elements, they must be grouped on the screen. Wickens and Andre (1990) recommend grouping display elements with physical proximity, common color, or a common border. On the other hand, to accurately check the status of a single display, this display must be kept separate from the other animated activity. Each animated screen design must therefore balance the design requirements for information integration tasks against those for focused attention tasks.

## 3 METHODOLOGY

In this section we discuss the preparation and conduct of the experiment. The following sections describe our approach: 1) a description of the simulation model; 2) a discussion of the animation software and the resulting animation displays; 3) a summary of the model violations presented in the experiment; and, 4) a description of the experimental equipment and procedure.

### 3.1 Simulation Model

The simulation model we used is from a SLAM (Simulation Language for Alternative Modeling) textbook (Pritsker, 1986). We chose a simple model to provide our experimental subjects with a familiar and easily understood situation. The system models two-way traffic flow through a restricted one-lane road segment. The timing of a traffic signal at the one-lane segment compensates for the unequal traffic volume in each direction. The interarrival times for each direction of traffic are random variables. The light is timed to yield a low average waiting time for all cars. Figure 1 is a diagram of the traffic simulation model.

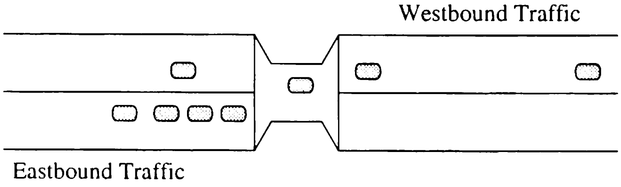

Figure 1: Traffic Analysis Model

We modified the original SLAM model for the requirements of the experiment. The interarrival distributions and light timing were different for each model violation scenario to give subjects a unique animation for each of the problems we presented to them.

### 3.2 Animation Software and Displays

The traffic model was animated using Proof Animation © (hereafter referred to as Proof). We created three different animations to present moving icons and bar graphs individually and in combination. Below, we discuss how Proof is used to animate a simulation model. Then, we will describe the design of each of the three animation displays used in this research.

Proof is a PC-based, "post-processing" animation software package. The animation is based on ASCII commands generated from state changes in the simulation and can therefore be used with a variety of simulation and programming languages. As a "post-processed" animation, Proof displays recorded simulation output and cannot be used interactively.

Two files are required to animate a simulation with Proof: a layout file and a trace file. The layout file describes how the background text and graphics will look in the animation. The layout file also contains blueprint information for the dynamic animated elements like the shape, size, color of objects and paths to guide their motion.

The trace file is the ASCII command file that drives the animation. It has the timing sequence for the animated activity and the commands to change the graphical display in accordance with the pre-recorded state changes from the simulation. Together, the layout and trace files provide all the information to animate any simulation.

We created three different animations to examine the moving icon and bar graph display formats individually and in combination. The animations were designated as follows:

- **CARS.** Car icons travel along the roadway, stopping at the traffic signal as necessary.
- **BARS.** Bar graphs for each road segment (inbound, middle, outbound) indicate the number of cars on each segment at any point in time.
- **ALL.** A combined presentation of both moving car icons and bar graphs.

Hereafter, we will refer to the animations as "CARS", "BARS", and "ALL". A static depiction of the combined (ALL) animation is shown in Figure 2. The CARS animation is the lower half of the animation in Figure 2 and the BARS animation is the upper half. Traffic lights are present in all of the animations.

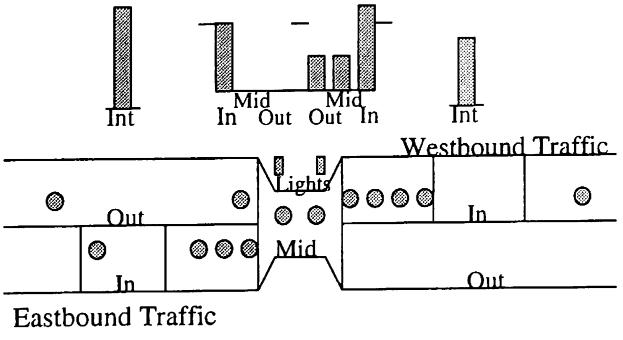

Figure 2. "ALL" Animation

The CARS animation was a birds-eye view of the model in operation. All cars had a common shape and color to help the user integrate information on traffic flow. Users compared the relative motion of cars to the status of the traffic signal to judge model validity.

The BARS animation was an alternate representation which focused on the changing level of traffic on each roadway segment. Bar graphs were grouped to help users compare and contrast the movement of traffic from one roadway segment to the next. Corresponding bars from each side had a common color to aid in their comparison. Interarrival bars ("Int") displayed the time between the arrival of successive cars from each direction. Users compared the levels and rates of change of the bar graphs with the status of the traffic light to judge model validity.

The ALL animation combined the icon and bar elements from the other animations. Users viewed both displays simultaneously to judge model validity.

### 3.3 Summary of Model Violations

There were eight violations of model assumptions in the experiment. Each violation represented a deviation from an assumption used in the development of the simulation model. The subjects, acting as supervisors, viewed each animation to see if the model operated in accordance with its stated assumptions. The violations of model assumptions (problems) are shown below.

1) Cars Pause - Each car briefly stops at a green light before proceeding.

2) Car Speeds Differ - Westbound cars travel twice as fast as eastbound cars

3) Encroachment - Cars enter the system with insufficient physical separation.

4) False Start - Cars move into the one-lane segment before the light turns green.

5) Constant Arrival - Car arrivals from one direction are NOT random.

6) Exit Queue - Cars from one direction wait in line to exit from the system.

7) Long Average Queue - The average queue lengths for each direction are distinctly different.

8) Two-way traffic - Traffic light timing allows two-way traffic on the one-lane road segment.

### 3.4 Experimental Equipment and Procedure

This subsection describes the software, hardware, and procedure used to collect objective performance data in the experiment. The animation software was Wolverine Software Corporation's Proof Animation Version 1.1. The animation ran on a 486DX/2-50 CPU and was displayed on a 14" VGA monitor. The experiment was conducted in an isolated room under ordinary fluorescent office lighting.

The experimental procedure included pretraining and two demonstration animations before objective data was collected. Each subject reviewed the model assumptions along with a description of the traffic model itself. A description of each animation display element was provided and all questions were answered. Subjects then viewed a demonstration animation to observe valid operation of the model. After a second review of model assumptions, subjects again viewed the demonstration animation and asked any final questions. Then each of the eight model violations (problems) was shown in a random order and the objective data was collected. The total time required for the experiment was 30 to 40 minutes.

## 4 MEASURES OF EFFECTIVENESS

We used objective and subjective measures of performance to determine which combinations of animation presentation and speed were best for displaying violations of model assumptions. This paper presents the results from the analysis of the objective data. Swider (1994) presents complete results for both measures of performance.

Objective performance data included the accuracy of each subject's identifications of problem behavior as well as how long it took to notice a problem. The response time indicated how much simulation output each subject viewed before detecting a problem. Each subject viewed up to 480 seconds of simulation output in each animation. If no violation was observed before the animation had finished, 480 seconds was recorded as a response time. Subjects viewed twice as much simulation output per second at the faster presentation speed.

## 5 RESULTS

Fifty-four graduate students and faculty from the Air Force Institute of Technology completed the experiment. We will present summary results for the response time and problem identification accuracy data first, followed by the Analysis of Variance results for response times.

### 5.1 Summary Results

Each subject viewed all eight problems under one combination of animation display and presentation speed. Figure 3 shows the mean response times (RT) at each of the six combinations of animation type and presentation speed. The first three bars represent the slower speed presentations of car icons and bar graphs (SA), bar graphs only (SB), and car icons only (SC). The final three bars represent the corresponding animations (FA,FB,FC) at the faster presentation speed. Figure 3 shows that subjects viewed significantly less simulation output at the slower presentation speed prior to their responses.

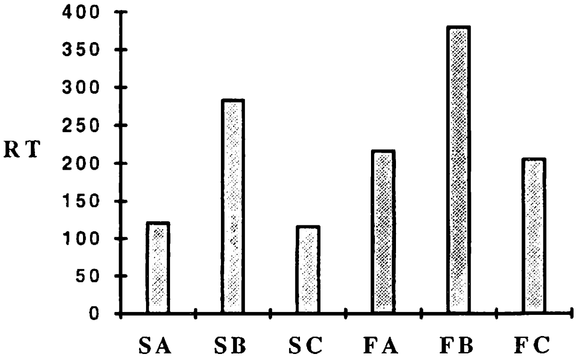

Figure 3: Average Response Time for Each Combination of Animation and Speed

Response times for each violation of model assumptions are shown in Figure 4. Response time was significantly different for problems 3 and 5 and similar for the remaining problems. Problem 3 (Encroachment) had the fastest response and problem 5 (Constant Arrival) had the slowest response.

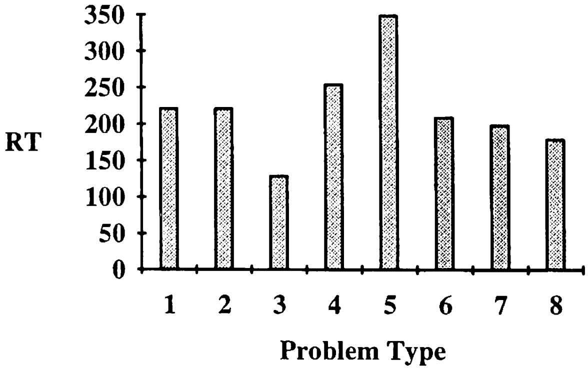

Figure 4: Average Response Time by Problem

Problem identification accuracy data is also shown in two ways. First, Figure 5 shows the identification accuracy for each problem at each of the presentation speeds. Problem 4 (False Start) was the only problem with a significantly lower accuracy at the faster presentation speed. Figure 6 shows the identification accuracy for each combination of problem and animation display. Identification accuracy was similar for animations with moving icons ( CARS, ALL ). Identification accuracy was generally inferior for BARS animations with no identifications at all for problem 1 ( Cars Pause ) and problem 2 (Car Speeds Differ).

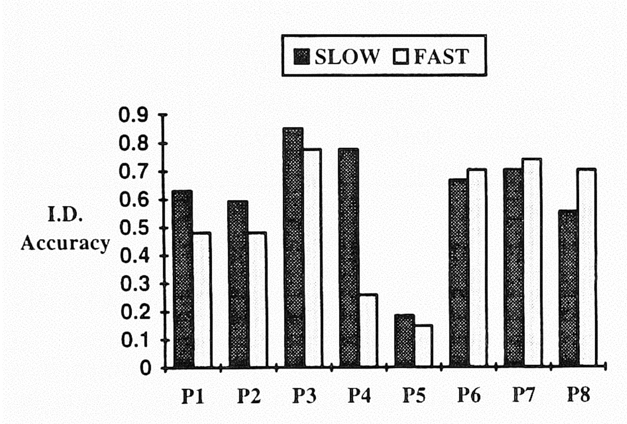

Figure 5: Identification Accuracy by Problem and Presentation Speed

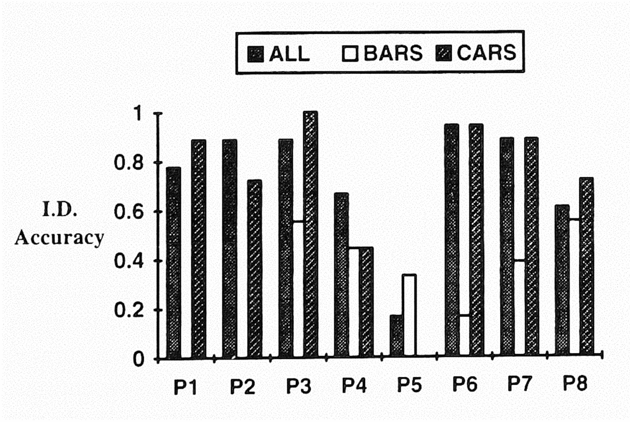

Figure 6: Identification Accuracy by Problem and Animation Display

### 5.2 Analysis of Variance Results

The experiment employed a 3 x 2 x 8 factorial design ( Animation Display x Presentation Speed x Problem Scenario ) with repeated measures on the third factor (problem scenario). Each subject viewed all problem scenarios, but only saw one combination of the other two factors. The analysis of variance for the repeated measures design was based on an analysis method developed by Winer (Winer et al., 1991).

The ANOVA results are shown in Table 1. The results indicated that presentation speed, animation type, and problem scenario all had a significant effect on response time. There was also a significant interaction between animation type and problem scenario.

Table 1: ANOVA Summary for Response Time

| SOURCE | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Between Subjects Factors | | | | | |
| Speed | 1 | 945378.9 | 945379 | 30.16 | .000 |
| Animation | 2 | 2690407.4 | 1345204 | 42.92 | .000 |
| Speed x Animation | 2 | 804.7 | 402.4 | .01 | .987 |
| Error Between | 48 | 1504530.3 | 31344.4 | | |
| Within Subjects Factors | | | | | |
| Problem | 7 | 1532057.9 | 218865. | 13.88 | .000 |
| Speed x Problem | 7 | 80937.7 | 11562.5 | .73 | .397 |
| Animation x Problem | 14 | 1010550.7 | 72182.2 | 4.58 | .015 |
| Speed x Animation x Problem | 14 | 207673.7 | 14833.8 | .94 | .398 |
| Error Within | 336 | 5299856.4 | 15773.4 | | |

Response times were greater at the faster presentation speed. Subjects had to view a significantly greater amount of each animation at the faster speed before they were ready to identify a problem. The faster presentation speed displayed each animation twice as fast as the slower speed presentation, making problems more difficult to see.

Problems were harder to see at the faster speed because there were more animated events to view each second. Table 2 shows the average number of discrete animated events (DAE) displayed per second for each problem. Discrete animated events included light changes, car arrivals, car departures, and car transitions from one roadway segment to the next. The majority of these animated events represented valid movement of traffic. Animations with twice as many animated events each second made greater perceptual and information-processing demands on each subject. As a result, most subjects required more exposure to each problem scenario before deciding that the observed behavior was abnormal.

Table 2: Discrete Animated Events (DAE) Displayed Per Second at Each Presentation Speed

| Problem # / Title | # of Discrete Animated Events (DAE) | Average DAE /second (Slow Speed) | Average DAE / second (Fast Speed) |
|---|---|---|---|
| 1 / Cars Pause | 432 | 3.6 | 7.2 |
| 2 / Car Speeds Differ | 432 | 3.6 | 7.2 |
| 3 / Encroachment | 468 | 3.9 | 7.8 |
| 4 / False Start | 432 | 3.6 | 7.2 |
| 5 / Constant Arrival | 432 | 3.6 | 7.2 |
| 6 / Exit Queue | 360 | 3.0 | 6.0 |
| 7 / Long Avg. Queue | 452 | 3.8 | 7.6 |
| 8 / Two-way Traffic | 636 | 5.8 | 10.6 |
| Overall Average | 456 | 3.8 | 7.6 |

The objective results also indicated an interaction between animation type and problem scenario. For problems displayed in the movement of many cars, response times were similar across all animation types. They included problem 5 (Constant Arrival), problem 7 (Long Avg. Queue), and problem 8 (Two-Way Traffic). These problems are shown in Figure 7 with dashed lines. The remaining problems were visible in the movement of one or a few cars. Subjects had shorter response times for the remaining problems when they viewed them with moving icon animations.

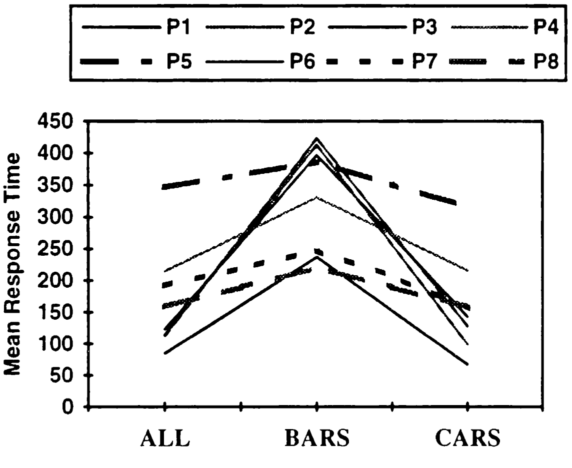

Figure 7: Interaction Plot of Animation and Problem Types for both Presentation Speeds

We used the Tukey method of multiple comparisons to determine if the differences between response times for each problem scenario were statistically significant. The mean responses were compared at a family $\alpha$ level of .05. The results of the comparison are shown in Table 3. Mean response times for problems with the same letter code (A,B, or C) were NOT significantly different. Problem 5 (Constant Arrival) had the longest response time and problem 3 (Encroachment) had the shortest.

Table 3: Tukey Multiple Comparison Results for Problem Scenarios

| Problem # / Title | Mean Response Time | Tukey's HSD Grouping | | |
|---|---|---|---|---|
| 5 / Constant Arrival | 349.5 | A | | |
| 4 / False Start | 254.0 | | B | |
| 2 / Car Speeds Differ | 221.6 | | B | |
| 1 / Cars Pause | 221.2 | | B | |
| 6 / Exit Queue | 209.0 | | B | |
| 7 / Long Avg. Queue | 198.3 | | B | C |
| 8 / Two-way Traffic | 179.2 | | B | C |
| 3 / Encroachment | 129.5 | | | C |

## 6 CONCLUSIONS AND RECOMMENDATIONS

In this study, animation was used to determine if a simulation model operated in accordance with its assumptions. Each subject viewed one of six combinations of presentation speed and animation type. Eight problem scenarios were presented, each one depicting a violation of a model assumption. Subjects viewed each scenario and identified violations as soon as they observed them. Objective measures included problem identification accuracy and response times.

### 6.1 Conclusions

Our objective results indicated that the slower presentation speed was superior to the faster speed and that animations with moving icons were superior to animations with bar graphs. A slower presentation speed resulted in significantly shorter response times with the same or better problem identification accuracy. For problems involving the movement of a few entities, animations with moving icons had significantly shorter response times with the same or better identification accuracy. For problems involving information integration, bar graph displays had response times and identification accuracy similar to those for the iconic displays.

### 6. 2 Animation Guidelines

Viewing an animation to discover invalid operation is a complex task. The viewer must understand the display elements, the relationships between these elements, and the appearance of violations of model assumptions if/when they occur. The results of this experiment and similar animation research (Carpenter, 1993), were combined with proximity compatibility display principles to offer the following six practical guidelines for effective animations.

1) **Use pictorial displays with moving icons for simulation models with moving entities**. Pictorial displays provide a concrete, intuitive representation of the model. Moving icons are best for models involving the interaction of moving entities (Carson and Atala, 1990; Carpenter, 1993; Zhao and Pirasteh, 1992).

2) **Design the display with the validation tasks in mind.** Focus attention on display elements by separating them from other displays. Integrate the information from individual elements by grouping them close together, within closed contours, or with a common color code. In this study, bars representing interarrival times were kept separate while bars representing traffic volume were closely grouped.

3) **Indicate problems with dynamic contrasts.** Problems with a subtle indication may be lost in the other animated activity. In this research, problems indicated with subtle contrasts were more difficult to notice.

4) **Set/adjust the presentation speed to make discrete differences visible**. Differences between significant discrete animated events should be visible in the animation. In this research, discrete differences of one-quarter second were significantly more difficult to notice than differences of one-half second.

5) **Avoid overloading the user with too much visual information.** A large number of discrete animated changes can overload the user. This overload can occur if a complicated animation is shown at a fast speed. If the user cannot comfortably scan the entire animation, he may focus on a small portion of the screen and miss important information. Subjects in this research ignored portions of the animation when they felt overloaded.

6) **Train the user to effectively scan the animation for potential problems.** Users will validate the model more effectively if they understand the animation display elements, the relationship between elements, and indications of potential problems in the animated output. This training will also prevent users from overreacting to normal animated behavior which they may not understand.

## REFERENCES

Carpenter, Michael L. 1993. Using Animation in the Validation of Simulation Models. Master's Thesis, Department of Operational Sciences, Air Force Institute of Technology, Wright-Patterson AFB, Ohio.

Carpenter, Michael L. and others. Animation: What's Essential for Effective Communication of Military Simulation Model Operation. In *Proceedings of the 1993 Winter Simulation Conference,* 1081-1088, Institute of Electrical and Electronic Engineers, Los Angeles, California.

Carson, John S. 1989. Verification and Validation: A Consultant's Perspective. In *Proceedings of the 1989 Winter Simulation Conference,* 552-558, Institute of Electrical and Electronic Engineers, Washington, D.C.

Carson, John S. and Onala M. Atala. 1990. Using Computer Simulation for Rapid Transit Operating Strategies. In *Proceedings of the 1990 Winter Simulation Conference,* 798-801, Institute of Electrical and Electronic Engineers, New Orleans, Louisiana.

Cyr, Rodney W. 1992. Using Animation to Enhance a Marine-Terminal Monte Carlo Simulator. In *Proceedings of the 1992 Winter Simulation Conference*, 1000-1003, Institute of Electrical and Electronic Engineers, Arlington, Virginia.

Haigh, Peter L. 1992. Using Mogul 2.0 to Produce Simulation Models and Animations of Complex Computer Systems and Networks. In *Proceedings of the 1992 Winter Simulation Conference*, 399-404, Institute of Electrical and Electronic Engineers, Arlington, Virginia.

Henriksen, James O. and Nancy J. Earle. 1992. Proof Animation: The General Purpose Animator. In *Proceedings of the 1992 Winter Simulation Conference*, 366-370, Institute of Electrical and Electronic Engineers, Arlington, Virginia.

Kalasky, David R. and Deborah A. Davis. 1991. Computer Animation with Cinema. In *Proceedings of the 1991 Winter Simulation Conference*, 122-127, Institute of Electrical and Electronic Engineers, Phoenix, Arizona.

Law, Averill M. and W. David Kelton. 1991. *Simulation Modeling and Analysis*, New York: McGraw-Hill.

Pritsker, A. Alan B. 1986. *Introduction to Simulation and SLAM II*, Third Edition. West Lafayette, Indiana: Systems Publishing Corporation.

Sargent, Robert G. 1992. Validation and Verification of Simulation Models. In *Proceedings of the 1992 Winter Simulation Conference,* 104-114, Institute of Electrical and Electronic Engineers, Arlington, Virginia.

Schlesinger, et al. 1979. Terminology for Model Credibility. *Simulation* **32**: 103-104.

Swider, Christopher L. 1994. The Effective Use of Animation in Simulation Model Validation. Master's Thesis, Department of Operational Sciences, Air Force Institute of Technology, Wright-Patterson AFB, Ohio.

Wickens, Christopher D. and Anthony D. Andre. 1990. Proximity Compatibility and Information Display: Effects of Color, Space, and Objectness on Information Integration. *Human Factors,* **32(1)**: 61-77.

Winer, B.J., Donald R. Brown, and Kenneth M. Michels. 1991. *Statistical Principles In Experimental Design,* New York: McGraw-Hill.

Zhao, Xu and Reza M. Pirasteh. 1991. Airline-Catering Plant Material Handling System Analysis with Simulation and Scaled Animation. In *Proceedings of the 1991 Winter Simulation Conference*, 402-409, Institute of Electrical and Electronic Engineers, Phoenix, Arizona.

## AUTHOR BIOGRAPHIES

**CHRISTOPHER L. SWIDER** is a suitability analyst for the Air Force Operational Test and Evaluation Center at Kirtland AFB, New Mexico. He received a B.S. degree in Electrical Engineering from the United States Air Force Academy in 1981, a M.S. degree in Engineering Management from the University of Dayton in 1991, and a M.S. degree in Operations Research from AFIT in 1994. His research interests are simulation, animation, and reliability analysis. His mailing address is: HQ AFOTEC/SAL, Kirtland AFB, NM, 87117.

**KENNETH W. BAUER, JR.** is an Associate Professor in the Department of Operational Sciences, Graduate School of Engineering, at the Air Force Institute of Technology. He received a B.S. in mathematics from Miami University (Ohio) in 1976, a M.E.A. from the University of Utah in 1980, a M.S. from AFIT in 1981, and a Ph.D. from Purdue University in 1987. His research interests are in the statistical aspects of simulation. His mailing address is: AFIT/ENS, WPAFB, OH, 45433-0538.

**THOMAS F. SCHUPPE** is the Dean, Graduate School of Logistics and Acquisition Management, Air Force Institute of Technology, Wright-Patterson AFB, Ohio. He received a B.S. in Mechanical Engineering from the University of Wisconsin, a M.S. in Systems Engineering from AFIT, and a Ph.D. in Operations Research from Ohio State University. His primary research interest is in simulation modeling of complex man-machine systems. He is currently a member of ORSA, the Air Force Association, and the Daedalions.