INCREMENTAL TRANSFER AND COST EFFECTIVENESS OF GROUND-BASED FLIGHT TRAINERS IN UNIVERSITY AVIATION PROGRAMS

Esa M. Rantanen and Donald A. Talleur
Institute of Aviation, University of Illinois at Urbana-Champaign
Savoy, Illinois

Use of ground-based flight training devices in flight training is attractive for several reasons. In addition to undeniable safety aspects and immunity to weather, ground trainers also offer benefits in terms of training effectiveness, typically measured by the time or number of trials saved over training exclusively in an airplane. A review of 19 studies from the past 56 years that have investigated transfer of training effectiveness from ground trainers to airplane revealed, however, that unambiguous conclusions about the best use of these devices are difficult to discern. The reason for the lack of valid data and widely varied results are the large number of intervening variables present in flight training as well as the difficulty of conducting sound research on this topic.

INTRODUCTION

Transfer of training refers to the beneficial effects of prior learning of given skills on learning of subsequent material or skills. In the aviation domain, the transfer of training concept is critical for evaluation of the benefits of pilot training in devices (e.g., simulators) other than an airplane when learning flight skills to a given performance criterion. For example, if a pilot received some amount of training in a simulator and was able to reach criterion performance in an airplane in less time than another pilot who had trained exclusively in the airplane, we would note that positive transfer of training from the simulator to the airplane had occurred. On the other hand, if the pilot had learned ‘bad habits’ in the simulator that had to be unlearned in the airplane, thus extending the time to criterion performance beyond that of the airplane-only pilot, this would be a case of negative transfer.

Transfer of training is quantifiable by several variables. One can consider the number of training sessions or the total time necessary for training to criterion. Counts of errors made while performing a task can also be used to quantify the extent of transfer. Clearly, the cost of using a simulator versus cost of actual flight time is an important factor to be considered, and as this comparison is frequently and unequivocally in favor of the former, transfer of training from ground-based flight training devices has enjoyed a long and sustained research activity (Roscoe, 1971; Flexman, Roscoe, Williams, & Williges, 1972; Povenmire & Roscoe, 1973; Taylor & Stokes, 1986; Taylor et al., 1999, 2002, 2005).

There are also several robust formulas for calculating the amount of transfer (Roscoe, 1971, 1972). The simplest formula is for percent transfer, which is obtained by dividing the difference in the number of trials or time to criterion between a control group and experimental group (that received training in a simulator) by the control group’s trials (or time). However, this formula does not consider the time spent in the simulator. To account for prior simulator training, Roscoe (1971, 1972) proposed cumulative- and incremental transfer effectiveness functions (CTEF and ITEF, respectively). In these functions, the numerator, expressing the difference between control and training groups, is divided by the total training (time or number of trials) received by the training group. The CTEF is given by the equation

$$CTEF = \frac{Y_0 - Y_x}{X}$$

where Y_0 is the time, trials, or errors required by the control group to reach performance criterion and Y_x is the corresponding measure for the experimental group after having received X training units in a simulator, which is also the denominator (Roscoe, 1972). The incremental transfer effectiveness function (ITEF) is given by the equation

$$ITEF = \frac{Y_{x-\Delta x} - Y_x}{\Delta X}$$

where Y_x is the time, trials, or errors of the experimental group to reach a performance criterion, $Y_{x\text{-}\Delta x}$ is the amount of time, trials, or errors required for an experimental group to reach performance criterion after receiving $X–\Delta X$ training units on a prior task, and ΔX is the incremental unit of time, trials, or errors during prior practice on another task (Roscoe, 1972).

Few empirical studies have employed an ITEF design, however. While the reasons for the lack of incremental transfer studies are unclear, time and financial constraints when running these types of studies may top the list. Also, although the ITEF concept proposed by Roscoe (1972) is solid, its implementation is not always clearly understood. Explained simply, the numerator of the ITEF function is the difference in time, trials, or errors of two experimental groups to reach a performance criterion after receiving prior training. The denominator is the difference in amount of prior training (usually referred to as the increment of time) between the two experimental groups being compared. Also, the ITEF formula will yield the same result as the CTEF formula when comparing any one experimental group with the control (no prior training) group.

The ITEF and CTEF functions will yield negatively decelerating curves, or diminishing transfer effectiveness ratios as the number of trials or hours in ground trainer increases. At some point, additional ground-based training will cease to be cost effective. The law of diminishing returns adequately describes this relationship between extra training and resultant benefit. One problem inherent to these functions is that they do not correctly account for the cost of negative transfer. In other words, in the case of negative transfer, a small number of trials or hours in a simulator will yield a very large negative value for TER; this value decreases as clearly detrimental time in the trainer increases. We will return to this issue in results.

In this paper we will review a number of transfer of training studies conducted between 1949 and 2005 and provide a meta-analysis of their respective transfer effectiveness functions. These results will be compared to the theoretical work of Roscoe (1971, 1972). Cost effectiveness of Frasca FTDs and PCATDs in a university flight curriculum is also examined and the many variables involved in the implementation of such trainers in flight curricula are discussed.

METHOD

We performed a literature search and review of published research on transfer of training or training effectiveness that met four specific criteria: (1) the primary transfer target had to be a fixed- or rotary-wing aircraft, (2) a ground-based training device had to be integrated into a full flight training curriculum leading to certification or rating in aircraft, (3) total time to completion had to be used as the primary dependent measure and to calculate percent transfer and/or TER, and (4) the experimental design had to be a multi-group incremental transfer design or a two-group conventional transfer design. A total of 19 studies were reviewed, from 1949 to 2005, from the University of Illinois, the Federal Aviation Administration (FAA), and the U.S. military. The very small number of studies recovered for this review was due to the strict criteria employed; in particular, our desire to study transfer effectiveness necessitated the inclusion of a control group in the experimental design. To the best of our knowledge, only three studies have employed such a complete design: Povenmire and Roscoe (1973) and Taylor et al. (2002, 2005).

RESULTS

The reviewed research can be classified into four groups in terms of the method of simulator integration in the training program: (1) studies that left the integration of prior simulator training up to the instructors providing the training (Povenmire 1970, 1973), (2) studies that used integrated simulation in all training lesson and trained each lesson in the simulator to proficiency (Caro, 1972; Chapman, 1966; Flexman, Matheny, & Brown, 1950; Taylor et al., 1999; Woodruff, 1974), (3) studies that used the simulator only for certain lessons and specific flight maneuvers within those lessons (Caro & Isley, 1966; Kotora & Siebert, 1983; Stewart, Dohne, & Nullmeyer, 2002; Taylor et al., 2002, 2005), and (4) studies where the exact method of simulator integration could no be determined (Crook, 1967).

The percent transfer was calculated from the reported data. Plotting percent transfer against ground training hours shows a clearly positive and fairly linear trend, as depicted in Figure 1. The large variability between studies is noteworthy, however, and bespeaks of the difficulty of making comparisons between different studies with different experimental designs, different ways in which ground trainers were implemented in the curricula, and different fidelity of the trainers. A linear regression of percent transfer against ground training hours was nevertheless statistically significant, F(1, 31) = 23.69, $p < .01$, $R^2 = .433$.

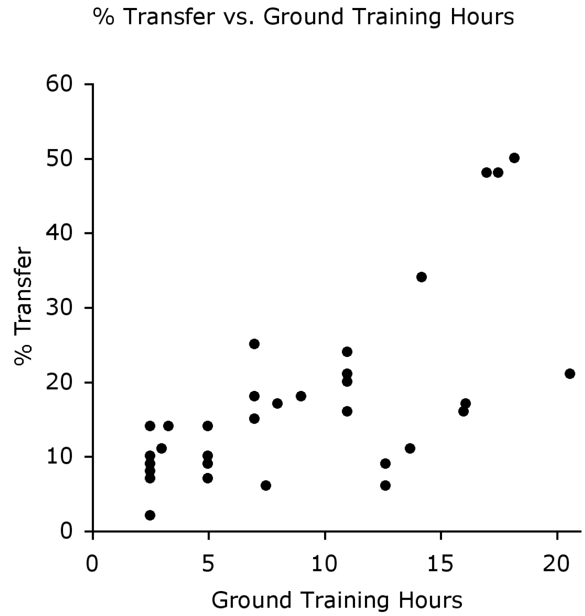

Figure 1. Percent transfer plotted against ground trainer hours. The study by Caro (1972) is excluded from the plot, as with over 40 hours of ground training it represented an outlier among the rest of the studies. Also, one study by Chapman (1966) was not plotted due to non-significant negative transfer results (-3% and –1%). A fairly linear positive trend can be seen in the plot, albeit relatively poorly fitting with an R^2 = .433. A linear trend nevertheless provided the best fit.

While the results concerning percent transfer were quite straightforward and seem to indicate a clear benefit of using ground trainers, the TER provided a more complicated picture, albeit arguably a more realistic one, as this variable also considered the hours spent on ground trainers. In a plot of TER vs. ground trainer hours it is difficult to see any clear trends, although an overall reduction in TER values as hours in ground trainers increased could be discerned (Figure 2). However, a linear regression model proved clearly inadequate to represent the data (R^2 = .02) and more complex models (e.g., third-order polynomial, R^2 = .16), while providing marginally better fits, are simply not interpretable. Hence, our conclusion is that the disparities in the reviewed studies prevent statistical treatment of the meta-data they provide.

We also tested whether type of training had an impact on simulator use effectiveness. In a comparison of 13 contact (visual flight) studies results to 24 instrument flight study results, no difference was found for observed transfer or training effectiveness; $F(1,35) = 0.44$, $p = .51$ and $F(1,35) = 0.11$, $p = .74$ respectively.

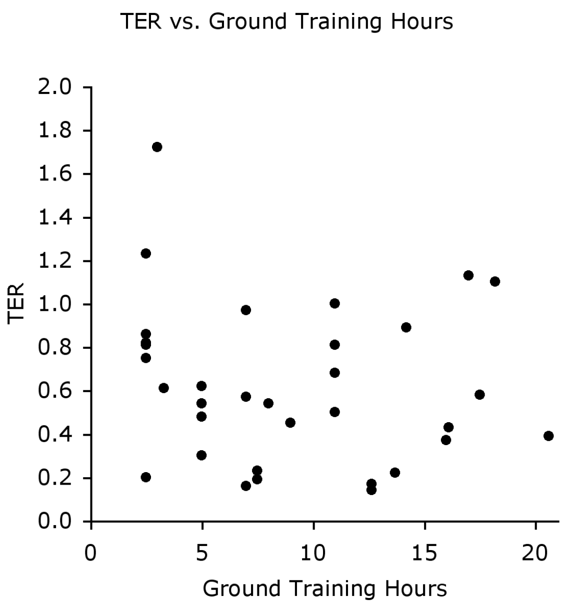

Figure 2. The picture is more complicated when TER is plotted against ground training hours; here, it is difficult to see any clear trends.

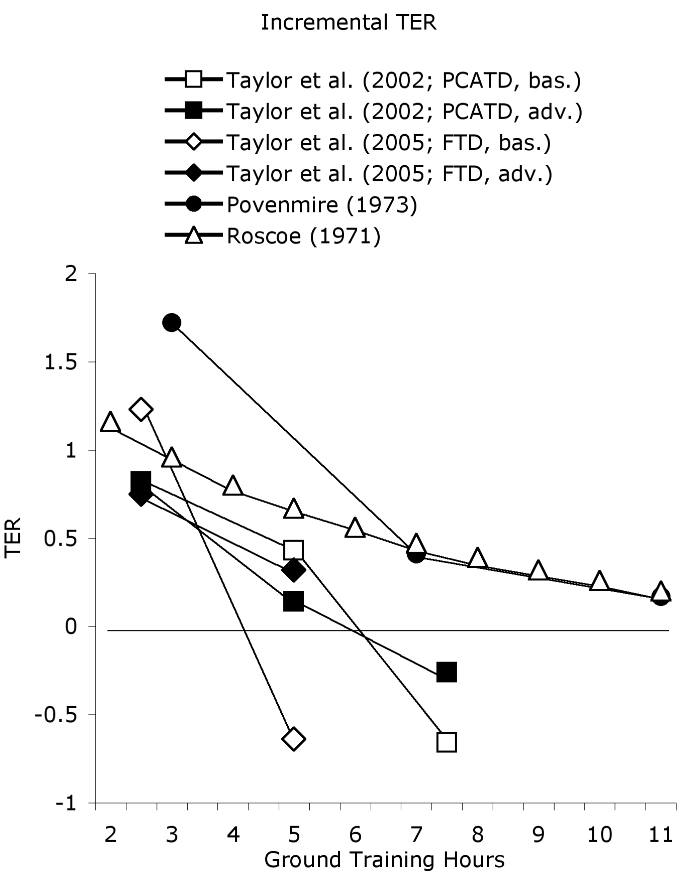

Figure 3. Comparison of incremental transfer effectiveness functions (ITEFs) from the three complete experiments. Roscoe's (1971) data is hypothetical, but has served as a theoretical foundation for all of the reviewed research. Note the much steeper functions from the empirical studies.

Finally, we compared the three incremental transfer effectiveness studies to the theoretical work by Roscoe (1971, 1972). The main finding was that the empirical ITEFs were lower (i.e., yielded smaller TER values) and much steeper than predicted by Roscoe, and in three cases yielded negative TER (Taylor et al., 2002, both basic and advanced instruments, and Taylor et al., 2005, basic instruments; see Figure 3). Together, these results show very rapidly diminishing benefits of using ground trainers beyond only a few hours (e.g., 5 hrs).

The negative TERs found in Taylor et al. (2002, 2005) are difficult to interpret, however, and the linear trends shown in Figure 3 are in fact misleading, as the ITEF is negatively decelerating only for positive values of TER. An ITEF yielding negative TER values is a mirror image of the positive function about the x-axis, meaning that more of clearly detrimental time spent in a ground trainer will yield diminishing negative TER values (see Figure 4). This fact makes interpretation of results such as from Taylor et al. (2002, 2005) difficult.

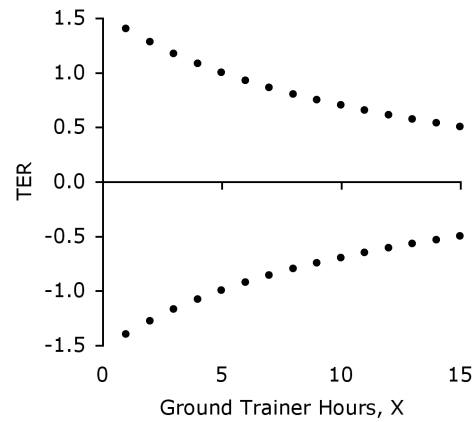

Figure 4. Hypothetical data showing the symmetry of the ITEF function about the x-axis.

DISCUSSION

The most interesting result was the wide range of effectiveness observed by these studies. However, this can be explained in part by the methods used during simulator training and the maneuvers chosen for training. In studies that actually looked at incremental amounts of time used in the simulator (Povenmire & Roscoe, 1973; Taylor et al., 2002, 2005), all produced the expected diminishing returns for additional hours. That is, after about 5 hours of simulator training on average, additional hours produced very little in terms of effectiveness. In many studies, transfer of training effectiveness values were far below values purported by some of the earliest simulator transfer work. However, the measures of transfer frequently differed from study to study and direct comparison with the earlier simulator studies was difficult. Much early work was concerned with error reduction and, as such, did not record total time to reach a desired skill competency or used transfer formulas that differed from those as reported by Roscoe (1971), and now regularly used in flight transfer research studies (e.g., Williams & Flexman, 1949). More recent studies exhibit a shift to establish effectiveness in terms of actual aircraft hours saved as a result of simulator training (e.g., Povenmire & Roscoe, 1973; Kotora & Siebert, 1983; Stewart et al., 2002; Taylor et al., 2002, 2005).

Some of the studies reviewed easily produced a one-for-one savings of aircraft time as a result of simulator time; others produced much less savings. The relationship between the amount of prior simulator training and the resultant savings in aircraft time (43% correlation) was statistically significant, which means that if implemented properly, simulator training leads to time saved in the aircraft. However, this finding is based on only the 19 studies reviewed that, on average, did not use more than 10 hours of prior simulator training. In terms of simulator use, one clear aspect of the studies that showed the most effectiveness was the very stringent instruction on how the simulator was to be used and where (when) it was to be integrated into the training program. On the other hand, some of the best transfer effectiveness occurred in studies where the instructors were required to use the simulator for a specific number of hours, but they were allowed to use it in any way that would most benefit the student (e.g., Povenmire & Roscoe, 1970, 1973). Programs of instruction that required simulator training in all lessons were only successful in terms of time saved when each lesson involved primarily new material, with little review of previous material (e.g., Taylor et al, 1999, 2002, 2005). The U.S. military (e.g., Caro & Isley 1966; Caro, 1972, Stewart et al., 2002) showed some success with this method in several helicopter training programs, and some civilian contact training programs (e.g., Crook, 1967; Chapman, 1966) have also had relative success, but to a lesser extent.

Training Type: Visual and Instrument Training

As noted in the results, there were few differences between visual and instrument simulator training results. The procedural aspects of instrument flight clearly make simulator an attractive training tool. However, the lack of difference between visual and instrument is odd because visual flying is thought to be far less procedural than instrument flying. We may only surmise that the procedural aspects (maneuver setup, etc.) of visual flying can be effectively learned in the simulator, and it is in this area that time is saved during actual flight. It goes without saying that a student who is clear on the procedures for performing a maneuver, and has even an inkling of what to expect, will do better in the aircraft and likely master the maneuver in less time.

Implementation

Our meta-analysis indicates that implementation is an important aspect of an integrated training program. If the simulator time is to be severely limited, it would appear that a thorough training course analysis should be performed to determine which maneuver(s) can be most effectively trained. This may mean that an occasional student will not benefit as much as the next, but a goal of maximizing the average flight time savings across all students in a program can still be reached. If the prior simulator training time allowed in a course of instruction is unlimited, the available data seem to suggest that leaving the content to the discretion of the instructor may be more effective than strict guidance on the simulator's use. One caveat warranted for this method of training is that instructors should be carefully standardized on the use of the simulator and should be simulator proficient on the maneuvers they intend to teach. In this way, the instructor will be aware of the simulator's limitations as well as its strengths, and can use the device for those maneuvers for which it is best suited. Training beyond 10 hours in the conventional simulator (without visual system), while not likely to produce negative results, appears to serve little use in terms of time saved in the aircraft. Newer simulators are likely to have some sort of basic visual system, and some are quite impressive. These devices offer new cues to be learned but the exact nature of their effectiveness is not well documented. However, such simulators may well enhance the utility and potential cost effectiveness beyond the 10-hour limit.

Simulator Fidelity

There are several aspects to be considered when interpreting the results presented in Figure 3 and the departure of the empirical data from Roscoe's (1971, 1972) theoretical work. First, simulation fidelity of the PCATD is substantially lower than that of the Link trainers used by Roscoe. Therefore, it is plausible that the results of Taylor et al. (2002) reflect transfer of merely procedural elements of training, leaving most of the psychomotor flying skills to be learned in the airplane. Second, given the fidelity issues of current ground-based aviation trainers, their use for training of basic flying skills may be called into question. It could be postulated that training focused on procedural aspect of flight and emergency situation management in ground-based trainers would, apart from being inherently safer, also result in higher transfer to the airplane. That the FTD data also fared poorly relative to the Roscoe's theoretical data is not necessarily a reflection on fidelity issues, although there are clear differences between the two devices. In the Taylor et al. (2005) studies, it was noted that the sample sizes were small and the instructors less experienced than in previous studies conducted by Taylor's group.

Other Considerations: Safety and Cost

It has long been the case that ground simulation provides advantages in terms of both safety and cost. Unfortunately, the meta-analysis of studies for this paper revealed few studies that emphasized practice of emergency procedure, tasks which would seem to be ideally suited for ground practice. The exceptions were most of the military studies (e.g., Kotora et al., 1983; Caro & Isley, 1966; Caro, 1972; Stewart et al., 2002) that likely placed a heavy emphasis on mission success, and tended to spend significant amounts of time on emergency practice. Of course, the nature of helicopters and military fixed-wing trainers would seem to require a higher level of emergency procedural skill than that required in most civilian training regimes.

As for the issue of cost, the relative difference in cost of military training aircraft to their civilian counterparts makes a clear case for simulator use prior, as well as during a flight training program. In many cases, even transfer effectiveness as low as .1 may be cost effective when considering the hourly cost of a simulator versus that of the target training aircraft. In the general aviation domain, however, the cost of top-of-the line simulators is often an issue, particularly for small flight schools.

CONCLUSION

Major research questions still remaining are whether maneuvers should be trained piecemeal and then integrated them into a training flight in the aircraft or to 'train like we fly'. The latter idea might require applying the line-oriented flight training (LOFT) concept used by the airlines in the general aviation domain. Another potential research paradigm is approaching simulator training from a whole new perspective. Specifically, part-task training and ways that instructors can teach the fundamentals of aircraft control and procedural application out of context from specific maneuvers are areas ripe for research. From an instructor's point-of-view it seems that most of the difficulty students experience with complex maneuvers is due to a lack of solid aircraft control skills coupled with poor procedural knowledge.

As the research on flight training effectiveness carried out at the University of Illinois during the past several years has made clear, transfer of training studies are complicated, expensive, and time-consuming to conduct. Difficulty of accomplishing adequate experimental validity and results that are significant in both practical (effect size) and statistical (p-value) terms also severely limits the size of body of evidence to draw conclusions about effective and appropriate use of simulators in flight training. On the other hand, over five decades of transfer of training studies in the aviation domain surely has accumulated ample data for meta-analytic treatment. However, only a fraction of those studies have used conventional methods of determining transfer to an operational requirement and many fail to report data or other statistics that would make it possible to reconstruct transfer or TER values (Hays et al., 1992). The limited evidence from the University of Illinois studies reviewed here nevertheless point to a possible conclusion: it may indeed not be the most effective and cost-efficient use of ground-based flight training devices to train basic flight skills using traditional training techniques. Instead, the potential of these devices in practicing emergency procedures or part-task training should be systematically investigated. Scenario based training may offer a boost to transfer effectiveness as well, by enforcing the associations of tasks that would otherwise be practiced individually.

REFERENCES

Caro, P. W. (1972). *Transfer of instrument training and the synthetic flight training system* (TR HumRRO-PP-7-72). Alexandria, VA: Human Resources Research Org.

Caro, P. W., Jr., & Isley, R. N. (1966). Helicopter trainee performance following synthetic flight training. *Journal of the American Helicopter Society, 11*(3), 38-44.

Chapman, G. C. (1966). *An experimental assessment of a ground pilot trainer in general aviation* (Technical Report ADS-63). Columbus, OH: OSU Department of Aviation.

Crook, W. G. (1967). *Experimental assessment of ground trainers in general aviation pilot training* (FAA-ADS-67-5). Washington, DC: FAA Aircraft Development Service.

Flexman, R. E., Matheny, W. G., & Brown, E. L. (1950). Evaluation of the school link and special methods of instruction. *Aeronautics Bulletin* (No. 8). Univ. of Illinois

Flexman, R. E., Roscoe, S. N., Williams, A. C., & Williges, B. H., (Eds.). (1972). *Studies in pilot training: the anatomy of transfer*. (Aviation Research Monographs, Vol. 2, No. 1). Savoy, IL: University of Illinois, Institute of Aviation.

Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Flight simulator training effectiveness: a meta-analysis. *Military Psychology, 4* (2), 63-74.

Kotora, J., & Siebert, W. F. (1983). Training effectiveness evaluation of device A/F37A-T59. *Proceedings of the International Conference on Simulators*, Sep. 26-30, 1983, The University of Sussex.

Povenmire, H. K., & Roscoe, S. N. (1970). *An evaluation of ground based flight trainers in routine primary flight training* (Technical Report ARL-69-1/LF-69-1). Savoy, IL: University of Illinois Aviation Research Laboratory.

Povenmire, H. K., & Roscoe, S. N. (1973). Incremental transfer effectiveness of a ground-based general aviation trainer. *Human Factors, 15*(6), 534-542.

Roscoe, S. N., (1971). Incremental transfer effectiveness. *Human Factors, 13*(6), 561-567.

Roscoe, S. N., (1972). A little more on incremental transfer effectiveness. *Human Factors, 14*(4), 363-364.

Stewart, J. E., II, Dohme, J. A., & Nullmeyer, R. T. (2002). U.S. Army initial entry rotary-wing transfer of training research. *Int'l J. of Aviation Psychology, 12*(4), 359-375.

Taylor, H. L., & Stokes, A. F. (1986). Flight simulators and training devices. In J. Zeidner (Ed.), *Human Productivity Enhancement: training and human factors in systems design, Vol 1.* (pp. 81-129). Praeger Press.

Taylor, H. L., Lintern, G., Hulin, C. L., Talleur, D. A., Emanuel, T. W., Jr., & Phillips, S. I. (1999). Transfer of training effectiveness of a personal computer aviation training device. *Int'l J. of Aviation Psychology, 9*(4), 319-335.

Taylor, H. L., Talleur, D. A., Emanuel, T. W., Jr., & Rantanen, E. M. (2005). Transfer of training effectiveness of a flight training device (FTD). *Proc. 13th International Symposium on Aviation Psychology.* Oklahoma City, OK.

Taylor, H. L., Talleur, D. A., Emanuel, T. W., Jr., Rantanen, E. M., Bradshaw, G. L., & Phillips, S. I. (2002). *Incremental training effectiveness of personal computer aviation training devices (PCATD) used for instrument training.* (ARL-02-5/NASA-02-3). Savoy, IL: University of Illinois, Aviation Research Lab.

Williams, A.C.,& Flexman, R.E. (1949). *Evaluation of the School Link as an aid in primary flight instruction*. Aeronautics Bulletin No. 5, University of Illinois, Institute of Aviation.