THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, *12*(4), 359–375

U.S. Army Initial Entry Rotary-Wing Transfer of Training Research

John E. Stewart, II and John A. Dohme
U.S. Army Research Institute

Robert T. Nullmeyer
Air Force Research Laboratory

Early fixed-wing research demonstrated that potential cost and training benefits could be derived from simulation-augmented primary flight training. More recent research in this area has been the exception, not the rule. This is especially true for rotary-wing aircrew training research. Currently, the U.S. Army does not use simulation in the primary (contact) phase of initial entry rotary-wing (IERW) training. Research performed by the Army Research Institute showed that a combination of synthetic flight simulation and criterion-based training during the primary phase of IERW had the potential for saving training time and costs in the aircraft. This research was performed using a low-cost simulator based upon the UH-1 helicopter. In the 4 quasi-experiments reported, positive transfer effectiveness ratios (TERs) were observed for most flight maneuvers pretrained in the simulator; student pilots in the simulator group required fewer iterations than control participants to reach proficiency on most flight maneuvers in the UH-1 training aircraft. As the visual display and flight modeling systems were upgraded, greater TERs were observed, and differences among groups tended to become significant.

Simulators are frequently integrated into training systems without evaluating their effectiveness. One possible explanation for this was proposed by Caro (1973), who observed that most personnel who design and integrate simulators are engineers, not behavioral scientists. Caro also expressed concern that much more attention has been paid to the development of the simulator itself than to the training program that supports it. Caro (1976) stated that those who integrate and employ

Requests for reprints should be sent to John E. Stewart, ARI Rotary Wing Aviation Research Unit, Building 5100 (Attn. TAPC-ARI-IR), Fort Rucker, AL 36362-5354. E-mail: stewartj@rwaru-emh1.army.mil

Report Documentation Page	*Form Approved OMB No. 0704-0188*

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
DEC 2002	**Journal Article**	**00-01-1999 to 00-01-2000**

4. TITLE AND SUBTITLE **U.S. Army Initial Entry Rotary-Wing Transfer of Training Research**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER **62202F**
6. AUTHOR(S) **John Stewart; John Dohme; Robert Nullmeyer**	5d. PROJECT NUMBER **1123** 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory/HEA,Warfighter Training Research Division,6030 South Kent Street,Mesa,AZ,85212-6061**	8. PERFORMING ORGANIZATION REPORT NUMBER **AFRL; AFRL/HEA**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory/RHA, Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ, 85212-6061**	10. SPONSOR/MONITOR’S ACRONYM(S) **AFRL; AFRL/RHA** 11. SPONSOR/MONITOR’S REPORT NUMBER(S) **AFRL-RH-AZ-JA-2002-0001**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Published in The International Journal of Aviation Psychology, 12(4), 359-375 (2002)

14. ABSTRACT
Early fixed-wing research demonstrated that potential cost and training benefits could be derived from simulation-augmented primary flight training. More recent research in this area has been the exception, not the rule. This is especially true for rotary-wing aircrew training research. Currently, the U.S. Army does not use simulation in the primary (contact) phase of initial entry rotary-wing (IERW) training. Research performed by the Army Research Institute showed that a combination of synthetic flight simulation and criterion-based training during the primary phase of IERW had the potential for saving training time and costs in the aircraft. This research was performed using a low-cost simulator based upon the UH-1 helicopter. In the 4 quasi-experiments reported, positive transfer effectiveness ratios (TERs) were observed for most flight maneuvers pretrained in the simulator; student pilots in the simulator group required fewer iterations than control participants to reach proficiency on most flight maneuvers in the UH-1 training aircraft. As the visual display and flight modeling systems were upgraded, greater TERs were observed, and differences among groups tended to become significant.

15. SUBJECT TERMS
Fixed-wing research; Transfer of training; Rotary-wing research; Training; Primary flight training; Aircrew training research; Simulation; Transfer effectiveness; Training effectiveness ratios;

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Public Release**	**17**	

simulators in aviation training systems are best described as artisans. They have shown considerable skill in getting simulators, even bad ones, to perform their training missions. However, their efforts do not provide a common knowledge base for the functional requirements of effective simulators. This leaves training system developers without much guidance.

Caro's observations of 2 decades ago are echoed by Salas, Bowers, and Rhodenizer (1998), who conclude that though simulation technology has undergone incredible evolution, the evolution of training has shown little progress. In short, evidence of training effectiveness is lacking for most military aviation simulation and training systems, because little research has been conducted. What has been said about fixed-wing aviation simulation can be stated even more emphatically for rotary-wing aviation. Hays, Jacobs, Prince, and Salas (1992) conducted an extensive review and meta-analysis of aviation training effectiveness research. Of an original total of 247 articles and technical reports, we were able to locate 26, which met the criteria of transfer of training (TOT) research (i.e., training was conducted in the simulator; training effectiveness was measured in terms of transfer to the aircraft). They found only 7 studies that pertained to rotary-wing training. A comparison of the fixed-wing jet to helicopter TOT studies showed that simulator-based training resulted in improved performance for jet student pilots, but not for helicopter trainees. The paucity of helicopter TOT studies precludes any interpretation of these findings.

HELICOPTER TOT RESEARCH

Rotary-wing simulation research has investigated the effects of functional systems such as motion cueing (Ricard, Parrish, Ashworth, & Wells, 1981) and visual display system characteristics (Kaempf & Blackwell, 1990; Kaempf, Cross, & Blackwell, 1989; Westra, Shepard, Jones, & Hettinger, 1987) on the performance of specific maneuver tasks (e.g., autorotation; shipboard landing). The training effectiveness of helicopter simulation in the context of the primary training syllabus has received scant attention. This is understandable, since the U.S. Army, which relies almost exclusively on helicopters for its aviation assets, does not use simulators for the primary (visual flight) phase of initial entry rotary-wing (IERW) training. It does, however, use simulators without visual display systems for the instrument phase of IERW.

This does not mean that Army aviation has not experimented with surrogate primary trainers and simulators; a pioneering study more than 3 decades ago showed promising results. Caro and Isley (1966) reported the results of an Army TOT experiment involving the Whirlymite, a lightweight, single-place helicopter tethered by an articulated arm to a ground-effects machine. Its handling characteristics were the same as those of a lightweight free-flying helicopter. They pretrained student pilots for either 3.75 or 7.25 hr on the device, and compared their flight training performance to that of a non-pre-trained control group. Results indicated that the two experimental groups were ready to solo significantly earlier than the control participants, though training time for subsequent evaluation flights did not differ. Checkride grades for the experimental and control groups were not significantly different, nor were they different for the 3.75 versus the 7.25 hr experimental groups. However, Whirlymite training was found to produce a significant reduction in eliminations of students from flight training.

Caro (1972) reported a quantitative, though not experimental, evaluation of the Army's then-new UH-1 Synthetic Flight Training System (SFTS), which consisted of a high-fidelity cockpit representing the UH-1 helicopter and a 5 *df* motion platform. The SFTS has no visual display system, which was considered unnecessary for instrument training. Caro reported that introduction of the simulator reduced instrument training time in the aircraft by approximately 90%. Prior to the introduction of the SFTS and the revamped instrument training program, 60 hr aircraft time and 26 hr in a modified Link 1-CA-1 Trainer were required to obtain the Army Standard Instrument Card. The introduction of the new simulator and revised training program reduced aircraft training time to 6.5 hr, supplemented by approximately 43 hr simulator time. The length of the instrument training course was reduced from 12 to 8 weeks.

We attempt to integrate four quasi-experiments, which comprised a program of research (Dohme, 1995) conducted by the U.S. Army Research Institute (ARI), for the Behavioral and Social Sciences, at its Fort Rucker, Alabama field unit. This research, which examined low-cost simulation in the context of IERW training, took place in the late 1980s and early 1990s.

THE TRANSFER EFFECTIVENESS RATIO

Before discussing the ARI research, it will be necessary to introduce the reader to a metric of training effectiveness, used primarily in aviation research settings. The transfer effectiveness ratio (TER) is a snapshot measure of the effectiveness of simulator pretraining expressed as the ratio of pretraining in the simulator to the savings achieved in the aircraft. It is calculated using the following equation (adapted from Williges, 1980):

$$\text{TER} = \frac{C_I - E_I}{E_{I\,(sim)}}, \qquad (1)$$

where C_I is the number of control group (no simulator training group) iterations to criterion in the aircraft, E_I is the number of experimental group (simulator pretraining group) iterations to criterion in the aircraft, and $E_{I\,(sim)}$ is the number of experimental group iterations to criterion in the simulator. TER is the ratio of savings in aircraft maneuver iterations to the number of iterations performed to criterion in the simulator by the experimental students (which is a training cost in that operation of the simulator and student pilot time expend economic resources). Examining the curves in Figure 1 reveals that successive sets of simulator iterations yield diminishing reductions in flight iterations. The cumulative transfer effect ratio (CTER) plotted in Figure 1 is a special case of the TER, when multiple combinations of simulator iterations or trials are used as the independent variable. The data in Figure 1 are strictly notional and are presented here for illustrative purposes. They are not based on any actual simulation research.

Roscoe (1971) proposed the CTER as a more dynamic measure of transfer effectiveness:

$$\mathrm{CTER} = \frac{(y_o - y_i)}{x_i}, \quad (2)$$

where x_i represents the number of iterations performed in a simulator, y_i is the number of iterations needed in the aircraft to demonstrate criterion performance after x_i simulator training, and y_o is the number of iterations that would be required in the aircraft if no simulator were available. Each point on the CTER function represents a TER. For the first point on the CTER function, 10 simulator iterations reduced iterations needed for students to demonstrate criterion-level behavior in the aircraft

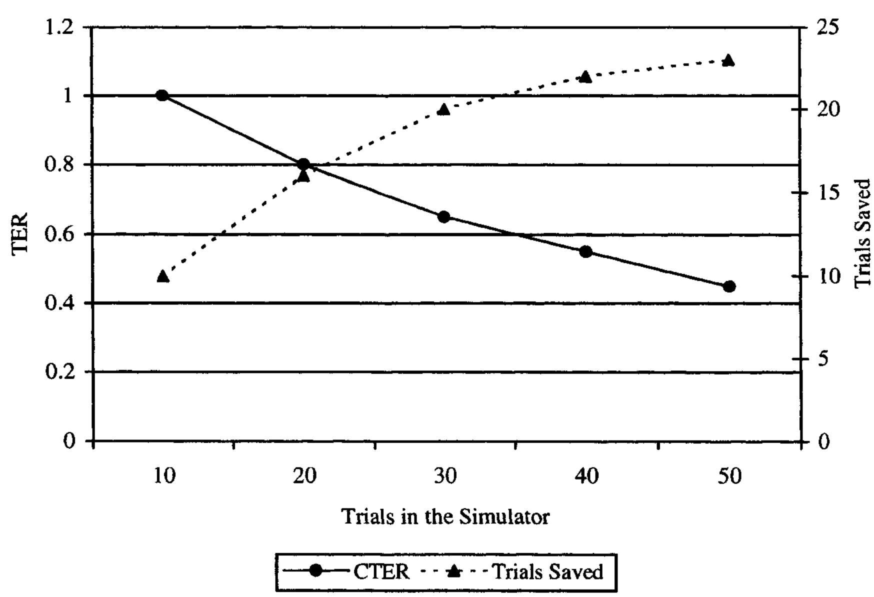

FIGURE 1 Effects of pretraining in the simulator on transfer effectiveness ratio (TER) and number of trials (iterations) saved in the aircraft.

by 10 in-flight iterations for a TER of 1.0. CTER is similar conceptually to TER. TER measures the effects of simulator pretraining for a particular number of trials or iterations; CTER is concerned with the cumulative effects (see Figure 1). The main distinction is temporal. The training researcher is concerned with the CTER when he or she is interested in the effects of training along a continuum of simulator trials as illustrated in Figure 1. The CTER conveys more information than the TER in that it illustrates the actual shape of the transfer effectiveness curve, whereas the TER simply conveys the effects of training at one particular point, usually the criterion for successful completion of that phase of training. In the illustration, after 40 simulator trials, 22 trials are saved in the aircraft, so TER = .55.

THE ARMY RESEARCH INSTITUTE IERW SIMULATION PROJECT

The Training Research Simulator

ARI developed the training research simulator (TRS) from an existing UH-1 SFTS instrument simulator. Three out-the-window displays were added, with 27 in monitors and collimating optics, providing a forward view for both pilot and co-pilot and a right-side view for the pilot. (Helicopters are typically flown from the right seat.) One image generator drove the forward visual displays; another controlled the right-side visual display. Besides a visual display system, the TRS had the high-fidelity cockpit, hydraulic control loaders, a seat-shaker and the 5 *df* motion base of the SFTS. The TRS first "flew" in 1986. During its life span, three types of image generators were evaluated for training effectiveness: a very-low-cost system (SGI IRIS 2400T; Silicon Graphics, Inc., Mountain View, CA), a low-cost system (BBN 120 TX/T; Bolt, Baranek & Newman Technologies, Cambridge, MA), and a moderate-cost system (ESIG 500H; Evans & Sutherland, Salt Lake City, UT). For the entire system, the maximum measured transport delay was 108 msec.

Since the newly developed simulator had unknown training capabilities, it was necessary to determine whether training in the UH-1 TRS provided any TOT to the UH-1 aircraft before it could be evaluated as a substitute for the aircraft. The TOT paradigm was chosen as the best metric to evaluate the effectiveness of the simulator, since it provided a separate estimate of training effectiveness for each maneuver and could serve to estimate the cost-effectiveness of simulator-based training.

Overview of the TOT Experiments

Four experiments (Dohme, 1991, 1995) were embedded in the IERW course structure, and in all of them IERW students served as participants. These were quasi-experiments, performed in an operational training setting, under conditions where rigorous experimental control was impossible to attain and where multiple upgrades and modifications to the simulator had to take place simultaneously. Likewise, the experimenter had little control over the availability and scheduling of student participants. Some constraints not under the experimenter's control, such as participant weight considerations when assigning student pairs to training aircraft, also compromised random assignment. These exigencies are typical whenever research is conducted in a field setting.

Eight primary phase maneuvers were selected for evaluation: takeoff to hover, hover taxi, hovering turns, hovering autorotation, normal takeoff, traffic pattern, normal approach, and landing from a hover. Hovering autorotation was chosen because it was the maneuver for which TOT to the aircraft was expected to be minimal. In the opinion of the investigators, this maneuver maximized the differences in the aerodynamic modeling and control strategies between the simulator and the aircraft; in short, it was a maneuver considered much more difficult to perform in the TRS than in the UH-1. In all four experiments, simulator motion was used. In Experiments 1 through 3, 2 weeks of training time was added to the experimental participants' syllabus to accommodate TRS training. This was done to assure that the student would graduate on time with his or her class, in the event that problems with the TRS necessitated additional training time.

There were other specific variations among these experiments. Experiment 1 was a process evaluation of the TRS to investigate TOT to the UH-1. Participants were student pilots who had completed primary phase (i.e., visual) training in the TH-55, which was soon to be replaced by the UH-1 as the Army's primary training helicopter. Experiments 2 and 3 were conventional TOT experiments, employing neophyte trainees. Experiment 4 was a substitution experiment in which 7.8 hr of UH-1 aircraft time was replaced with 9 hr of simulator time. Weather conditions on the flight line reduced the number of aircraft flight hours from the originally planned 9, precluding a simple hour-for-hour substitution.

The main performance metric in these experiments entailed iterations to criterion (i.e., proficiency) rather than the more customary flight hours to criterion. The criterion was defined as three consecutive iterations, which met the official training standards for a particular maneuver. Hours to criterion were considered an unsatisfactory benchmark because the exact number of hours each student pilot required to meet each training objective was unknown. The IERW program follows a fixed-hour curriculum in which every student pilot gets the same training regimen; if it calls for 0.5 hr of hover training on a given training day, each student pilot gets that amount of training. Therefore, it is impossible to review the students' logbooks and determine how many hours were required to meet the training standard for a given maneuver.

Flight grades were only used in Experiment 4. Their limited use in this research was based upon several reasons. In the Army IERW course, the daily training grade is a letter (A, B, C, U) whereas the pre-checkride (put-up) and checkride grades are numerical, based on a 100-point scale. The modal daily grade is a B. If the instructor pilot (IP) believes a grade of A, C, or U is warranted, he or she must justify this in writing. Furthermore, if a student fails a checkride and passes a subsequent one the grade on the second checkride is 70, regardless of performance. Consequently, grades were only considered appropriate for the substitution experiment, where the main dependent measures were traditional indicators of student progress through the program.

Experiment 1

Introduction. The simulator was configured with the very-low-cost image generators (limited to 300 polygons, 17 Hz update rate, no surface texturing nor moving models in the display database) and a National Aeronautics and Space Administration (NASA) "Uncle" aerodynamic model (Talbot & Corliss, 1977). The TRS was evaluated by selecting a random sample of 10 flight students who had already completed the primary phase of IERW training in the TH-55. All had soloed the aircraft and had demonstrated proficiency in a variety of emergency procedures as well as in the common rotary wing flight maneuvers. It was hypothesized that all these student pilots, who had an average 46 hr flight training, would demonstrate positive TOT from the simulator to the UH-1 since they had already demonstrated basic helicopter piloting skills. Students who had also graduated from TH-55 primary phase of IERW served as controls. They went directly to the UH-1 without any intervening training. The hypothesis was that the simulator-trained students would meet the criteria for soloing the UH-1, as written in the contact (UH-1) phase program of instruction, in fewer maneuver iterations than would the control group.

Method. Five IPs, all qualified to teach primary phase, were each given 2 students for 1 hr of training per day in the simulator to the same standards they would use on the flight line. All experimental participants were trained to standard on each maneuver in the TRS. A comparison group of 24 student pilots, drawn from the same class, served as controls. Iterations to standard in the aircraft served as the dependent measure for both groups. To ensure an unbiased evaluation, flightline IPs were not informed which of the students had been pretrained in the simulator. Age, gender, and demographic data were elicited via a questionnaire administered in a flight training class prior to the experiment. The main purpose of the questionnaire in all of these experiments was to screen out those potential participants who were rated (i.e., licensed) pilots. The mean self-reported age of the sample of 34 participants was 23.73 years (SD = 1.50); median age was 23.68 years (Mode = 22). The 8 youngest participants were

22 years old; the oldest, 28. All were men. Thirty-three participants described themselves as White; one, as African American.

Results. Three outcomes were possible for each of the eight maneuvers: (a) positive TOT, (b) no TOT, or (c) negative TOT. The actual results of the experiment provided all three outcomes depending upon the particular maneuver. Given the primitive nature of the visual imagery from the low-cost image generators, this result was not surprising. Another possible factor is the dissimilarity between the TH-55 and the UH-1. The former is a two-place, piston-engine helicopter, with direct control linkages and a manually operated throttle; the latter is larger, heavier, with hydraulically augmented controls and a turbine engine with a governor-operated throttle.

Student pilots required an average of 6.3 hr (SD = .82), or 102 iterations (SD = 16.77), to perform to standard in the TRS. The minimum number of TRS hr was 5; the maximum, 7; for iterations, the minimum and maximum were, respectively, 80 and 139. TER values from Experiments 1 through 4 are presented in Table 1. The data show that six of the eight maneuvers demonstrated a moderate positive TOT, one (normal approach) was essentially zero, and one (normal takeoff from a hover) demonstrated a small negative TOT. For Experiment 1, the average TER across all eight maneuvers was .18.

The same data are presented in another form in Table 2. *Iterations transfer ratio* (ITR) values are reciprocals of the TERs. The ITR is also called the *flight substitution ratio.* ITRs provide the number of simulator iterations required to save one iteration in the aircraft. It is a practical index of the amount of TOT expected in an applied training scenario, especially when relative costs are at issue.

Table 3 presents means and standard deviations for iterations to standard, for each of the maneuvers performed. Across all maneuvers, experimental participants required fewer iterations to standard in the UH-1 (M = 70.10; SD = 22.69)

TABLE 1
Transfer Effectiveness Ratios for Transfer of Training Experiments 1–4

	Aircrew Training Manual Maneuver Tasks							
Experiment	*Takeoff to Hover*	*Hover Taxi*	*Hover Turns*	*Hovering Autorotation*	*Normal Takeoff*	*Traffic Pattern*	*Normal Approach*	*Land from Hover*
1	.18	.32	.28	.26	–.05	.38	.06	.25
2	.19	1.12	–.002	–.28	.63	.43	.49	.72
3	.32	.87	.81	.50	.54	.82	.36	.53
4	.32	–.04	.47	.17	.13	–.04	.22	.59
Combined	.26	.51	.36	.16	.28	.37	.26	.52

TABLE 2
Iteration Transfer Ratios for Transfer of Training Experiments 1–4

	Aircrew Training Manual Maneuver Tasks							
Experiment	*Takeoff to Hover*	*Hover Taxi*	*Hover Turns*	*Hovering Autorotation*	*Normal Takeoff*	*Traffic Pattern*	*Normal Approach*	*Land from Hover*
1	5.6	3.1	4.2	3.6	Neg.	2.6	16.7	4.0
2	5.3	.89	Neg.	Neg.	1.6	2.3	2.0	1.4
3	3.1	1.1	1.2	2.0	1.9	1.2	2.0	1.9
4	3.1	Neg.	2.1	5.9	7.7	Neg.	4.6	1.7
Combined	3.8	2.0	2.8	6.3	3.6	2.7	3.8	1.9

than did controls (M = 87.83, SD = 39.99). This difference was not significant via a Mann–Whitney U test. (This nonparametric was used for all four experiments, due to departures from normality of distribution and homogeneity of variance.) An examination of Table 3 reveals that, for all maneuvers with the exception of normal takeoff, the experimental group required fewer iterations to attain proficiency in the UH-1. Differences, however, were not significant for any maneuver.

Experiment 2

Introduction. Experiment 1 demonstrated that, for most maneuver tasks, the low-cost simulator produced positive TOT to the aircraft with students already familiar with the basic primary phase flight maneuvers, though differences among groups did not attain significance. Based on these results, ARI upgraded the image generators and made improvements to the aerodynamic simulation model before conducting additional TOT research. The image generators were upgraded from the original IRIS 2400Ts to BBN 120 TX/Ts. This improved out-the-window imagery in four important ways: the frame update rate increased from 17 to 30 Hz; the displayed polygon count, from 300 to 1,000; the capability of surface texturing was added; and a realistic terrain model (Fort Knox, KY) replaced the primitive Flat Earth model. The same TOT paradigm was followed.

Method. A random sample of 10 students with no prior flight experience was selected. Twenty-four nonselected student pilots in the same flight class served as controls. Age and demographic data were collected as in Experiment 1. Participants' self-reported mean age was 24.21 years (SD = 2.16); median age was 23.50 years (Mode = 23). The oldest was 30 years of age. The 7 youngest participants were 22-years-old. Thirty-two were men. Twenty-nine described

TABLE 3
Mean Iterations to Standard in the Aircraft, Experiments 1–4

	Takeoff to Hover		*Hover Taxi*		*Hover Turns*		*Hovering Autorotation*		*Normal Takeoff*		*Traffic Pattern*		*Normal Approach*		*Land from Hover*	
Maneuvers	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Experiment 1																
Experimental[a]	8.40	3.69	7.00	2.94	8.20	2.70	9.30	3.97	11.40	5.12	9.80	6.68	9.20	4.39	6.80	2.10
Control[b]	11.38	9.25	9.88	5.86	10.17	6.27	11.46	6.04	10.88	6.52	14.29	9.86	10.75	3.93	9.04	4.27
Experiment 2																
Experimental[a]	19.90	11.74	20.50	9.22	26.10	11.03	21.60	5.04	18.30	7.33	18.30	5.89	16.80	4.98	15.70	6.38
Control[b]	22.88	8.17	29.25	15.42	26.08	11.88	19.50	9.62	23.63	7.81	24.08	9.66	26.00	9.85	23.38	10.51
Experiment 3																
Experimental[a]	14.00	9.68	13.80	12.68	15.70	10.83	15.30	10.02	11.70	10.23	13.20	12.86	14.40	10.23	14.90	9.20
Control[c]	22.62	13.28	25.81	9.63	24.48	10.70	19.00	10.66	18.81	9.50	23.52	13.00	23.33	10.35	20.24	10.34
Experiment 4 (substitution)																
Experimental[a]	11.20	4.24	20.90	10.57	18.30	8.96	12.80	9.58	16.80	9.67	19.50	11.48	19.30	11.28	13.30	9.31
Control[d]	17.79	8.46	20.37	13.40	24.42	13.53	14.74	7.71	17.63	9.43	18.89	11.17	23.58	10.33	19.05	9.20

[a]$n = 10$. [b]$n = 24$. [c]$n = 21$. [d]$n = 19$.

themselves as White; 3, as African American; 1, as Native American; 1, as Pacific Islander. Students were trained on the same eight target maneuvers. As in Experiment 1, training was conducted in the TRS by qualified Primary Phase IPs to the standard criteria for each maneuver.

Results. An average of 9.17 hr (SD = 3.08; minimum = 4, maximum = 12.70) and 91 iterations (SD = 29.13; minimum = 48, maximum = 133) were required per student to meet the criteria in the TRS. The TERs for Experiment 2 are presented in the second row in Table 1. The related ITRs are presented in Table 2.

Results from Experiment 2 revealed an average TER of .41, with one maneuver (hovering turns) displaying essentially zero TOT and one (hovering autorotation) displaying moderate negative TOT. It was not known if the improvement in average TER in Experiment 2 was either the result of improvements to the TRS or the result of using neophytes as participants. The degree of improvement in flight skills would be expected to be large for neophytes during the early training hours compared to that of the trainees from Experiment 1 who had already completed the primary phase of IERW training.

The total iterations to standard in the aircraft were compared via a Mann–Whitney U test. The experimental group required fewer iterations (M = 157.20, SD = 41.53) than the control group (M = 194.79, SD = 55.32). The difference approached significance ($z = -1.80, p < .07$). Table 3 presents means and standard deviations for each maneuver. Differences between experimental and control groups were clearly significant for normal approach ($z = -3.28, p < .001$), approaching significance for landing from a hover ($z = -1.92, p < .06$), hover taxi ($z = -1.68, p < .09$), and for traffic pattern flight ($z = -1.72, p < .09$).

Experiment 3

Introduction. Lessons learned from the preceding experiments were incorporated into the highest fidelity configuration of the TRS. Further improvements included the following: A two-channel Evans & Sutherland ESIG 500H calligraphic/raster image generator was installed to improve the quality of the out-the-window imagery. This increased the frame update rate from 30 to 50 Hz. A greater level of detail management and texturing as well as weather capabilities were added. A visual database was developed to model the Fort Rucker, AL, training area of operations.

Method. The same experimental procedure was followed as in Experiment 2: Iterations to standard in the aircraft were assessed for 10 student pilots pretrained in the TRS and for a comparison group of 21 student pilots from the same class, who were not pretrained in the simulator. Mean age for this sample was 23.03 years (SD = 1.52). Median age was 22 (Mode = 22), with the youngest 3 participants being 21-years-old; the oldest 3 participants, 26. Twenty-nine were White; 2 were African American. Twenty-nine were men.

Results. Experimental students were able to perform all maneuvers to standard in the TRS in a mean of 9.5 hr (SD = 1.78). The minimum time in the TRS was 7 hr; the maximum was 12 hr. For iterations in the TRS, the M = 119.10 (SD = 20.98), with a minimum of 93 and maximum of 168. The TER values for Experiment 3 are presented in the third row in Table 1; and the derived ITRs, in row 3 of Table 2. In this experiment, a positive TOT was observed for all eight maneuvers. The overall average TER was .55; average ITR was 1.82:1, suggesting that approximately two iterations in the TRS saved one in the aircraft.

Experimental student pilots required a mean of 113 iterations (SD = 73.39) to attain proficiency in the aircraft; for controls, the M = 178.05 (SD = 67.57). This difference was significant ($z = -2.35, p < .02$). Means and standard deviations for each maneuver appear in Table 3. Differences were significant for hover ($z = -1.99$, $p < .05$), hover taxi ($z = -2.60$, $p < .01$), hover turn ($z = -2.20$, $p < .03$), traffic pattern flight ($z = -2.08$, $p < .04$), normal takeoff ($z = -2.14$, $p < .03$), and normal approach ($z = -2.20, p < .03$).

Experiment 4: Substitution of Simulator Hours for Flight Hours

Introduction. At the request of the Commanding General at Fort Rucker, ARI performed a fourth experiment to evaluate whether the low-cost simulator could substitute for aircraft flight training. Experiments 1 through 3 had demonstrated that that the TRS could produce a positive training benefit if introduced as an adjunct to an existing fixed-hour training course. However, an adjunct training schedule offered no cost savings over the current IERW course.

Method. Since the results of Experiments 2 and 3 suggested that the average student pilot meets the training criterion for all eight target maneuvers in approximately 9 flight hr, Experiment 4 replaced 7.8 hr of aircraft time with 9 hr of simulator time. Because the ESIG 500 was on loan to ARI and had to be returned, the simulator was configured as it was for Experiment 2, using two BBN 120 TX/Ts for out-the-window imagery. Ten randomly selected IERW student pilots were trained in the TRS; 19 non-selected students served as controls. The mean age for the 29 student pilots was 24.47 years (SD = 2.42). Median age was 24 (Mode = 23) with the youngest being 19-years-old and the oldest 2 being 29. All were White; 27 were men. The main dependent measures of interest in Experiment 4 were not TERs, but instead measures of student pilot progress in training. It had already been demonstrated that the TRS produced positive TOT to the aircraft. The issue in Experiment 4 was whether training in it could substitute for aircraft training on an (approximate) hour-for-hour basis. The simplest measure of equivalency was whether any students in either the experiment or the control group were eliminated from training or set back to an earlier phase of training.

Results. There were no setbacks or eliminations in primary phase for either group. When flight checkride grades were compared at the end of primary phase, the control group had higher grades ($M = 85.47$; $SD = 2.72$) than the experimental group ($M = 82.40$; $SD = 6.29$). This difference was not significant ($z = -.67$, $p < .50$).

TERs and ITRs were calculated for the 10 experimental students. These are presented in Tables 1 and 2. A positive TOT was achieved on six of the eight target maneuvers; there were slightly negative TERs on the hover taxi and traffic pattern flight maneuvers. The average TER for Experiment 4 was .21. The experimental group performed a mean of 117.20 iterations ($SD = 40.33$) across all maneuvers, in the TRS (minimum = 51; maximum = 162). The same group required a mean of 132.10 iterations ($SD = 57.98$) to attain proficiency in the aircraft; for the control group, $M = 156.47$, $SD = 67.94$. This difference was not significant. Takeoff to a hover was the only maneuver where a significant difference was observed ($z = -1.98$, $p < .05$), with this difference favoring the experimental group. No other differences were significant.

Overall, the substitution experiment was considered a success because of the following:

1. No experimental group students were set back or eliminated.
2. There were no significant differences in primary phase grades.
3. Net savings in training costs were approximately $36,000 for 10 trainees.
4. Overall positive TOT was demonstrated.

Summary: Experiments 1 through 4

The four TOT quasi-experiments used random samples of 40 Army students as they progressed through IERW flight training. The remaining students in the same classes served as control participants. Conclusions should be generalizable to current and future Army trainees. The last row in Table 2 displays average ITR values for the four combined experiments. Although each individual experiment evidenced considerable variability across the eight target maneuvers, the combined results show that, overall, the simulator produced positive TOT on all maneuvers. Though the overall TERs may seem modest, one must take into account the greater training and cost efficiencies of simulators. The overall TERs indicate that, for the configurations used in these four studies, the TRS was most effective for training students how to hover taxi and to land from a hover. For both these maneuvers, two iterations in the TRS saved one in the aircraft. Not only are the hourly operational costs of helicopter simulators low when compared to those of the training aircraft, but more iterations of a specific maneuver can be performed within a given time frame. In the simulator, the student pilot does not have to take off and fly to the practice area to rehearse the maneuvers; if a maneuver is failed, the student can practice it again, without the distraction of other tasks.

We must again emphasize that the four studies were conducted in a field setting, and hence, some factors are inevitably confounded. For this reason, they had to be analyzed separately, and comparison among experiments must be undertaken with caution. Variability across maneuvers and experiments was probably attributable to changes in simulator configuration, differences in instructor pilot experience, skill level, attitude toward simulation, and differences in student pilot abilities. Still, it is interesting to note that more performance differences between experimental and control group participants became significant, with each upgrade to the TRS. The same trend was evident with regard to TERs.

One issue in the results of these experiments that might require clarification is the very large difference between the number of iterations required among student pilots in Experiment 1 to reach proficiency in the aircraft and those in the three subsequent experiments. The first cohort of student pilots reached proficiency in fewer iterations, even though the configuration of the simulator represented the lowest level of fidelity of the three examined. The reader should recall that the participants in Experiment 1 had completed the primary phase of flight training in the TH-55. Since they had previously performed all target maneuvers, one would expect transfer between the TH-55 and the UH-1 (as well as the TRS). Thus, these differences could probably be explained by the transfer of previously acquired flight skills to aircraft and simulator.

In Experiment 4, practice in the TRS was time limited. For this reason, not all experimental participants were able to perform all tasks to standard in the time allotted. This was the case for 3 student pilots. Looking at the instances of performances not to standard, it seems that the most challenging maneuver tasks were hover taxi (one), hover turn (two), traffic pattern flight (two), and normal approach (one). It is difficult to speculate, after the fact, the extent to which TERs would have been affected had all participants been allowed to practice to standard, as they had in Experiments 1 through 3.

Finally, it is noteworthy that there were no reported instances of simulator sickness in any of these four experiments. The reasons for this negative finding are unclear, though on a post hoc basis, it is possible that the low transport delay latency of the system could have partially accounted for the absence of symptoms.

DISCUSSION

Implications of ARI IERW Experimental Findings

The experiments described previously were conducted to evaluate the feasibility and practicality of training IERW students in the primary helicopter piloting skills using low-cost simulation. Since the research was conducted using random samples of Army IERW trainees as research participants with training embedded into the IERW program at Fort Rucker, the results are considered to be directly generalizable to that training program. Although the course structure has changed somewhat since this research was completed, the following conclusions are supported by this research:

1. Low-cost simulation is effective in training neophyte students in the basic flight control skills underlying helicopter pilotage.
2. Training in low-cost visual simulators can substitute for in-aircraft training with no significant loss in trainee performance. However, as was learned from Experiment 4, it may be necessary to provide more maneuver iterations in the simulator than in the aircraft to meet the same standards. The combined research data from Experiments 1 through 4 suggest that approximately three iterations in the low-cost simulator are required to replace one iteration in the aircraft.
3. Training in a low-cost simulator can show positive TOT to the aircraft, provided that the visual out-the-window scene and the aerodynamic flight model offer the trainee at least moderate fidelity. This is best illustrated by comparing the results of Experiments 2 and 3.
4. Improvements in the quality of the out-the-window visual scene such as more polygons displayed, textured surfaces, and faster scene update rates resulted in greater TOT.

Applications and Recommendations for Future Research

Currently, IERW uses the concept of the flight training class. All students are assigned to a class with each class following a fixed program providing multiple modes of instruction such as the classroom, procedures training devices, simulation, and flight training, all on a fixed schedule. Experience shows that students do not all learn the material and meet the training objectives at the same rate. In the IERW program, these differential learning rates are handled by mechanisms such as the setback in which a student is reassigned to another class at an earlier stage of training so as to repeat a portion of the curriculum.

There are disadvantages to the class-based (lockstep) program. A student who is learning rapidly and is ahead of the fixed curriculum is nonetheless required to stay with the curriculum. The authors know of numerous students who have already met the training objectives of a given training phase and yet continued to fly to meet the flight time requirement to complete that training phase. Although additional flight time has positive consequences such as improving the student's confidence and honing overall flight skills, it is expensive. In short, the class-based curriculum is optimized to benefit the slowest learners.

Enhancing IERW Training Effectiveness

The authors attempted to demonstrate the current status of the need for simulation-augmented IERW training and its potential advantages, especially with regard to savings in training effectiveness and efficiency. We also presented evidence to underscore the urgency of the need; it is difficult to find definitive data that could provide the guidance needed to develop and validate an effective simulation-based IERW training system. However, existing research appears promising, in that advantages of simulation for rotary-wing primary flight training were demonstrated. Unfortunately, these studies were conducted at different times with different simulator configurations, making generalizability difficult. Comparisons among different configurations were all made after the fact. Nonetheless, they do demonstrate that further research could provide Army aviation with an empirical database regarding the functional requirements of an IERW simulator and how it might best be integrated into a training program to improve training quality and save flight time and costs.

The Army still uses the UH-1 SFTS for instrument training. It is interested in more up-to-date, cost- and training-effective alternatives to these old simulators as well as in the potential training benefits that could result from the addition of visual display systems to primary training simulators. This could possibly allow a broader range of flight maneuvers to be trained in the simulator, including those contact maneuvers now practiced solely in the TH-67, the current primary training helicopter.

The technology behind the TRS is dated. A training simulator like it can be constructed today using PC technology. At this writing, a high-end PC with a visual image accelerator card can produce high-quality imagery for out-the-window scenes at a fraction of the cost of the image generators employed in these experiments. Likewise, software flight models of much higher resolution can now be developed to run on a PC platform at a relatively low cost. ARI, at its Fort Rucker Field Unit, is currently involved in the development and adaptation of such software flight models. It is also exploring the integration of the latest image generator technology to low-cost simulator platforms. ARI is in the process of undertaking a new series of TOT experiments, using a low-cost, PC-based simulator which represents the TH-67. These experiments will systematically compare the transfer effectiveness of the new simulator to that of the much older and more complex UH-1 SFTS simulators. The training of contact flight maneuvers in the new simulator will also be investigated, examining the same target maneuvers used in the earlier TRS studies. If the results of the planned TOT experiments are promising, the Army will likely incorporate simulator training into both the contact and the instrument phases of IERW training, thus reducing training costs by saving flight time in the aircraft.

REFERENCES

Caro, P. W. (1972). *Transfer of instrument flight training and the synthetic flight training system* (Prof. Pap. 7–72). Alexandria, VA: Human Resources Research Organization.

Caro, P. W. (1973). Aircraft simulators and pilot training. *Human Factors, 15,* 502–509.

Caro. P. W. (1976). *Some factors influencing transfer of simulator training* (Prof. Pap. 1–76). Alexandria, VA: Human Resources Research Organization.

Caro, P. W., & Isley, R. N. (1966). Helicopter trainee performance following synthetic flight training. *Journal of the American Helicopter Society, 11,* 38–44.

Dohme, J. A. (1991). *Transfer of training from a low cost helicopter simulator.* Paper presented at the Sixth International Symposium on Aviation Psychology, Columbus, OH, 4 April 29 May.

Dohme, J. A. (1995). The military quest for flight training effectiveness. In W. E. Larsen, R. J. Randle, & L. N. Popish (Eds.), *Vertical flight training* (NASA Reference Publication 1373 No. DOT/FAA/CT-94/83). Moffett Field, CA: National Aeronautics and Space Administration.

Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992) Flight simulator training effectiveness: A meta-analysis. *Military Psychology, 4,* 63–74.

Kaempf, G. L., & Blackwell, N. J. (1990). *Transfer of training study of emergency touchdown maneuvers in the AH-1 Flight and Weapons Simulator* (Res. Rep. No. 1561). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kaempf, G. L., Cross, K. D., & Blackwell, N. J. (1989). *Backward transfer and skill acquisition in the AH-1 Flight and Weapons Simulator* (Res. Rep. No. 1537). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Ricard, G. L., Parrish, R. V., Ashworth, B. R., & Wells, M. D. (1981). *The effects of various fidelity factors on simulated helicopter hover* (Tech. Rep. No. IH-231). Orlando, FL: Naval Equipment Training Center.

Roscoe, S. N. (1971). Incremental transfer effectiveness. *Human Factors, 13,* 561–567.

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *International Journal of Aviation Psychology, 8,* 197–208.

Talbot, P. D., & Corliss, L. D. (1977). *A mathematical force and moment model of a UH-1 helicopter for flight dynamics simulations.* NASA TM-73254.

Westra, D. P., Sheppard, D. J., Jones, S. A., & Hettinger, L. J. (1987). *Simulator design features for helicopter shipboard landings: II Performance experiments* (Tech. Rep. No. 87–041). Orlando, FL: Naval Training Systems Center.

Williges, B. H. (1980). Measurement of transfer of training. In Roscoe, S. N. (Ed.), *Aviation psychology.* Ames: Iowa State University Press.

Manuscript first received January 2000