



**Aviation Human Factors Division
Institute of Aviation**

**University of Illinois
at Urbana-Champaign
1 Airport Road
Savoy, Illinois 61874**

**The Effectiveness of a
Personal Computer Aviation Training Device,
a Flight Training Device, and an Airplane
in Conducting Instrument Proficiency Checks**

Volume 1: Subjective Pilot Performance Evaluation

**Henry L. Taylor, Donald A. Talleur,
Esa M. Rantanen, and Tom W. Emanuel, Jr.**

Final Technical Report AHFD-04-12/FAA-04-5

November 21, 2004

Prepared for

**Federal Aviation Administration
Civil Aerospace Medical Institute
Oklahoma City, OK**

Cooperative Agreement DTFA 2001-G-037

Foreword

This study was prompted by the FAA Advisory Circular (AC) No. 61-126 (1997), which authorized the use of a Personal Computer Aviation Training Device (PCATD) to be used for 10 of the 15 hours authorized for an approved ground training device. The advisory circular, however, did not authorize the use of PCATDs for Instrument Proficiency Checks (IPCs). The study was supported under Federal Aviation Administration (FAA) cooperative agreement DFTA2001-G-037 with the Institute of Aviation, University of Illinois at Urbana-Champaign, during September 2001-November 2004. The study was sponsored by FAA Headquarters Flight Standards Service, General Aviation and Commercial Division. Dennis B. Beringer, Civil Aerospace Medical Institute (CAMI), served as the contracting officer's technical representative for FAA-CAMI. This report is Volume 1 of a two volume final report and is in the process of review and approval and is not at present an official FAA document. Consequently, the views expressed herein do not necessarily represent official FAA positions. Volume 2 will cover results obtained from objective pilot performance measures employed in the project. Semi-annual, annual, and published reports and presentations of the work including reports of the airborne flight data recorder (FDR) and development of objective pilot performance measures are listed in Appendix A.

We express our appreciation to Mary Wilson, who scheduled participants, and to Karen Ayers, who assisted with report formatting. We also thank Bill Jones, David Boyd, Sybil Phillips and Donald Talleur, who served as the check pilots. We also thank the Institute of Aviation flight instructors who provided familiarization training in the Flight Training Device, the PCATD, and the airplane, as well as the instrument pilots for their participation in the study.

Executive Summary

To maintain instrument currency, instrument pilots must meet the recent instrument experience requirements of FAR 61.57(c) or (d) every six months. The requirements may be conducted in an airplane or an approved flight training device (FTD). If an instrument pilot fails to meet recent experience requirements within the previous 12-month period, an instrument proficiency check (IPC) must be successfully accomplished to regain instrument currency. The present study compared the performance of pilots receiving an IPC in a (PCATD), a FTD or an airplane (IPC #1) with their performance in an airplane (IPC #2). The comparison of performance in a PCATD to that in an airplane investigates the effectiveness of the PCATD as a device in which to administer an IPC. Currently, the PCATD is not approved to administer IPCs. The comparison of performance in a FTD with performance in an airplane will help determine whether the current rule to permit IPCs in a FTD is warranted. The comparison of the performance in a PCATD and a FTD permits a comparison of the relative effectiveness of the respective devices. Finally, the comparison of performance of pilots receiving IPC #1 in an airplane and IPC #2 in an airplane with a second CFII will permit the determination of the reliability of IPCs conducted in an airplane.

This study involved 75 participants (25 participants in each group: FTD, PCATD and airplane). Each participant agreed to refrain from instrument flight (either in flight or in a ground-based device) between IPCs #1 and #2. They also flew a familiarization flight in the FTD, the PCATD and the airplane prior to being randomly assigned to one of the three groups (FTD, PCATD and airplane). The participating instrument pilots in the study were in one of four categories of instrument currency: (1) instrument current, (2) within one year of currency, (3) between one and two years of currency, and (4) between 2 and 5 years of currency and they were balanced among the three groups. Pilots in the 2 to 5 year category received up to five hours of instrument proficiency training in either a FTD or a PCATD prior to the experiment

The results indicated no significant differences in performance by instrument pilots on an IPC given in either a PCATD, and FTD or an airplane. Performance on the IPC of the PCATD group was statistically indistinguishable from both the airplane and the FTD groups. In addition, there was no difference in performance between the aircraft and the FTD groups. These findings present compelling evidence that the FAA should permit the use of PCATDs to give IPCs.

It was expected that performance on IPC #1 would be a good predictor of performance on IPC#2. The results indicated that the prediction was no better than chance. The change in performance between IPC #1 and IPC #2 for all participants was statistically significant, but none of the comparisons experimental groups were significant. Analyses to determine the performance changes between IPC #1 and the IPC #2 for each experimental group were conducted and improvement and deterioration ratios were calculated. The improvement and deterioration ratios for the three groups were very similar.

Only 24 of the 75 participants passed IPC #1 (32%), and 51 failed (68%). Of the 53 instrument current pilots, only 19 (36%) passed IPC #1 and 34 failed (64%), but 30 passed IPC # 2 (57%) in the aircraft and 23 failed (43%). These results are comparable to those of an earlier study by Taylor, Talleur, Bradshaw, Emanuel, Rantanen, Hulin and Lendrum (2001).

Table of Contents

Foreword.....	i
Executive Summary.....	ii
Table of Contents.....	iii
1. Introduction.....	1
1.1. Background.....	1
1.2. Transfer of Training Studies.....	1
1.3. Recency of Experience Studies.....	2
1.4. Purpose of the Study.....	3
2. Method.....	3
2.1. Participants.....	3
2.2. Apparatus.....	4
2.3. Procedure.....	4
3. Results.....	5
3.1. IPC #1 and IPC# 2 Pass/Fail Rate by Group.....	5
3.2. Performance of Changes from IPC #1 to IPC #2.....	7
3.3. IPC #1 and IPC #2 Pass/Fail Rates by Currency Status.....	9
3.4. Change in Maneuver Performance between IPC #1 and IPC#2.....	10
3.5. Demographic Factors.....	13
3.6. Inter-Rater Reliability.....	13
4. Discussion.....	15
5. Recommendations for Rulemaking.....	17
6. Summary and Conclusions.....	17
7. References.....	18
Appendix A. Reports and Presentations.....	21
Appendix B. Flight Experience and Biographical Data Questionnaire.....	23
Appendix C. Training Sessions.....	27
Appendix D. VFR Flight Familiarization Flight.....	28
Appendix E. IPC 1-PCATD Form.....	29
Appendix E. IPC 1-Frasca Form.....	34
Appendix E. IPC 1-Aircraft Form.....	39
Appendix E. IPC 2-Aircraft Form.....	44
Appendix F. FAA IPC Checkpilot’s Post-Project Questionnaire.....	49
Appendix F. IPC Checkpilot Post-Project Questionnaire Results.....	52

1. Introduction

1.1. Background

During the past ten years there has been an increased emphasis on the use of personal computers in flight training (Campbell, 1998; Miller, 1996; Kolano, 1997). Williams (1994) summarized the findings of a joint industry-FAA conference concerned with the development and use of personal computers, which documented this emphasis. Koonce and Bramble (1998) provided an overview of the use of personal computer-based flight training devices (PCATDs). Today the personal computer is a viable tool for presenting realistic, high-quality, full-size graphic representations of aircraft instrument displays. Current technology can also provide aerodynamic models and realistic flight controls that closely mimic those experienced in flight and are as accurate as FAA approved flight training devices (FTDs). It should be noted that PCATDs also accept control inputs from low-fidelity devices that range from computer keyboards, single joysticks, and yoke/pedal combinations of varying quality (Peterson, 1993). PCATDs also have navigation databases, which provide unlimited geographic coverage. However, while PCATDs provide many features required to practice instrument tasks, their fidelity may be low in areas such as displays, switches, out-of-cockpit scenes, control loading and flight dynamics, which are usually thought to be important in instrument training. Nevertheless, and most importantly, PCATDs offer a low-cost alternative for instruction of instrument tasks at a cost of less \$6,000.

Empirical evaluations by Phillips, Hulin and Lamermayer (1993), Ortiz (1994) and Dennis (1994) have found positive transfer of training from personal computers to the airplane, but these studies have been limited in dealing primarily with individual instrument tasks. A report by Hampton, Moroney, Kirton, and Biers (1994) reported that students trained in a PCATD performed as well on instrument procedures in the airplane as students trained in a Frasca 141 FTD. No airplane control group was used in this study, and so it was not possible to determine the transfer effectiveness of the PCATD or the Frasca 141. Karp (2001) recently described the use of PCATDs in the classroom. He found that the use of PCATDs in the classroom was effective but the PCATD s was not used to substitute for flight time in the airplane.

1.2. Transfer of Training Studies

Taylor, Lintern, Hulin, Talleur, Emanuel, and Phillips (1996, 1999) and Taylor, Talleur, Phillips, Emanuel, and Hulin (1998) investigated the effectiveness of using a PCATD to develop instrument skills taught in instrument flight training and to determine the extent that these skills transfer to the aircraft. A commercially available PCATD was used to teach instrument tasks to students at the Institute of Aviation of the University of Illinois. To evaluate transfer of training, the performance of a group of participants trained in a PCATD and later trained to criterion in an airplane (PCATD group) was compared to the performance of a control group of participants trained only in the airplane (Airplane group). All new maneuvers and procedures for the PCATD group were introduced and trained to proficiency in a PCATD prior to training and skill validation in the airplane. All new maneuvers for the Airplane group were introduced and trained to proficiency in the airplane. The dependent variables were the number of trials to criterion in the airplane for the two groups, time to complete each flight lesson in the airplane, and course

completion times. The results indicated that the PCATD was an effective training device for teaching instrument tasks. Transfer savings were generally positive and statistically significant when new tasks were introduced, but lower transfer was found when reviewing previously learned instrument tasks. An evaluation of course completion times showed a saving, which was statistically significant, of about four hours in the airplane for the PCATD group compared to the Airplane group. The transfer effectiveness ratio (TER) was 0.15 or a savings of 1.5 flight hours for each ten hours of PCATD time. The low overall TER was due to the methodology of including the review of instrument tasks previously trained to proficiency in the overall TER.

Williams and Blanchard (1995) discussed the development of qualification guidelines for PCATDs. As result of the Taylor et al. (1996) study, the Federal Aviation Administration (FAA) published an advisory circular concerned with the qualification and approval of PCATDs (U.S. Department of Transportation, 1997). The advisory circular permitted the use of PCATDs in instrument training programs conducted under FAR Part 61 and FAR Part 141 and authorized the substitution of a PCATD for 10 of the 15 hours authorized for an approved ground training device. The advisory circular did not authorize the use of PCATDs for instrument proficiency checks (practical tests) nor for instrument recent flight experience requirements.

In a follow-on study concerning incremental transfer of training effectiveness Taylor, Talleur, Emanuel, Rantanen, Bradshaw, and Phillips (2001) found that the PCATD was effective in teaching basic instrument tasks to private pilots. Prior training in the PCATD for 5, 10, or 15 hours resulted in a smaller number of trials in the airplane for each of the three PCATD groups when compared to the Airplane group which was trained only in the airplane. However, the transfer effectiveness ratio was not a simple function of the amount of practice in the PCATD. Although it seems reasonable to believe that greater training in the PCATD would reduce the amount of training needed in the aircraft this prediction was not borne out. For five of the eight instrument tasks, the PCATD 10-hour group needed the fewest number of trials in the airplane, for two tasks the PCATD 5-hour group had the fewest number of trials in the airplane and the PCATD 15-hour group had one task with the fewest number of trials in the airplane. Of course, all groups benefited to some extent from their practice. The mean times to complete the flight lesson in the airplane for the four flight lessons in which there was prior training in the PCATD were less for all three PCATD groups than for the Airplane group.

1.3. Recency of Experience Studies

A study by Taylor, Talleur, Bradshaw, Emanuel, Rantanen, Hulin and Lendrum (2001) and Talleur, Taylor, Emanuel, Rantanen, & Bradshaw (2003) determined the effectiveness of PCATDs for maintaining instrument currency. One hundred six instrument current pilots were divided in four groups: a PCATD group, a FTD group, and airplane group, and a control group. The pilots in each group received an initial instrument proficiency check (IPC #1). During a six-month period following IPC #1, the pilots in three groups received recurrent training in a PCATD, a Frasca FTD, or an airplane, respectively. The fourth (control) group received no training during the six-month period. After this time, the pilots in each group flew an instrument proficiency check (IPC #2). The comparison of IPC #1 and IPC #2 indicated that both the PCATD and the Frasca FTD were more effective in maintaining instrument proficiency when compared to the control group and at least as effective as the airplane. The study also found that of 106 instrument current pilots, only 45 (42.5%) were able to pass IPC #1. Of the group who received an IPC in a Frasca FTD to regain currency, only 22 of 59 (37.3%) were able to

subsequently able to pass IPC #1 in an airplane. This study established the effectiveness of PCATDs for use in instrument currency training. However, the question of whether PCATDs are effective for administering the IPC was not demonstrated in this project. Based on the data above, a question concerning the effectiveness of the Frasca FTD in administering an IPC also arose.

1.4. Purpose of the Study

Although the PCATDs have been shown to be effective in training basic instrument flight skills as well as in maintenance of these skills to meet recency of experience requirements, a critical element, using PCATDs to administer and IPC flight has been conspicuously missing from the research literature. The purpose of the present study was to compare the performance of pilots receiving an IPC (IPC #1) in a PCATD, a FTD or an airplane with their performance in an airplane (IPC #2). Currently, the PCATD is not approved to administer IPCs. The comparison of performance in a FTD with performance in an airplane was to determine whether the current rule to permit IPCs in a FTD is warranted. Finally, the comparison of performance of pilots receiving IPC #1 in an airplane and IPC #2 in an airplane with a second CFII will permit the determination of the reliability of IPCs conducted in an airplane and in accordance with current IPC requirements.

2. Method

2.1. Participants

One hundred and five participants were originally approved for the project. Due to funding constraints the total number of pilots participating in the experiment was reduced to 75 (25 per group). All participants were instrument pilots, but not all were instrument current when they began the experiment. The participants agreed to refrain from instrument flight for the duration of the study (about 2 weeks). They also agreed not to use a PCATD for instrument training during this period. Volunteer participants were recruited within a 75-mile radius of Champaign, IL. Their participation was solicited using a mail survey, which was sent to all instrument-rated pilots in the area. A total of 267 pilots responded with a statement of interest. Pilot Experience and Biographical Data Questionnaires (Appendix B) were mailed to the 267 instrument pilots who expressed interest. A total of 179 pilots returned the questionnaire, which collected information about the pilot's experience and instrument currency status. Of these 179 pilots, 146 were placed in the potential participant pool. Additional participants were later added to the pool. The average age of the participants was 49 years with a range of 20 to 80 years. Average total flight experience was 1909 hours with a range of 142 to 13,000 hours. Average experience in aircraft similar to the type used in the experiment was 1572 hours with a range from 140 to 11,438 hours.

The instrument pilots considered potential participants for the study were in one of four categories of instrument currency: (1) instrument current, (2) within one year of currency, (3) between one and two years of currency, and (4) between 2 and 5 years of currency. Pilots in the 2 to 5 year category received up to five hours of proficiency training in either a FTD or a PCATD prior to the experiment (see Appendix C). The device used for this proficiency training was balanced across participants. Several potential participants failed to reach proficiency (a level of performance indicating they should be able to pass an IPC in the aircraft) and were

subsequently released from the project prior to their involvement in the experiment. No additional participants in the 2-5 year category were included in the study after the number of participants in the study was reduced to 75. All participants had the option of receiving payment for flight time flown during the experiment, as well as mileage costs to and from Willard Airport in Savoy, IL, where all sessions took place.

2.2. Apparatus

Two FAA approved Elite PCATDs with a Piper Archer performance model and two FAA approved Frasca 141 FTDs with a generic single-engine, fixed-gear, fixed-pitch propeller performance model were used. The FTDs were approved for instrument training towards the instrument rating, instrument recency of experience training, and instrument proficiency checks (IPCs), as well as for administering part of the instrument rating flight test. Two single engine 180 hp Beechcraft Sundowner aircraft (BE-C23) with a fixed-pitch propeller and fixed undercarriage were used as the testing aircraft for IPC #1 and IPC #2. An airborne flight data recorder (FDR) was installed in each aircraft to record flight data during the IPC flights (Lendrum, Taylor, Talleur, Hulin, Bradshaw, & Emanuel, 1999, 2000).

2.3. Procedure

The study directly compared the performance of pilots receiving an IPC in a PCATD, a FTD or an Airplane (IPC #1) with performance in an airplane (IPC #2). The comparison of performance in a PCATD to that in an airplane investigates the effectiveness of the PCATD as a device in which to administer an IPC. Currently, the PCATD is not approved to administer IPCs. The comparison of performance in a Frasca and the airplane will determine whether the current rule to permit IPCs in a FTD is warranted. Finally, we compared the performance of pilots receiving IPC #1 in an airplane and IPC #2 in an airplane with a second certified flight instructor, instruments (CFII). This comparison will permit the determination of the reliability of IPCs conducted in an airplane.

Each participant received a Visual Flight Rules (VFR) familiarization flight in the FTD, the PCATD and the airplane prior to being assigned to an experimental group (see Appendix D). The participants also received a review of the aircraft systems and instrumentation in each device. The order in which each participant received their familiarization flights was randomized with a constraint that each experimental group had a balance of the participants receiving each of the six possible familiarization orders. Following the familiarization flights, participants were randomly assigned to one of the three groups (FTD, PCATD and airplane) with a constraint that the four currency categories and familiarization orders were balanced among the groups.

All pilots began the experiment with a baseline instrument proficiency check (IPC #1) in either an Elite PCATD, a Frasca FTD or a Beechcraft Sundowner airplane, according to which group they were assigned. IPC #1 was flown with a certified flight instructor, instruments (CFII), who acted both as a flight instructor and as an experimental observer. The IPC is a standardized test of the instrument pilot's instrument skills. The types of maneuvers, as well as completion standards for an IPC, are listed in the instrument rating practical test standards (PTS) (U.S. Department of Transportation, 1999). A flight scenario, that follows the current guidelines for the flight maneuvers required by the PTS, was used for the IPC. This scenario was used to collect baseline data and to establish the initial level of proficiency for each participant who participated in the project. Table 1 shows the experimental design.

Table 1. *Experimental Design*

GROUP	Fam. Flight	Initial IPC flight (IPC#1)	Final IPC flight (IPC#2)
Airplane	In Sundowner, Frasca and Elite	In Sundowner	In Sundowner
FTD	In Sundowner, Frasca and Elite	In Frasca	In Sundowner
PCATD	In Sundowner, Frasca and Elite	In Elite	In Sundowner

The IPC #1 flight contained seven maneuvers (VOR approach, holding pattern, steep turns, unusual altitude recovery, ILS approach, VOR partial panel approach and ATC procedures and communication). The CFII's for the IPC #1 flight used a form that was designed to facilitate the collection of three types of data (Phillips, Taylor, Lintern, Hulin, Emanuel, & Talleur, 1995) (see Appendix E). First, within each maneuver there were up to 24 variables (e.g., altitude, airspeed) which were scored as pass/fail indicating whether performance on those variables met PTS requirements. Second, the flight instructor judged whether the overall performance of the each maneuver was pass/fail. Third, the CFII recorded if the overall performance of the participant met the PTS for the IPC.

3. Results

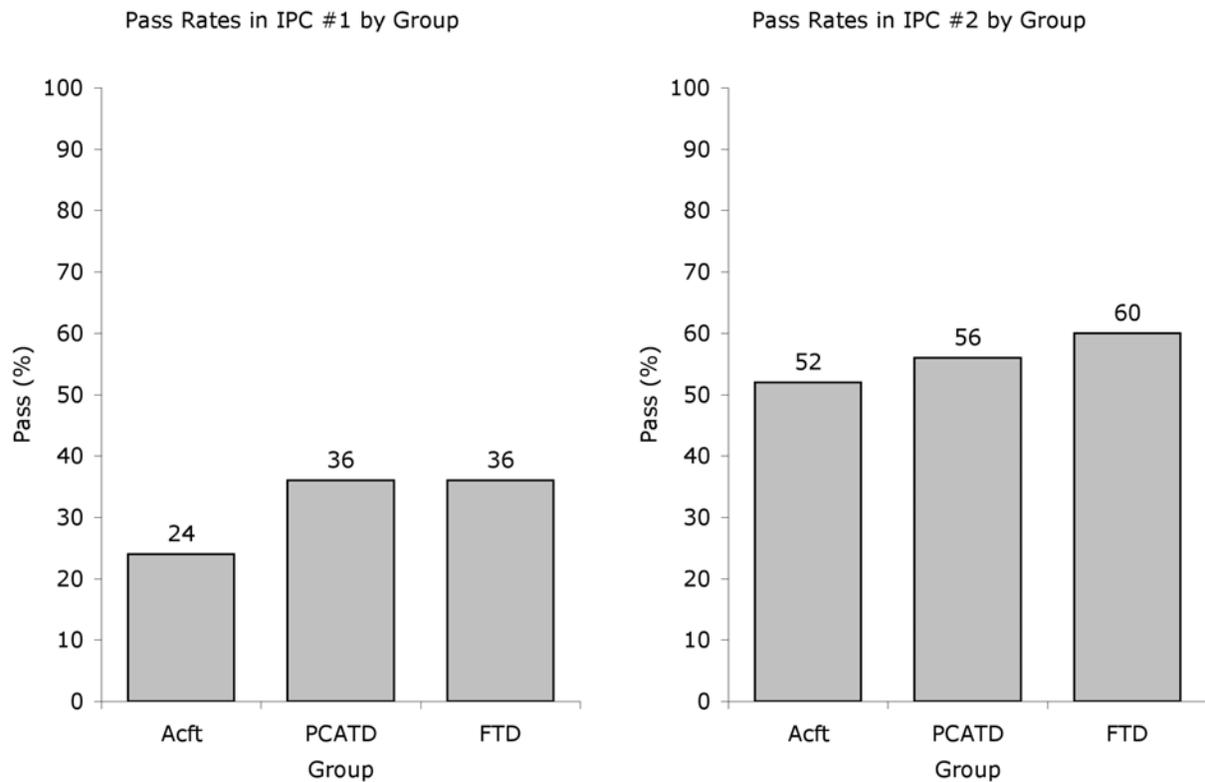
As indicated above, three types of data were collected. The primary data presented first concerns the pass/fail rate on IPC #1 and 2 by the experimental group, which permitted a comparison of the effectiveness of the PCATD and FTD with the airplane. The next level of analysis evaluated the performance of the three groups at the maneuver level for IPC #1 and 2. The third level of analysis was concerned with the elements in each maneuver. Demographic variables were analyzed to determine if they had any effect on the experimental outcome. Finally the inter-rater reliability of the check pilots was analyzed to rule out any contamination of the data by CFII performance.

3.1. *IPC #1 and IPC# 2 Pass/Fail Rate by Group*

Table 2 presents the number and percentage of pilots that passed/failed IPC #1 and IPC #2 for each of the three experimental groups and for the total participants. Figures 1 and 2 shows the differences between pass rates for the three groups for IPC #1 and IPC #2, respectively. Inspection of Figures 1 and 2 indicate few differences between groups for the number of participants who passed IPC #1 and IPC #2. A total of 24 of 75 participants (32%) passed the IPC #1 and a total of 42 of 75 participants (56%) passed the IPC #2 flight across all three experimental groups. Chi-square tests were used to analyze the IPC #1 and IPC #2 data to determine whether the treatment (assignment to group) had an effect on the pass/fail ratio for the IPC #1 and IPC #2 flights respectively. The treatment effect on the IPC #1 pass/fail ratios was not statistically significant, $\chi^2(2, N=75) = 0.32, p = 0.85$. Neither was the treatment effect statistically significant for the IPC #2 pass/fail ratio, $\chi^2(2, N=75) = 1.1, p = 0.58$.

Table 2. Pass/Fail rates for each experimental group

Group	N	IPC#1				IPC#2			
		Pass	(%)	Fail	(%)	Pass	(%)	Fail	(%)
Aircraft	25	6	(24)	19	(76)	13	(52)	12	(48)
FTD	25	9	(36)	16	(64)	14	(56)	11	(44)
PCATD	25	9	(36)	16	(64)	15	(60)	10	(40)
Total	75	24	(32)	51	(68)	42	(56)	33	(44)



Figures 1 and 2. Pass rates in IPC #1 and IPC #2 by experimental group.

A series of planned pairwise comparison tests were performed between and among the experimental groups. The first comparison evaluated the performance of the PCATD group on IPC #2 with the Aircraft group; the difference between these groups was not significant, $\chi^2(2, N = 50) = .32, p > .10$. The next comparison evaluated the performance of the PCATD group on IPC #2 with the FTD Group but yielded no significant difference, either, $\chi^2(2, N = 50) = 0.08, p > .10$. The final comparison evaluated the performance of the Aircraft group on IPC #2 with the Frasca group and showed a non-significant difference between the groups, $\chi^2(2, N = 50) = 0.08, p > .10$.

3.2. Performance of Changes from IPC #1 to IPC #2

Analysis of the change of performance that took place between the IPC #1 and IPC #2 flights was made in order to understand the effectiveness of the three devices in conducting IPCs. It was expected that performance on IPC #1 would be a good predictor of performance on IPC#2. Table 3 shows a comparison of the pass/fail rates for IPC #1 and IPC #2. Of the 24 participants who passed IPC #1 only 14 also passed IPC #2 (58%), and of the 51 participants who failed IPC #1 only 23 (45%) subsequently failed IPC #2 (a total of 37). Twenty-eight participants, who failed IPC #1 subsequently passed IPC #2 and 10 of the participants who passed IPC #1 subsequently, failed IPC #2 (a total of 38). Therefore, performance on IPC #1 predicted the performance on IPC# 2 only at the chance level. Indeed, the McNemar change in performance analysis between IPC #1 and IPC #2 for all participants was significant; $\chi^2 (1, N = 75) = 8.53, p < .005$.

Table 3. Comparison of IPC #1 vs. IPC #2 Pass/Fail

		IPC#2		
		Pass	Fail	Total
IPC#1	Pass	14	10	24(32%)
	Fail	28	23	51(68%)
	Total	42 (56%)	33 (44%)	75 (100%)

Table 4 shows the IPC #1 and IPC # 2 pass/fail for the aircraft group. The McNemar test for change in performance between IPC # 1 and IPC #2 was significant; $\chi^2 (1, N = 25) = 4.45, p < .05$. A continuity adjustment was applied since 2 cells had frequencies less than 5. With the continuity adjustment the test for a change in performance between IPC #1 and IPC #2 was not significant, however.

Table 4. Aircraft Group. IPC #1 vs. IPC #2 Pass/Fail

		IPC#2		
		Pass	Fail	Total
IPC#1	Pass	4	2	6 (24%)
	Fail	9	10	19 (76%)
	Total	13 (52%)	12 (48%)	25 (100%)

Table 5 shows the IPC #1/IPC # 2 pass/fail for the FTD group. The McNemar test for change in performance between IPC # 1 and IPC #2 was not significant; $\chi^2 (1, N = 25) = 1.92, p > .10$.

Table 5. *FTD Group. IPC #1 vs. IPC #2 Pass/Fail*

		IPC#2		Total
		Pass	Fail	
IPC#1	Pass	5	4	9 (36%)
	Fail	9	7	16 (64%)
Total		14 (56%)	11 (44%)	25 (100%)

Table 6 shows the IPC #1/IPC # 2 pass/fail for the PCATD group. The McNemar test for change in performance between IPC # 1 and IPC #2 was not significant; $\chi^2 (1, N = 25) = 2.57, p > .10$.

Table 6. *PCATD Group. IPC #1 vs. IPC #2 Pass/Fail*

		IPC#2		Total
		Pass	Fail	
IPC#1	Pass	5	4	9 (36%)
	Fail	10	6	16 (64%)
Total		15 (60%)	10 (40%)	25 (100%)

The IPC #1 and IPC #2 pass-fail rates by group are presented in Table 7. Analyses to determine the performance changes between IPC #1 and the IPC #2 for each experimental group were conducted and improvement and deterioration ratios are presented (Figure 3). Participants who failed IPC #1 in the aircraft, FTD, or PCATD and passed IPC #2 are included in the improvement ratio and participants who passed the IPC #1 and failed IPC #2 are included in the deterioration ratio. The improvement ratios for the three groups are very similar. The airplane and FTD groups had 36% of their participants who failed IPC #1 pass IPC#2, while the PCATD group had 40%. The airplane group had only 8% who failed IPC # 2 after passing IPC # 1 (the deterioration ratio) while both the FTD and the PCATD groups both had deteriorations ratios of 16%.

Table 7. IPC #1 and IPC #2 pass/fail by group.

Group	Pass IPC #1		Fail IPC #1		Total
	Pass IPC #2	Fail IPC #2	Pass IPC #2	Fail IPC #2	
Aircraft	4	2	9	10	25
FTD	5	4	10	6	25
PCATD	5	4	9	7	25
Total	14	10	28	23	75

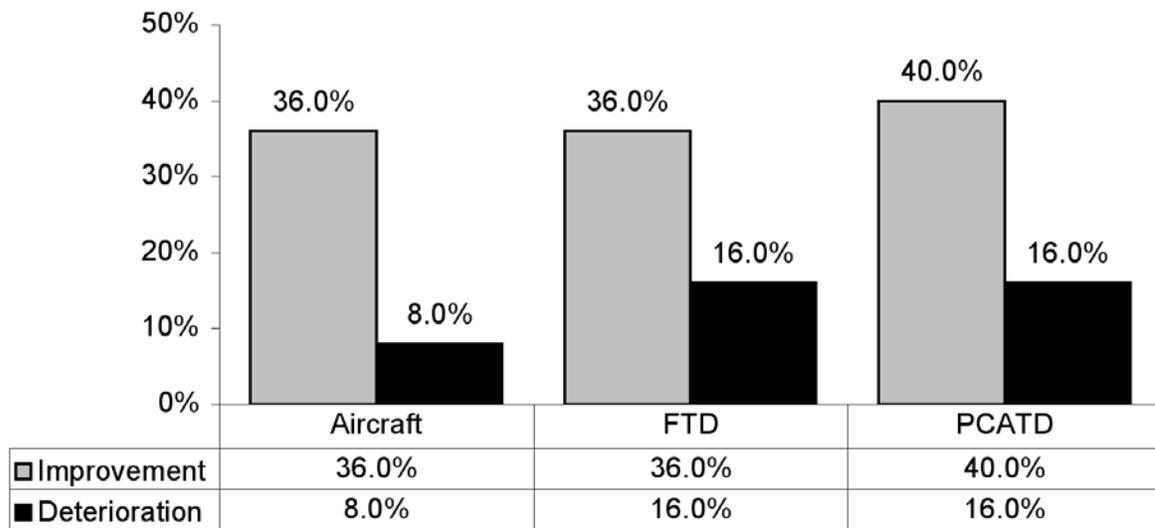


Figure 3. Skill improvement/deterioration by group

3.3. IPC #1 and IPC #2 Pass/Fail Rates by Currency Status

As indicated earlier, only a third of the instrument pilots passed IPC# 1 but over half passed IPC# 2. Table 8 illustrates the number of participants that passed/failed IPC #1 flight by prior currency status. Of the 75 participants who completed IPC #1 only 24 passed (32%) and 51 failed (68%). Of the 53 instrument current pilots, only 19 (36%) passed IPC #1 and 34 failed (64%), but 30 passed IPC # 2 (57%) in the aircraft and 23 failed (43%). The number of participants in the remaining currency groups is too small to permit meaningful analysis.

Table 8. *IPC #1 and IPC #2 pass/fail rates by currency status*

Status	IPC#1				IPC#2			
	Pass		Fail		Pass		Fail	
	N	%	N	%	N	%	N	%
Current	19	36	34	64	30	57%	23	43%
1 year	2	29	5	71	6	86%	1	14%
1-2 year	1	100	0	0	1	100%	0	0%
2-5	2	14	12	82	5	13%	9	87%
Total	24			51	42		33	

The next level of analyses, more detailed than the previous analyses, was by the maneuver, control and procedural elements, and for the individual maneuvers in the IPC flight. Each of these analyses is presented below.

3.4. Change in Maneuver Performance between IPC #1 and IPC#2

An analysis of the changes in maneuver performance that occurred between IPC #1 and IPC #2 was performed to determine if there were systematic changes in performance when considering the overall change in number of maneuvers passed. There were seven maneuvers to be scored in both IPC #1 and IPC#2. An overall “maneuver change score” (+1, 0, -1) for each maneuver was determined for each participant; a positive score represents an improvement from IPC #1 to IPC #2, while a negative score represents a loss of skill from IPC #1 to IPC #2. The maneuver change scores for the seven maneuvers were then summed for each participant. The participant’s overall performance change from IPC #1 to IPC #2 could range from -7 to 7. These scores were then standardized and analyzed using a single factor ANOVA to determine if there was a difference between experimental groups. The change in maneuver performance between IPC #1 and IPC #2 was not significant, $F(2,74) = .78, p > 0.10$.

3.4.1. Control and procedural element change in performance analysis

An analysis of the change in performance at the maneuver element level was performed to determine if any particular control or procedural elements contributed significantly to the overall maneuver pass/fail judgment. This analysis goes beyond the maneuver pass/fail change score analysis at the maneuver level in that it examines whether performance on certain component skills of the maneuvers in Table 9 (i.e., procedural and psychomotor skills) were influenced by experimental assignment. There were several different elements for each maneuver that received a pass/fail score for both IPC #1 and IPC#2, so it was possible to assign an overall ‘maneuver element change score’ to each participant in the same manner as described earlier. The participant received a ‘1’ or ‘0’ assignment depending on whether they passed or failed a given element within a given maneuver. Maneuver elements fell into two distinct categories: procedural and psychomotor (referred to as ‘control’). Therefore, two separate maneuver element change scores were computed. The procedural element change score consisted of those maneuver elements that were not directly related to aircraft control, but rather the execution of

instrument procedures. The control element change score consisted of those maneuvers that specifically indicated how well the participant controlled the aircraft flight path and performance during flight in the process of executing the aforementioned procedures. Only four of the six flight maneuvers scored had both procedural and control maneuver elements. The remaining three maneuvers consisted of either pure procedural or pure control elements (ATC communication was not analyzed here since it has only one procedural element); the unusual attitude has only one control element; the steep turn consists of only control elements).

Four separate ANOVA were used to compare all three experimental groups for a change in performance on the VOR approaches, hold, and ILS approach on either procedural or control elements between IPC #1 and IPC #2 (each ANOVA was a 2 x 3 mixed design). None of the ANOVAs showed a main effect for group assignment

Table 9. *Number of procedural/control elements on each maneuver*

Maneuver	Procedural	Control	ANOVAs	McNemar, χ^2
VOR	11	13	p > .10	
HOLD	9	11	p > .10	
Steep Turn	0	6	p > .10	
Unusual Attitude	0	1		p > .10
ILS	4	8	p > .10	
VOR, partial panel	5	8	p > .10	
ATC	1	0	NA	NA

The Unusual attitude and ATC communication contained only one control and procedural element respectively, and as such, a pass or failure on those single elements determined whether the maneuver was passed or failed. A McNemar χ^2 analysis was used to compare the performance on the unusual attitude maneuver since, due to only one element, a pass or failure could be tallied by a simple count. The analysis showed that this maneuvers' performance did not differ significantly between IPC #1 and IPC #2 for the three experimental groups. The ATC procedural maneuver did not have a sufficient number of failures in any of the three groups to warrant analysis. From these analyses, we infer that the performance changes observed between IPC #1 and IPC #2 were not related to the group assignment.

3.4.2. *Change in performance between the IPC flights for individual maneuvers*

Since the previous analysis considered the performance on all six maneuvers simultaneously, the pattern of maneuvers passed in IPC #1 and IPC #2 could not be determined. It was not necessary to "pass" all maneuvers in order pass IPC #1 and IPC #2. Therefore a change in the individual maneuvers passed, while leaving the total number of maneuvers passed unchanged, could easily result in a participant who had not previously passed IPC #1, passing IPC #2. Performance changes between IPC #1 and IPC #2 for each group, considering each maneuver individually, were compared. Since this analysis compares pass/fail judgments for only one

maneuver within each experimental group, the appropriate analysis method was the McNemar Chi-square test for intervening activity effect. The results of this post-hoc analysis are shown in Table 10. The columns in Table 10 represent the individual maneuvers that were performed on IPC#2. 'VOR' is the VOR instrument approach. This maneuver requires both psychomotor skill and is also highly procedural. 'Hold' is the holding procedures and also requires a combination of psychomotor and procedural skill. 'Steep T' represents the steep turns (one to the left and one to the right) that were performed and is almost entirely dependent on psychomotor skill. 'Uns. Att.' is the unusual attitude recovery and is also mainly dependent on psychomotor skill. 'ILS' is the ILS instrument approach and requires coordination of both psychomotor and procedural skills. 'VOR PP' is the VOR partial panel approach and requires the same skills as the earlier VOR approach. 'ATC' is the participants' performance on air traffic control (ATC) communications throughout the flight and is mainly a procedural skill. Improvement by an experimental group on a given maneuver is indicated by a (+) preceding the McNemar chi-square statistic. A (-) indicates that deterioration on that maneuver occurred. In order for an individual maneuver to show a significant change in performance between IPC #1 and IPC #2 for a given group, each McNemar statistic must be rejected against a Chi-Square critical value determined by $df=1$ and the desired p-value. Rejection of the null hypothesis for no improvement was made if, $\chi^2 (1, N=25) \geq 3.84$. One result of the analysis shows that all three training groups (i.e., Aircraft, FTD or PCATD) showed a trend of improvement in performance on most of the seven maneuvers between IPC #1 and IPC#2. Several maneuvers showed a significant improvement in performance prior to continuity correction. These differences, except for the VOR approach in the Frasca experimental group, all became non-significant after continuity correction. Thus, any determination of improvement on individual maneuvers between IPC #1 and IPC #2 is less clear. It is note worthy that both the PCATD and Frasca groups performance deteriorated on ATC communication. Although the result was not significant, if the trend is an important one, it indicates a possible drawback to practicing communications outside of the actual ATC environment.

Table 10. *McNemar Statistics for change in performance between IPC #1 and IPC #2 for individual maneuvers; (+) indicates improvement and (-) indicates deterioration from IPC #1 to IPC#2.*

Maneuvers: Group	VOR	Hold	Steep T.	Uns Att	ILS	VOR PP	ATC
Aircraft	(+) 0.40	(+) 0.82	(-) 0.40	(+) 0.40	(+) 0.09	(+) 0.67	(+) 0.33
PCATD	(+) 0.5	(+) 3.77*	(+) 0.5	(+) 4.00**	(+) 4.50**	(+) 4.50**	(-) 1.00
Frasca	(+) 7.36**	(+) 3.00*	(+) 0.50	(-) 2.00	(+) 1.60	(+) 4.50**	(-) 1.00

**Significant at $p < .05$

*Significant at $p < .10$

3.5. Demographic Factors

The next analyses compared the demographic between the three experimental groups. The results show that all demographic effects were balanced among the groups and hence did not influence any of the between-groups results (see Table 11).

Table 11. *Between participants/within group ANOVA results of demographic factors.*

Factor	F	p
Prior Currency	0.00	1.00
Age	0.27	0.76
Flight Time	0.65	0.52
Recent Exp.	0.13	0.88

3.6. Inter-Rater Reliability

Three analyses were performed to (1) determine that all 4 check pilots were assigned to equal proportions of IPC #1 and IPC #2 flights between experimental groups, (2) examine the extent of agreement on the check pilots' ratings on the same participant, and (3) assess overall interrater reliability.

A Friedman test for j matched groups was completed to determine if the IPC check pilots had given a proportionally equal number of IPC #1 and IPC #2 sessions within each experimental group. Rankings were established for each check pilot (see Table 12). No significant difference were found, $\chi^2(2, N = 12) = .375, p > .10$, and $\chi^2(2, N = 12) = .5, p > .10$ respectively for IPC #1 and IPC#2, indicating that any variability introduced by an individual check pilot was evenly distributed between each of the three experimental group.

Table 12. *Inter-rater reliability*

IP	IPC#1						IPC#2					
	A	A-Rank	P	P-Rank	F	F-Rank	A	A-Rank	P	P-Rank	F	F-Rank
1	8	3	5	2	4	1	5	2	4	1	14	3
2	3	1	9	3	7	2	11	3	10	2	3	1
3	5	3	3	1.5	3	1.5	3	2	1	1	4	3
4	9	2	8	1	11	3	6	2	10	3	4	1
	Tj	9		7.5		7.5	Tj	9		7		8

Since each participant received IPC #1 and IPC #2 by two different check pilots, it was important to rule out any possible bias effect from a particular pairing of check pilots. In doing so, all possible pairings of the four check pilots within participants for IPC #1 and IPC #2 were compared (see Table 13).

Table 13. *Paired inter-rater analysis of all possible pairings of check pilots and ratings on IPC #1 and 2; Agree = both pass or fail, Disagree = change in IPC score.*

IP Pair	Observed			Expected	
	Agree	Disagree	Totals	Agree	Disagree
BJ/JB	9	11	20	9.87	10.13
BP/PB	4	1	5	2.47	2.53
BT/TB	9	7	16	7.89	8.11
JP/PJ	1	1	2	0.98	1.01
JT/TJ	9	12	21	10.36	10.64
PT/TP	5	6	11	5.43	5.57
Totals	37	38	75		

No association between agreements/disagreements and any check pilot pairings, $\chi^2 (5, N = 75) = 2.76, p > .10$, was found indicating that no one particular pairing of check pilots scored IPC performance significantly different than any other pairing.

In conventional test-retest situations, it is customary to calculate the inter-rater reliability for all raters (e.g. check pilots) who administer the same test. In this project, a single inter-rater score for all four check pilots would not take into account that the same check pilot never gave the same participant the IPC test twice. Under these circumstances, it is appropriate to consider each check pilot's inter-rater reliability across participants to whom they administered the IPC, since each check pilot gave a proportionally equal number of IPC sessions to each experimental group and to each currency level. A z-test was then performed on the group of four inter-rater reliability scores and no significant difference was found (see Table14).

Table 14. *Inter-rater Agreement (number of Agreements/number of IPCs given)*

IP	Agreement	
1	0.55	
1	0.44	Mean =0.50
3	0.53	SD = 0.05
4	0.48	Z-test p-value = .66

Although no significant difference between check pilot inter-rater reliability was found, this analysis may be confounded by a change in participant performance as a result of group assignment. However, it would be unlikely, considering the lack of differences in performance between experimental groups, that participant performance has significantly affected the inter-rater reliability.

After all experimental sessions were complete, a short questionnaire was administered to the four IPC checkpilots to capture their opinions on the use of the aircraft, PCATD and FTD for administering IPCs (see Appendices F and G). Responses to the questions varied widely, making any meaningful analysis difficult. However, one point on which all checkpilots agreed was that all favored the use of an aircraft over either an FTD or PCATD for giving IPCs, regardless of prior knowledge about the particular IPC applicant. Also, in one open-ended question, 3 of the 4 checkpilots cited control fidelity issues as a hindrance to effectively administering IPCs in the ground training devices.

4. Discussion

This study has demonstrated that there are no significant differences in performance by instrument pilots on an IPC given in either a PCATD, and FTD or an airplane. No significant difference was found on IPC #1 among the three groups, which indicates that the participants performed the same regardless of the device in which they had the IPC. In addition there was no significant difference on IPC #2 indicating that the device in which the participants had IPC #1 had no influence on their performance on IPC #2 in the airplane. The planned comparisons showed that performance on IPC #2 of the PCATD group was statistically indistinguishable from both the airplane and the FTD groups. In addition, there was no difference in performance between the aircraft and the FTD groups. These findings present compelling evidence that the FAA should permit the use of PCATDs to give IPCs.

It was expected that performance on IPC #1 would be a good predictor of performance on IPC#2. A comparison of the pass/fail rates for IPC #1 and IPC #2 indicated that the performance on the baseline IPC was not a good predictor of performance on the final IPC. Only 58 percent of the participants who passed IPC #1 also passed IPC #2 and only 45 percent of the participants who failed IPC #1 also failed IPC #2. Only 49 percent of the participants either passed both tests or failed both tests, while 51 percent of the participants passed IPC #1 and failed IPC #2 or failed IPC #1 and passed IPC #2. Therefore performance on IPC #1 predicts performance on a second IPC at a chance level.

The McNemar change in performance between IPC #1 and IPC #2 for all participants was significant but the comparisons for the individual three groups were not significant. Some of the failures may be related to a lack of familiarity with the PCATD, the FTD and the Sundowner airplane, since few of the participants had flown either of the devices prior to the study. The familiarization flights in each of the devices were expected to provide sufficient familiarity with the devices to eliminate the problem but apparently failed to do so. It is possible that additional familiarity with instrument flying in each device, in addition to the VFR familiarization, was needed. The former was not done in order to minimize a possible training effect on group assignment.

Since there was a large change in performance between IPC #1 and IPC #2 for each of the groups, improvement and deterioration were computed for each group. The improvement and deterioration ratios for the three groups are very similar. Tests of significance indicated that for all three groups, the performance was no more likely to improve than to deteriorate.

Of the 53 participants who were instrument current, only 19 (36 %) passed IPC #1. The earlier study by Taylor et al. (2001) and Talleur et al. (2003) showed that 42 % of the instrument current pilots passed the initial IPC. The results from the current study are only slightly worse in this regard than those from earlier studies. In addition, most of the participants tested in the previous study had not taken an IPC after the test was standardized to include required maneuvers (thereby increasing the difficulty of the IPC test). This finding raises questions concerning the relationship between instrument currency and instrument proficiency. Less than half of the participants were able to demonstrate instrument proficiency in an IPC in the airplane. This suggests the need for the FAA to consider changing the recency of experience requirements for instrument currency. Taylor et al. (2001) made the same observation and the current study reinforces the concern that currency rules are inadequate for instrument pilots to maintain proficiency. As Taylor et al. (2001) suggested, an alternative approach would be to require a periodic IPC to demonstrate instrument proficiency in addition to the current currency requirements.

Analysis of the change in maneuver performance between IPC #1 and IPC #2 showed no significant differences between groups. In the study by Taylor et al. (2001) individual maneuvers performed during IPC #2 showed that the PCATD was more effective than either the Aircraft or the FTD in terms of the number of maneuvers that were scored as passes by the check pilot. One possible explanation for the result in the current study is that all participants were not instrument current prior to IPC #1 as was the case in the previous project.

The effectiveness of the PCATD for training specific maneuver elements (i.e., altitude control, airspeed control, navigation procedures, etc.) was observed by comparing performance on subsets of maneuver elements between the experimental groups. None of the group comparisons on procedural/control element change in performance were significant. The study by Taylor et al. (2001) showed a significant improvement for the FTD group on procedural elements on the hold relative to the Aircraft and Control groups, which is similar to the findings of Homan and Williams (1997) as well as Taylor and Stokes (1986), and Taylor (1985). The PCATD group showed a significant improvement on control elements for the ILS approach. This result appears to contradict the finding by Dennis and Harris (1998) that inferred that the PCATD was not effective for practicing psychomotor skills. However, it is well accepted that instrument flight tasks may require differing levels of psychomotor skills than the visual tasks such as those examined by Dennis and Harris (1998). The current study's finding is not unexpected, however. In Taylor et al. (2001), recent experience requirements were met using the PCATD and FTD (according to group assignment) as a means of maintaining instrument currency. This extra exposure to instrument practice in ground training devices accounted for the observed change in performance. Since the current project participants had no such practice in the PCATD, or FTD, significant improvements in maneuver element performance were not expected.

The effect of pilot experience as an explanation for observed variability in data has been reviewed by Taylor (1985) and Taylor and Stokes (1986). The participants in the present study had a wide range of piloting experience which could potentially affect piloting performance. A

biographical questionnaire was completed on each participant so that demographic data could be incorporated into the analysis. No significant differences for any demographic factors between groups were found; thus we conclude that the effect of pilot experience was balanced across all groups.

5. Recommendations for Rulemaking

Based on the findings of the study reported here, three specific recommendations for future rulemaking can be made:

1. We recommend that the FAA permit the use of approved PCATD to give Instrument Proficiency Checks.
2. We recommend that the FAA consider changing recency of experience requirements for instrument currency. This recommendation was also made by Taylor et al. (2001).
3. While not germane to the present study, the following recommendation from the Taylor et al. (2001) study is reiterated: “We recommend that the FAA permit the use of approved PCATD to meet recency of experience requirements.”

6. Summary and Conclusions

This study has demonstrated that PCATDs are as effective as either the airplane or the FTD in conducting Instrument Proficiency Checks. The study involved 75 participants (25 participants in each group: FTD, PCATD and airplane). Each agreed to refrain from instrument flight (either in flight or in a ground-based device) for the duration of the study. Each participant flew a familiarization flight in the FTD, the PCATD and the airplane prior to being randomly assigned to one of the three groups (FTD, PCATD and airplane). The instrument pilots in the study were in one of four categories of instrument currency: 1) instrument current, 2) within one year of currency, 3) between one and two years of currency, and between 2 and 5 years of currency and they were balanced among the three groups.. Pilots in the 2 to 5 year category received up to five hours of proficiency training in either a FTD or a PCATD prior to the experiment

The results indicated no significant differences in performance by instrument pilots on an IPC given in either a PCATD, and FTD or an airplane. Performance on the IPC of the PCATD group was statistically indistinguishable from both the airplane and the FTD groups. In addition, there was no difference in performance between the aircraft and the FTD groups. These findings present compelling evidence that the FAA should permit the use of PCATDs to give IPCs.

It was expected that performance on IPC #1 would be a good predictor of performance on IPC#2. A comparison of the pass/fail rates for IPC #1 and IPC #2 indicated that the performance on the baseline IPC was not a good predictor of performance on the final IPC. The results indicated that the prediction was no better than chance; only 58% percent of the participants who passed IPC #1 also passed IPC #2 and 45% percent of the participants who failed IPC #1 also failed IPC #2

The change in performance between IPC #1 and IPC #2 for all participants was statistically significant, but none of the comparisons of groups were significant. Analyses to determine the performance changes between IPC #1 and the IPC #2 for each experimental group were

conducted and improvement and deterioration ratios were calculated. The improvement and deterioration ratios for the three groups were very similar. Only 24 of the 75 participants passed IPC #1 (32%), and 51 failed (68%). Of the 53 instrument current pilots, only 19 (36%) passed IPC #1 and 34 failed (64%), but 30 passed IPC # 2 (57%) in the aircraft and 23 failed (43%). These results are comparable to those of an earlier study by Taylor, Talleur, Bradshaw, Emanuel, Rantanen, Hulin and Lendrum (2001).

This finding raises questions concerning the relationship between instrument currency and instrument proficiency. Of the participants who were instrument current, less than half of the participant population was able to demonstrate instrument proficiency in an IPC in the airplane. This suggests the need for the FAA to consider changing the recent experience requirements for instrument currency. Alternative approaches would include requiring a periodic IPC to demonstrate instrument proficiency in addition to the current currency requirements, and/or tailoring the IPC content, or recent experience requirements to the type of flight activities the pilot engages in regularly. For example, a pilot who does not regularly fly on instruments, may have to show instrument competency more often than pilots who fly in IMC regularly.

7. References

- Campbell, S. (1998, September). PC trainers in the real world. *IFR*, 14, 16-18.
- Dennis, K. A. (1994). *Computer-based simulation as adjunct to the initial stages of private pilot device training*. Cranfield University.
- Dennis, K. A. & Harris, D. (1998). Computer-based simulation as an adjunct to ab-initio flight training. *International Journal of Aviation Psychology*, 8(3), 261-276.
- Hampton, S., Moroney, W., Kirton, T. & Biers, D. W. (1994). *The use of personal computer-based training devices in teaching instrument flying: A comparative study* (Final Report). Daytona Beach, FL: Embry-Riddle Aeronautical University.
- Homan, W. J., & Williams, K. D. (1997). Effects of PC-based pretraining on pilot's performance in an approved flight-training device. *The Journal of Aviation/Aerospace Education and Research*, 7(3), 19-28.
- Karp, M. R. (2001). Using the PC-based aviation training device (PCATD) in the classroom. *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus, OH. The Ohio State University.
- Kolano, E. (1997, October). Personal computer primer: at the controls of a new instrument training device. *Flight Training*, 56-58.
- Koonce, J. M. & Bramble, M. J. (1998). Personal computer based flight training devices. *International Journal of Aviation Psychology*, 8(3), 277, 292.
- Lendrum, L., Taylor, H. L., Talleur, D. A., Hulin, C. L., Bradshaw, G. L. & Emanuel, T. W., Jr. (1999). Airborne flight data recorders. In R. S. Jensen, B. Cox, J. D. Callister, & R. Lavis

- (Eds.), *Proceedings of the 10th International Symposium of Aviation Psychology* (v. 2, pp. 1003–1007). Columbus, OH: The Ohio State University.
- Lendrum, L., Taylor, H. L., Talleur, D. A., Hulin, C. L., Bradshaw, G. L. & Emanuel, T. W., Jr. (2000). *PC data logger (a flight data recorder): Operation manual* (Technical Report ARL-00-8/FAA-00-5). Savoy, IL: University of Illinois, Institute of Aviation, Aviation Research Lab.
- Miller, W. (1996, May 1). Flying the computer. *Aviation Safety*, 12-14.
- Ortiz, G. A. (1994). Effectiveness of PC-based flight simulation. *The International Journal of Aviation Psychology*, 4, 285-291.
- Peterson, C., Jr. (1993, April). Controlling your simulator software. *Private Pilot*, 72-77.
- Phillips, S. I., Hulin, C. L. & Lamermayer, P. J. (1993). Uses of part-task trainers in instrument flight training. *Proceedings of the 7th International Symposium of Aviation Psychology* (pp. 743-746). Columbus, OH: The Ohio State University.
- Phillips, S. I., Taylor, H. L., Lintern, G., Hulin, C. L., Emanuel, T. W., Jr., & Talleur, D. A. (1995). Developing performance measures for evaluating personal computer-based aviation training devices within an FAR part 141 pilot training school. *Proceedings of the 8th International Symposium on Aviation Psychology*. Columbus: The Ohio State University.
- Talleur, D. A., Taylor, H. L., Emanuel, T. W., Jr., Rantanen, E., & Bradshaw, G. L., (2003). Personal computer aviation training devices: Their effectiveness for maintaining instrument currency. *The International Journal of Aviation Psychology*, 13(4), 387-399.
- Taylor, H. L. (1985). *Training effectiveness of flight simulators as determined by transfer of training*, Keynote Address, NATO Symposium on the Transfer of Training to Military Operational systems. Brussels, Belgium: Headquarters, NATO.
- Taylor, H. L., Lintern, G., Hulin, C. L., Talleur, D. A., Emanuel, T. W. & Phillips, S. I. (1996). *Transfer of training effectiveness of personal computer-based aviation training devices* (Tech. Rep. ARL-96-3/FAA-96-2). Savoy: University of Illinois, Aviation Research Laboratory.
- 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. *The International Journal of Aviation Psychology*, 9(4), 319-335.
- 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, 1*, 81-129. Praeger Press.
- Taylor, H. L., Talleur, D. A., Bradshaw, G. L., Emanuel, T. W., Jr., Rantanen, E., Hulin, C. L., & Lendrum, L. (2001). *Effectiveness of personal computers to meet recency of experience*

requirements. (Tech. Rep. ARL-01-6/FAA-01-1). Savoy: University of Illinois, Aviation Research Laboratory.

Taylor, H. L., Talleur, D. A., Emanuel, T. W., Jr., Rantanen, E. M., Bradshaw, G. L. & Phillips, S. I. (2001). *Incremental training effectiveness of personal computer aviation training devices (PCATD) used for instrument training devices* (Tech. Rep. ARL-02-5/NASA-02-3). Savoy: University of Illinois, Aviation Research Laboratory.

Taylor, H. L., Talleur., G. L., Emanuel, T. W., Jr., Rantanen, E. M., & Phillips, S. I. (2001). Incremental training effectiveness of personal computers used for instrument training. *Proceedings of the 11th International Symposium of Aviation Psychology*. Columbus, OH: The Ohio State University.

Taylor, H. L., Talleur, D. A., Phillips, S. I., Emanuel, T. W. & Hulin, C. L. (1998). Use of personal computers for instrument training: *Proceedings of the Royal Aeronautical Society: Low cost simulation – new opportunities in flight training* (pp. 5.1 – 5.14). London, UK.

U.S. Department of Transportation. (1997). *Qualification and approval of personal computer-based aviation training devices*, AFS-840 (AC No: 61-126). Washington, DC: Federal Aviation Administration, U.S. Department of Transportation.

U.S. Department of Transportation. (1999, March). *Instrument rating for airplane, helicopter and airship: Practical test standards* (FAA-S-8081-4C). Washington, DC: Federal Aviation Administration, U.S. Department of Transportation.

Williams, K. W. (1994). *Summary proceedings of the joint industry-FAA conference on the development and use of pc-based aviation training devices* (DOT/FAA/AM-94125). Oklahoma City, OK: Federal Aviation Administration, Civil Aero medical Institute.

Williams, K. W., & Blanchard, R. E. (1995). *Development of qualification guidelines for personal computer-based aviation training devices* (DOT/FAA/AM-96/6). Oklahoma City, OK: Federal Aviation Administration, Civil Aeromedical Institute.

APPENDIX A

REPORTS AND PRESENTATIONS

Semi-Annual and Annual Reports

Taylor, H L, Talleur, D A, Rantanen, E M, Emanuel, T W Jr., (2002). Comparison of the effectiveness of a Personal Computer Aviation Device, a flight training device, and an airplane in conducting instrument proficiency checks. Semi-annual report. April 10, 2002.

Taylor, H L, Talleur, D A, Rantanen, E M, Emanuel, T W Jr., (2002). Comparison of the effectiveness of a Personal Computer Aviation Device, a flight training device, and an airplane in conducting instrument proficiency checks. Semi-annual report. September 26, 2002.

Taylor, H L, Talleur, D A, Rantanen, E M, Emanuel, T W Jr., (2002). Comparison of the effectiveness of a Personal Computer Aviation Device, a flight training device, and an airplane in conducting instrument proficiency checks. A report. November 27, 2002.

Taylor, H L, Talleur, D A, Rantanen, E M, Emanuel, T W Jr., (2002). Comparison of the effectiveness of a Personal Computer Aviation Device, a flight training device, and an airplane in conducting instrument proficiency checks. Semi-annual report. April 9, 2003.

Taylor, H L, Talleur, D A, Rantanen, E M, Emanuel, T W Jr., (2002). Comparison of the effectiveness of a Personal Computer Aviation Device, a flight training device, and an airplane in conducting instrument proficiency checks. Semi-annual report. September 20, 2003.

Taylor, H L, Talleur, D A, Rantanen, E M, Emanuel, T W Jr., (2002). Comparison of the effectiveness of a Personal Computer Aviation Device, a flight training device, and an airplane in conducting instrument proficiency checks. Semi-annual report. April 2, 2004.

Published Reports

Emanuel, T W Jr., Taylor, H L, Talleur, D A, Rantanen, E M. (2003). Comparison of the effectiveness of a Personal Computer Aviation Training Device, a flight training device and an airplane in conducting an instrument proficiency check. Proceedings of the 12th International Symposium on Aviation Psychology, 356-360. April 12-17, 2003, Dayton, OH.

Taylor, H L, Emanuel, T W Jr., Talleur, D A, Rantanen, E M. (2003). The effectiveness of a Personal Computer Aviation Device, a flight training device, and an airplane in conducting instrument proficiency checks. FAA Annual Program Review. September 10, 2003, Reno, NV.

Johnson, N. R., Rantanen, E. M., & Talleur, D. A. (In press). Criterion setting for objective, Fourier analysis based pilot performance metrics. *Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society*.

Johnson, N. R., Rantanen, E. M., & Talleur, D. A. (2004). Time series based objective pilot performance measures. *International Journal of Applied Aviation Studies (IJAAS)*, 4(1), 13-29.

Presentations

Taylor, H L, Talleur, D A, Rantanen, E M, Emanuel, T W, Jr. (2004). The Effectiveness of a Personal Computer Aviation Training Device (PCATD), a Flight Training Device (FTD) and an Airplane in Conducting Instrument Proficiency Checks (IPC), Presented at the Research

Roundtable, Fall Education Conference of the University Aviation Association, October 8, 2004, Toronto, CA, Abstract to be published.

Taylor, H L, Talleur, D A, Emanuel, T W, Jr., Rantanen, E M. (2004). The Effectiveness of a Personal Computer Aviation Training Devices (PCATDs), and Flight Training Devices (FTDs) in Instrument Training, Maintaining Currency, and Instrument Proficiency Checks, Presented at the Institutional Workshop, Fall Education Conference of the University Aviation Association, October 9, 2004, Toronto, CA, Power Point slides to be published.

Taylor, H L, Talleur, D A, Rantanen, E M, Emanuel, T W, Jr. (2004). The Effectiveness of a Personal Computer Aviation Training Device (PCATD), a Flight Training Device (FTD) and an Airplane in Conducting Instrument Proficiency Checks (IPC), Presented at the FAA FY 04 General Aviation, Vertical Flight, and Aviation Maintenance Human Factors Program Review, Washington D. C., November 16, 2004 Poster.

Taylor, H L, Talleur, D A, Emanuel, T W, Jr., Rantanen, E M. (2004) The Effectiveness of personal computers (PCATDs) and flight training devices (FTDs) on instrument training for pilots. 75th Annual Scientific Program Meeting of the Aerospace Medical Association, May 2004, Anchorage, AK. (Abstract).

Taylor, H L, Talleur, D A, Emanuel, T W, Jr., Rantanen, E M (2004). The Effectiveness of Personal Computer Aviation Training Devices (PCATDs), flight Training devices (FTDs) and an airplane in conducting instrument proficiency checks. Annual Midyear Symposium of APA Division 19 and 21 and the Potomac Chapter of the HFES, March 5, 2004, Ft. Belvoir, VA.

Taylor, H L, Talleur, D A, Emanuel, T W, Jr., Rantanen, E M (2003). Comparison of the effectiveness of PCATDs, FTDs, and aircraft in conducting instrument proficiency checks. Research Roundtable, University Aviation Association Fall Education Conference, October 24, 2003, Dayton, OH.

Taylor, H L, Talleur, D A, Emanuel, T W, Jr., Rantanen, E M (2003). The effectiveness of a personal computer (PCATD), an aviation training device (FTD), and an airplane in conducting instrument proficiency checks. Technology Enhancements for Aviation Classrooms Seminar, University Aviation Association Fall Education Conference, October 22, 2003, Dayton, OH.

Taylor, H L, Emanuel, T W Jr., Talleur, D A, Rantanen, E M.(2003) The effectiveness of a Personal Computer Aviation Device, a flight training device, and an airplane in conducting instrument proficiency checks. FAA Annual Program Review. September 10, 2003, Reno, NV.

Taylor, H L, Talleur, D A, Emanuel, T W, Jr., Rantanen, E M, Bradshaw, G L, Phillips, S I. (2003) Effectiveness of personal computers for training instrument pilots, Annual Scientific Program Meeting of the Aerospace Medical Association, May 2003, San Antonio, TX,(Abstract).

Emanuel, T W Jr., Taylor, H L, Talleur, DA, Rantanen, EM. (2003). Comparison of the effectiveness of a Personal Computer Aviation Training Device, a flight training device and an airplane in conducting an instrument proficiency check. 12th International Symposium on Aviation Psychology, April 12-17, 2003, Dayton, OH.

Emanuel, T W, Jr., Taylor, H L, Talleur, D A, Rantanen, E A. (2002). Comparison of the effectiveness of a personal computer aviation training device, a flight training device, and an airplane in conducting instrument proficiency checks, Research Roundtable, University Aviation Association Fall Education Conference, September 13, 2002, Orlando, FL.(Abstract).