TRANSFER OF TRAINING EFFECTIVENESS EVALUATION FOR U.S. NAVY DEVICE 2B35

William V. Hagin, Wallace W. Prophet
Seville Research Corporation
Pensacola, Florida

and

Edwin P. Durall
Control Data Corporation
Washington, DC

ABSTRACT

Training effectiveness of Navy Device 2B35, a computer image visual system used in Navy Advanced Jet Phase UPT, was evaluated utilizing a transfer of training design. Comparisons were made for two simulator groups and a control group for the Familiarization, Night Familiarization, and Carrier Qualification stages of training; for the Weapons stage training, comparisons were made for a single simulator group and a control group. Results are presented separately for the various training stages. Implications for Navy UPT, the Navy VTXTS procurement, instructor training, and for visual simulation in general are drawn.

I. INTRODUCTION

The U.S. Navy was the first of the military services to exploit visual flight simulation for Undergraduate Pilot Training (UPT). In 1972, a prototype wide-angle computer-image visual display system was installed on a 2F90 operational flight trainer (OFT) for the TA-4J Advanced Jet phase training aircraft. Following an extensive evaluation (1) of its potential for enhancing that phase of Navy UPT, decision was made to procure two additional devices. This allowed one such visual trainer at each of the Navy's three Advanced Jet training sites. The two production devices, slightly modified versions of the original prototype, were designated Device 2B35 and were put into operational use in 1976.

Coincident with their installation, the Advanced Jet syllabus was revised (2) to provide approximately six hours in the visual trainer during the Familiarization Stage (FAM), eight hours of device training at the start of the Weapons Stage (WEP), and a brief session for practice of carrier deck emergency procedures prior to Carrier Qualification (CQ). It is interesting to note that any students who do not receive the prescribed FAM stage device flights, due to device down-time or other factors, are given two extra aircraft rides. However, students missing the WEP stage device training are not provided any added air time.

Device 2B35 and its antecedent prototype are Computer Image Generation (CIG) visual systems representative of early CIG visual technology. The visual system presents a 60° X 210° field of view by means of three rear-projection screens, each of which is illuminated by a light-valve projector. The scene is full color. The instructor-operator console is relatively unsophisticated, having the visual trainer controls and repeater instruments, a three-monitor TV display of the visual scene, and a teletype printout of selected performance parameters, such as bomb distance data, carrier wire engagement, and the like.

The 2B35 visual system is integrated with a Device 2F90 cockpit. The 2F90 OFT was introduced in the late 1960s. It has a limited pitch, roll, and heave motion system and is used in teaching instrument flight procedures. Because the CIG visual system and the 2F90 device to which it has been attached are inseparable when used in training, further references in this paper to the 2B35 will include both the visual system and the host 2F90 OFT.

As they gained operational experience with the 2B35, the Navy personnel became increasingly concerned about the real training worth of the device. While the training personnel were generally convinced that there was positive training value from the 2B35 for WEP training, they were equally convinced that there were serious undesirable effects from device use during FAM training. The possibility of negative transfer in FAM was of particular concern, because the three sessions in the device were being used to replace two aircraft FAM flights.

Concurrently, there was growing interest DoD-wide in formal evaluation of the training effectiveness of all aircrew training devices, and of flight simulators in particular. This position was endorsed by the Chief of Naval Operations (CNO). It was in response to CNO direction, and in recognition of the user concerns over Device 2B35's training effectiveness, that the Chief of Naval Education and Training (CNET) decided to undertake a rigorous appraisal of the training worth of the 2B35 by means of a transfer of training study. The Navy subsequently contracted with Seville Research Corporation to conduct such a study.

The Transfer of Training Effectiveness Evaluation (T^2E^2) study was to be conducted as a joint Navy-contractor effort. Seville would be responsible for the study design, preparation of implementing instructions and material, data analyses, and reporting of results. The Navy retained responsibility for the actual administration of all instruction and data collection, although Seville was to maintain close surveillance over these Navy activities. Navy guidelines for the conduct of the study were that Seville develop an evaluation design which would assure a fair appraisal of the 2B35's training worth, but it was also required that the evaluation would be minimally disruptive of ongoing training activities.

This paper reports the results of the consequent T^2E^2 study. It summarizes the necessary preparatory activities, briefly describes the study design, reports the principal effects of device training, and evaluates the probable utility of the 2B35 within the Navy's Advanced Jet pipeline. It also makes certain observations regarding the use of visual flight simulation in future Navy Jet Undergraduate Pilot Training.

II. METHOD

The study began with a review of the content and management of Navy jet training and an assessment of the 2B35's potential for training. These activities were accomplished through review of Chief of Naval Air Training (CNATRA) syllabi and training materials, interactions with the CNATRA staff, and on-site observations of Advanced Jet training operations at Training Wing 3, NAS Chase Field, Texas, the site selected by the Navy for the study.

Device 2B35 Training Potential

It quickly became apparent that there were fundamental problems with the 2B35 that would have to be solved before the transfer study could begin. Observation of students and instructors flying the 2B35 convinced the study team that it was neither providing proper cues nor eliciting the correct responses. The two most serious problems, close-in ball and power control and carrier-deck trapping, were believed due to errors in basic modeling which made successful performance of these tasks nearly impossible. User suspicions of negative transfer on tasks involving these device characteristics appeared justified. Fortunately, a priority corrective action program initiated by CNATRA and engineers/programmers from the Naval Training Equipment Center provided assurance that these and other discrepancies could and would be corrected prior to the start of T^2E^2 data collection.

Having assurance that the 2B35 could be made to perform properly, it was then necessary to identify the training stages for which suitable image-generating data bases were available. The data bases for ongoing FAM and WEP training were of obvious potential use. In addition, although they were not being used in training, data bases for Night Familiarization (NF) and Field Carrier Landing Practice/Carrier Qualification (FCLP/CQ) were judged to be of potential use. There were, thus, four training stages, FAM, NF, WEP, and FCLP/CQ, which could be examined during the study without extensive new data base development.

Evaluation Design

The first evaluation design consideration was the selection of a basic training strategy that could be effected within ongoing Navy UPT. The second was the technical concerns underlying development of an experimental design as such.

Training Strategies. Two training strategies were entertained. One would utilize an individual, train-to-proficiency training regimen, while the other would involve a lock-step, constant-time training strategy. The first approach had the advantage of providing information about time or trials required for training and would allow direct computation of both transfer ratios (TRs) and transfer effectiveness ratios (TERs). Unfortunately, this approach could not be implemented without making major changes to the Navy training syllabus and established scheduling procedures, changes that the Navy could not support. Since the fixed-time training regimen was also the training regimen in being, it was selected for the T^2E^2. The disadvantages of this regimen from the research viewpoint were more than offset by its operational suitability.

Experimental Design. The experimental design was basically a three group, repeated measures design for the FAM, NF, and FCLP/CQ stages. It provided for three levels of the simulator time variable at each of these stages. At the WEP stage, however, because of limitations on sample size, the design called for only two major groups, a simulator group and a nonsimulator control group.

Training Treatments. The 2B35 training treatments were administered in the following manner. Each student scheduled to begin the FAM Stage, was randomly assigned to one of three FAM treatment groups: Group A, which received four 2-hour sessions in the 2B35 followed by five flights in the TA-4J; Group B, which received two 2-hour 2B35 sessions prior to its five training flights in the TA-4J; and Group C, the control group, which flew seven flights in the TA-4J and received no 2B35 training. This assignment procedure continued until each group had twenty subjects.

As the three FAM groups approached WEP training, half of each group was randomly assigned to a WEP A simulator group, while the other half was assigned to a WEP C nonsimulator control group. This allowed examination of FAM-WEP interaction effects. The WEP A group received four 2-hour 2B35 training sessions, while the WEP C group, of course, received none. Each group (N=30) then received seven WEP training flights in the aircraft.

For FCLP/CQ training, the original A, B, and C groups were reformed, with the A group receiving three 1-hour 2B35 sessions and the B group one such session. The C group, of course received no 2B35 training other than the brief carrier deck emergency procedures training session received by all students during CQ training.

Performance Measurement

The existing Navy UPT grading system evaluates student pilot performance on a four-point scale: Unsatisfactory=1; Below Average=2; Average=3; and Above Average=4. This grading system was quickly discounted as useful for study purposes due to its lack of discrimination sensitivity. For example, more than 90% of the maneuver grades given were "average." Such "normative" grades may be adequate for local management of students' daily progress, but they were judged not sufficiently sensitive to be of use in training program/treatment effects evaluations.

As a consequence, a performance measurement system involving instructor recordings of observed student control over specified maneuver-critical aircraft parameters such as heading, altitude, time, and airspeed was developed for T^2E^2 use. Since this measurement approach has been found by many investigators to provide reliable and discriminating data in flight training studies, it was deemed appropriate for this effort.

Separate performance recording booklets were developed for FAM, NF, and WEP. As will be explained shortly, FCLP/CQ performance was measured differently. For the FAM, NF, and WEP maneuvers, the instructor was required to observe and record at selected times/places within a maneuver whether the student was within prescribed tolerances for specific parameters. Some parameters were measured only once during a maneuver, while others were sampled two or more times. Figure 1 illustrates the procedure. In the example shown, the instructor noted Takeoff rotation airspeed to be 5-10 knots high, attitude as proper, and rate as being over-controlled.

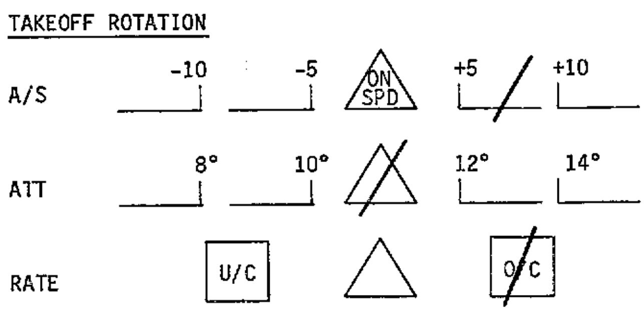

Figure 1. Sample Segment of a FAM Performance Recording Booklet.

Performance recordings were converted to error scores by considering all marks within the center triangles or squares "correct." Any other marks, left or right, counted as errors. By summing errors across all items in the maneuver, a maneuver total error score was obtained; by summing across parameters measured repeatedly, a parameter-control error score was derived. These maneuver and parameter scores were the primary data analyzed for the FAM and NF stages.

There was an additional source of data for the WEP stage. While the recording booklet was used to provide error scores for WEP pattern flying from the abeam position to the release point during 30° dive-angle practice bombing, clock angle and miss-distance scores were also obtained for six practice bomb drops on each training flight--device and aircraft. These bomb drop data thus provided a second data source.

The booklet recording procedure was not suitable for FCLP/CQ performance measurement, since an instructor is not on board the aircraft. The student is, however, under the radio control and guidance of the Landing Signal Officer (LSO) on all FCLP/CQ flights. Further, the LSO routinely observes each student's pattern and landing--in particular from final turn to touchdown--and records descriptive information about each pass using a standardized error notational shorthand. A procedure was developed whereby these data were routinely transcribed to Navy Landing Trend Analysis forms and subsequently converted to error scores for each student landing on each flight. Error scores thus derived were the error data employed for FCLP/CQ analysis. In addition, the LSO's subjective grades were also available for each of the 14 FCLP/CQ Flights, and these were analyzed separately.

Subjects

Students enter the Advanced Jet Phase of UPT in relatively small weekly or bi-weekly increments of from four to six per "class." Thus, to achieve the desired N of at least 20 subjects per FAM treatment group, it was necessary to use in the T^2E^2 all students who entered this training phase at Chase Field from June through September 1978. In all, a total of 64 students were involved. Of these, 59 were subsequently graduated as Naval Aviators.

Implementation

A number of special implementing procedures were required. Foremost among these were daily lesson guides which controlled the sequence and content of device instruction and defined the procedures to be followed during data collection. Second in importance were the instructor/manager training sessions conducted to assure standardized device training and recording booklet use. Last, but not least, was the close surveillance maintained by the research team over all T^2E^2 activities throughout the protracted training/data collection period which extended from June 1978 into March 1979.

III. RESULTS

The T^2E^2 study had three general objectives: (1) to determine whether Device 2B35 visual training produced demonstrable learning; (2) to discover whether such learning as was found transferred to the TA-4J aircraft; and (3) to identify those syllabus stages wherein 2B35 training would be of greatest utility assuming that meaningful transfer were obtained.

Learning in the 2B35

The FAM and WEP stages were the ones in which the device-trained group (Group A) received four consecutive 2-hour training sessions, thus allowing the clearest examination of across-trial learning. Using group percent error scores for each maneuver flown, learning curves were plotted for the four FAM maneuvers and for the WEP pattern data and bomb scores. All five curves reflected a substantial reduction in group percent error scores from the first to last trial. To illustrate, Figure 2 presents the Group A learning curves for the FAM Stage Full Flap Landing and the WEP pattern over the four simulator sessions.

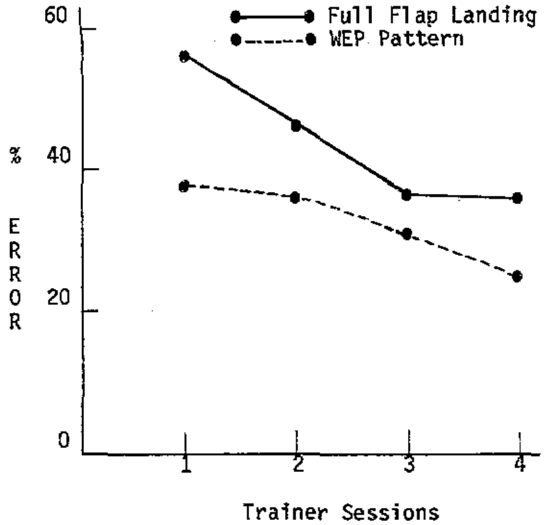

Figure 2. Percent Error on FAM Full Flap Landing and WEP Pattern by Simulator Session (Group A)

From these curves, it is apparent that learning did occur in the 2B35. The reduction in errors by the A Group was statistically significant for all of the FAM maneuvers but the Barrel Roll. For the WEP training, even though the reduction in errors from the first to the fourth simulator sessions was not statistically significant, the trend for that maneuver was also for reduction in errors. Also, that the WEP Pattern data did not asymptote, as was typical of the FAM curves, suggests that further WEP device training could be productive.

Since NF and FCLP/CQ device training was limited to two sessions for the A Group and only one session for the B Group, presentation of learning curves for NF is not useful. However, it should be noted that the NF A Group did show a statistically significant reduction in percent error score over the two device sessions for the Full Flap Landing, further evidence of learning in the 2B35.

During FCLP/CQ device training, both day and night Chase Field FCLP scenes were used for Groups A and B, but the carrier scene was used only during the A Group's training. As a consequence, device learning in this stage could be examined only by comparing Group A's night landing performance during their first and second 2B35 sessions. This analysis showed no significant difference over trials for the error scores derived from the Landing Trend Analysis forms. There was, however, significant improvement in LSO subjective grades over the two sessions, thus providing some indication that FCLP/CQ learning resulted from practice in the 2B35.

In summary, it appears that the types of visual tasks and maneuvers used in the T^2E^2 effort can be learned in the 2B35. While there were differences in learning rates and in absolute performance levels achieved across these maneuvers, there remains little doubt that the device can be used to learn visually cued maneuvers. The critical issue is, of course, whether learning in the 2B35 influences subsequent student performance in the TA-4J aircraft.

Learning in the Aircraft

The objective aircraft data for FAM, NF, WEP, and FCLP/CQ were analyzed as follows. The data from the FAM, NF, and WEP booklets and the derived error scores from the FCLP/CQ Landing Trend Analysis forms were analyzed on the Univac 1100 computer at the Arizona State University Computer Center. The program "Multivariance" was used for the univariate and multivariate analyses of variance, covariance, and regression. Supplementary analyses of these data were also performed on the Univac 1100 using programs from the Statistical Package for the Social Sciences. The .05 probability level was selected as the significance point for all analyses.

The WEP practice bomb score data were examined separately using univariate analyses and correlated t tests. These WEP data were analyzed first as circular error data and then, by trigonometric conversion, as the vertical and horizontal components of the circular error.

The LSO subjectives grades available from the FCLP/CQ flights were also analyzed by either univariate analyses or correlated t tests, as appropriate.

FAM Stage. The fact that the no-device time control group (Group C) received two more aircraft training flights than did Groups A and B allowed for two types of comparisons. Comparison one, the traditional transfer approach, compared the three FAM groups on their first, second, third, fourth, and checkride performances; this omitted the C Group's fifth and sixth training flights. In comparison two, the C Group's first and second flights were ignored, and inter-group comparisons were made between the A and B Groups' first flights against the C Group's third, the A/B second with C's fourth, and so on, up to the checkride. These arrangements are shown in Table 1.

TABLE 1

Device 2B35 versus Control Group Comparison Paradigms (FAM)

	Group	Flights						
Comparison 1	A/B	1	2	3	4			Checkride
	C	1	2	3	4	(5)	(6)	Checkride
Comparison 2	A/B			1	2	3	4	Checkride
	C	(1)	(2)	3	4	5	6	Checkride

Multivariate analyses (MANOVAs) for both FAM comparison arrangements yielded nonsignificant MVFs for treatments, but the MVFs across flights were significant for all maneuvers except the Straight-In Precautionary Approach (MVF=2.16; df=20.37; $p<.03$). While the nonsimulator C Group tended to make fewer errors than the two simulator groups (A & B), the performance of the control group across the seven aircraft training flights was not significantly different from the performance of the simulator groups which received only five aircraft training flights.

These results show, as would be expected, that all three groups improved with practice. However, the results do not lend any substantial support to a significant simulator-to-aircraft transfer effect, other than the fact that the simulator groups which received two fewer aircraft flights performed in a fashion not significantly different from that of the control group. On the other hand, neither do the results lend support to the converse inference that the all-aircraft training regimen was superior to the simulator-plus-aircraft regimen.

Unfortunately, subject limitations precluded having a fourth group which could have received only five aircraft flights and no 2B35 training. Such a group would have provided more conclusive evidence regarding the degree to which transfer from the 2B35 did or did not account for the A/B Group performance. Without such a group, the possibility remains that the FAM checkride can be "passed" as well with only five TA-4J flights and no device time as it is with seven aircraft rides.

Night FAM Stage. The NF error data from the three treatment groups yielded a significant MVF value favoring Group C (MVF=2.84; df=6, 110; p<.02). In view of the clearly better performance of the nonsimulator Group C, use of the 2B35 for NF training cannot be supported.

FAM Ball-Throttle Control. As indicated earlier, there were significant concerns over the power response and Fresnel Lens Optical Landing System (FLOLS) ball-tracking characteristics of the 2B35. Since substantial software and engineering changes to the 2B35 had been made to support the T^2E^2 effort, it was of special interest to determine whether the device might have negative transfer characteristics in these areas. The many power control and ball-tracking measures which had been obtained across the several FAM and NF Full Flap Landing maneuvers flown during the aircraft training flights allowed an examination of both treatment and practice effects of device training on these critical parameters.

Multivariate analysis of the data from the FAM stage showed no significant treatment effects. A similar analysis of the NF data indicated a significant treatment effect in favor of the A simulator group (MVF=3.00; df=12, 104; p<.01). The evidence from these two analyses slightly favors the trainer groups and, at the very least, supports a conclusion that subjects with 2B35 training time clearly were not inferior to their all-aircraft peers with respect to ball control and power control. In view of these outcomes, there is no reason to believe that the 2B35 training given A and B Groups resulted in any negative transfer in these task areas.

WEP Stage Error Data. The MANOVA of the WEP pattern error data yielded a significant treatment effect in favor of the simulator group (MVF=2.63; df=10, 47; p<.02). This was the most clear-cut area of simulator advantage in the study. The data also yielded, as expected, a significant effect across flights, i.e., a learning or practice effect (MVF=2.75; df=20, 37; p<.01).

The analyses of the error scores also showed a significant interaction effect between the two WEP treatments (A or C) and previous FAM/NF treatments (A, B, or C) (MVF=1.65; df=40, 74; p<.03). The pattern of this interaction suggests that, while WEP A treatment positively influenced WEP pattern performance overall, having had the previous 2B35 FAM/NF Group A training was slightly disadvantageous to later WEP flight performance, FAM/NF Group B membership had no clear effect on WEP performance, and FAM/NF Group C membership was slightly advantageous to later WEP performance. One may speculate that students and instructors who had already spent substantial time in the 2B35 (FAM/NF Group A) may have approached further exposure to the device for WEP training less positively than those being introduced to it for the first time.

WEP Stage Practice Bomb Scores. As previously described, practice bomb scores were first analyzed in terms of mean miss distance and then in terms of the mean horizontal and vertical error components of that miss distance. Univariate analyses of these three types of miss distances showed no significant treatment effects, though the A Group's miss distances tended to be somewhat less than those of the control group (Group C). The advantage of 2B35 training to the A Group that was apparent from the WEP objective pattern data analysis did not significantly influence their bombing accuracy, though the trend was for them to be more accurate.

Practice effects as shown by these bomb scores were interesting. While there was substantial improvement in overall miss distance over trials, the horizontal error component improved relatively little. The bulk of the improvement was in the vertical error components. Also, horizontal miss distances were significantly smaller than the vertical miss distances. These data suggest that bomb-run lineup is not a difficult skill, and consequently, there was little room for improvement in horizontal accuracy. On the other hand, the relatively greater vertical error component is related to the pattern elements such as dive angle and release altitude and seems to represent the more difficult aspect of the bombing task.

FCLP/CQ Stage. The objective FCLP/CQ error data were subjected to MANOVA analysis, and the subjective grades were examined by univariate analysis. Multivariate analysis of the LSO-derived error data showed no significant effects due to treatments, i.e., there was no evidence of any differences between simulator and nonsimulator training groups in the aircraft portion of FCLP/CQ. Further, there was no significant change in performance from early FCLP flights to the later flights, i.e., these LSO error data provided no evidence of FCLP/CQ skills learning in the aircraft.

While it is possible that there was no change in student performance over the 13 FCLP flights, i.e., no learning, it is more likely that the LSO error data reflect a shifting frame of reference as to what constitutes an error. This lack of discrimination in the LSO error data was disappointing in view of the usefulness of such data that has been reported by Brictson and Burger (3) in their evaluation of the A-7E Night Carrier Landing Trainer (NCLT).

As has been noted, the LSO subjective grade data were also available for FCLP/CQ flights.

These grades were subjected to univariate analyses of variance which also showed no significant treatment effects. However, all three groups made significant gains from their first three FCLP flights to their last three (A Group: $t=2.51$, $df=14$, $p<.01$; B Group: $t=5.08$, $df=17$, $p<.01$; C Group: $t=5.20$, $df=19$, $p<.01$).

The findings of a clear practice effect reflected by LSO grades is consonant with the preceding statement concerning the likelihood of a shifting scale in the LSO-derived error scores.

There were a number of problems other than measurement that further weakened the FCLP/CQ portion of the study. Among these were schedule conflicts that unavoidably disrupted the orderly flow of 2B35 training, missing data, etc. For these reasons, this evaluation of the 2B35 for use in FCLP/CQ must be viewed as inconclusive.

IV. DISCUSSION

The central purpose of this study was to conduct an evaluation of the training transfer effects of Device 2B35 for the visually cued tasks required in the Navy's Advanced Jet Undergraduate Pilot Training. The design chosen for the evaluation was necessarily one feasible of implementation within ongoing operational training that would provide the Navy with operationally meaningful transfer information.

The results of this study are generally supportive of visual simulation, though not strongly so, and they are consistent with the findings from similar efforts in which evidence of positive transfer has been obtained, but in which the effects have been modest.

Two recent Air Force UPT studies are of interest in this regard. In the first (4), transfer of basic contact skills from visual simulator training to the T-37 primary jet training aircraft was examined, while the second (5) examined transfer for aerobatic maneuvers. In both studies, evidence of positive transfer was obtained, but the effects were not dramatic. In addition to these two efforts dealing with UPT training, there have been several other research studies addressing the transfer of training potential of visual simulators for basic fighter maneuvers, aerobatics, transition skills, and weapons delivery. For example, one such study (6) conducted in a well-controlled experimental situation, found a consistent trend toward positive training transfer for Navy F-4 pilots on basic fighter maneuvering tasks--but, except for one maneuver, none of the effects was large enough to be statistically significant. Similar results were obtained by the Air Force during their evaluation (7) of the Simulator for Air-to-Air Combat (SAAC); a small positive trend, but not of statistical significance. In contrast, recent studies of the transfer effects of visual simulator training for weapons delivery appear to have the most impressive results. One in particular (8) found substantial positive transfer for air-to-surface weapons delivery.

As can be seen from the several studies cited, the use of visual simulation has produced moderately positive results for transition/familiarization type skills and somewhat stronger positive results for weapons delivery skills. Thus, the results of this T^2E^2 can be interpreted as providing modest additional support to the use of visual simulation for such military piloting skills.

In contrast, while the A-7E NCLT study previously cited (3) found a significant transfer effect from that trainer to night carrier qualification, no such effect was found here. Whether the divergence between that study's findings and the present results reflects differences in measurement sensitivity, differences in the way the device was used, or differences between day and night CQ cannot be determined here. However, the NCLT results do suggest FCLP/CQ as an area that might be worthy of further investigation in the UPT setting.

Findings

Learning in the 2B35. The results show clearly that students learned in the 2B35. While some of the maneuvers showed greater learning than did others, virtually all reflected improvement in task performance as a function of practice. It is apparent also that some of the device training tasks were relatively easy, while others were relatively more difficult. Since there is only one 2B35 device at each Navy Advanced Jet training site, this finding has implications for device use. Tasks easy to learn in the device should not be given much emphasis, particularly if the device time so consumed prevents time being spent on harder to learn tasks.

The most interesting 2B35 learning result, perhaps, is with reference to WEP training. For both the flight pattern skills and for bombing miss distance, it is clear that asymptote had not been reached at the end of four trainer periods. This, in combination with the positive transfer evidence and the fact that there is still room for considerable improvement in the inflight bombing skills, suggests that additional 2B35 WEP training beyond the four periods might be beneficial.

Transfer to the Aircraft. The transfer results provide general support for the continued use of the 2B35 in Navy Advanced Jet training, but its utility in the different stages of that training varies. Specific recommendations concerning the various stages are as follows.

FAM stage. Overall, the transfer data from the present study neither strongly support nor refute the use of the 2B35 for the various FAM stage maneuvers. The all-aircraft control group did not show any significant flight advantage over the device-trained groups, but neither do the data support the contention that the device produced negative transfer with reference to the critical flight skills of ball control and power control. The finding that device-trained students achieve in five aircraft flights a skill level that appears to be the equivalent of that achieved by control students in seven aircraft flights can be construed as supportive of continued use of the 2B35 for FAM instruction, but no clear cut performance advantage was shown for device-trained students. A saving of two aircraft flights is a saving of some consequence, but it is possible that students who received neither the two extra flights nor the 2B35 training might perform equally well. Such a determination, however, would have required a second control group, a requirement beyond study resources.

Based on all these considerations, it is concluded that continued use of the 2B35 for FAM training is warranted.

Night FAM stage. In view of the apparent lack of difficulty with the NF maneuvers, and since the control group showed some flight advantage over the trainer groups, it is concluded that the 2B35 will not provide significant training benefit in the NF stage of Advanced Jet training.

WEP stage. The results of the T^2E^2 effort with reference to 2B35 in WEP training provide relatively strong support for continued use of the device in this stage. The acquisition of WEP flight pattern skills in the aircraft is clearly enhanced by 2B35 training. It is of some interest to note, though, that the improvement in WEP aircraft flight pattern skills that results from the 2B35 training is not accompanied by statistically significant differences in bomb scores, though the differences did favor the trainer group.

The data concerning vertical and horizontal error components of bomb scores are of interest in terms of instructional emphasis. This finding, in combination with the fact that asymptotic performance level was not reached in either the trainer or the aircraft, suggests that further 2B35 WEP training might be beneficial, particularly to emphasize the obvious relevance of dive angle and release to the vertical error component.

FCLP/CQ stage. In view of problems experienced in the FCLP/CQ stage of the study and the lack of adequate data, no firm conclusion is drawn relative to the use of the 2B35 to support FCLP/CQ stage training. On an analytical basis, the device would seem to have potential for such use, but on the basis of the empirical results of this study, such use can neither be endorsed nor rejected.

Implications for the Future

This effort adds support to the growing body of literature that shows simulators can provide a positive contribution to the meeting of many visual training requirements, in particular in the areas of contact transition or familiarization training and visual weapons delivery. However, it also indicates that use of visual devices should be based on a careful analysis of the task training requirements, the device's capabilities, and the training system in which it will be employed.

The utility of any given visual device must be viewed in this systems context. For example, because of the Navy's use of the "instruments first, contact later" syllabus sequence in Advanced Jet training, the student has considerable skill in flying the TA-4J aircraft on instruments before he is introduced to the 2B35 visual simulator. This instructional sequence probably limits somewhat the 2B35's potential contribution to the acquisition of FAM aircraft skills. An alternative sequencing or use of a visual device in an earlier phase of training (e.g., the T-2 Basic Jet Phase) might result in a different potential. Future programs, such as the VTXTS, must examine such factors carefully if visual simulation and simulation in general are to contribute both effectiveness and efficiency to undergraduate training.

Though not specifically detailed in this paper, three general requirements for effective future simulator training were highlighted in the T^2E^2 effort and are worth noting. These are the requirements for (1) adequate maintenance and personnel support, (2) more objective measures of performance, and (3) instructor training in the instructional use of simulation. With adequate attention to these requirements and to the device and training system factors noted, visual simulation offers significant potential for improving future UPT programs.

REFERENCES

1. General Electric Company. Human factors evaluation 2F90 visual system (Contract N61339-72-C-0192). Daytona Beach, FL: General Electric Company, May 1973.

2. U.S. Navy. Curriculum, Advanced Jet (TA-4J) (CNATRA Instruction 1542.20B). NAS Corpus Christi, TX: Chief of Naval Air Training, September 1976.

3. Brictson, C. A., & Burger, W. J. Transfer of training effectiveness: A-7E night carrier landing trainer (NCLT) Device 2F103 (NAVTRAEQUIPCEN 74-C-0079-1). Orlando, FL: Naval Training Equipment Center, August 1976.

4. Martin, E. L., & Wagg, W. L. Contributions of motion to simulator training effectiveness: Study I - basic contact (AFHRL-TR-78-15). Brooks AFB, TX: Air Force Human Resources Laboratory, June 1978.

5. Martin, E. L., & Waag, W. L. Contributions of motion to simulator training effectiveness: Study II - aerobatics (AFHRL-TR-78-52). Brooks AFB, TX: Air Force Human Resources Laboratory, September 1978.

6. Payne, T. A., Hirsch, D. L., & Semple, C. A. Experiments to evaluate advanced flight simulation in air combat pilot training, Volume I. Transfer of learning experiment. Hawthorne, CA: Northrop Corporation, 1976.

7. U.S. Air Force Tactical Air Command. A continuation training program using the simulator for air-to-air combat (SAAC). Nellis AFB, NV: Tactical Fighter Weapons Center, March 1976.

8. Gray, T. H., & Fuller, R. R. Effects of simulator training and platform motion on air-to-surface weapons delivery training (AFHRL-TR-77-29). Brooks AFB, TX: Air Force Human Resources Laboratory, July 1977.

ABOUT THE AUTHORS

William V. Hagin

Dr. Hagin, who holds the PhD in Experimental Psychology, was Project Director for the research reported here. He is a Senior Staff Scientist with Seville Research Corporation and is Director of Seville's offices in Phoenix, Arizona. Dr. Hagin was formerly Technical Director of the Flying Training Division, Air Force Human Resources Laboratory.

Edwin P. Durall

Dr. Durall, who holds the PhD in Instructional Systems, has been involved in instruction and research in both civilian and military eduactional settings since 1957. He is currently employed by Control Data Corporation in Washington, D.C. At the time of the research reported here, he was employed as a Research Associate in Seville's aircrew training device research program.

Wallace W. Prophet

Dr. Prophet, who holds the PhD in Experimental Psychology, was Program Manager for the 2B35 effort. He is a Senior Staff Scientist and President of Seville Research Corporation. Dr. Prophet was formerly a Vice President of the Human Resources Research Organization (HumRRO) and Director of its Aviation Division.