Comparing the Accuracy of Performing Digital and Paper Checklists Using a Feedback Intervention Package During Normal Workload Conditions in Simulated Flight

William G. Rantz Ph.D.
Bryan W. Hilton B.S.

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2009

Part of the Other Psychiatry and Psychology Commons

Repository Citation
https://corescholar.libraries.wright.edu/isap_2009/37

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2009 by an authorized administrator of CORE Scholar. For more information, please contact library-core scholar@wright.edu.
This study examined whether pilots completed airplane digital and paper checklists more accurately when they received post-flight graphic and verbal feedback. Participants were 6 college student pilots with instrument ratings. The task consisted of flying flight patterns using a Frasca 241 Flight Training Device which emulates the Cirrus SR20. An alternating treatment, multiple baseline design across pairs with reversal was used. Visual inspection and statistical analysis of the data suggests that paper checklist accuracy does not differ significantly from digital checklist accuracy during normal workload conditions. The results also suggest that graphic feedback and praise can be used to increase the extent to which pilots use both digital and paper checklists accurately.

Understanding the knowledge and behaviors required to effectively manage risk are an integral component of the professional pilot training curriculum (Western Michigan University: Professional Flight Training Program, 2004). The aviation industry demands that professional pilot graduates understand the inherent risks associated with flight operations and that individuals must continue to practice comprehensive preflight planning, attention to detail, procedural discipline, and run the checklists as printed, Wilson, (2008). Checklists organize tasks into sequences of actions that configure the aircraft and prepare the crew for evolving events. “The major function of the checklist is to ensure the crew will properly configure the plane for flight, and maintain this level of quality throughout the flight, and in every flight” (Degani & Wiener, 1990, p. 7). Checklist devices or methods of presentations are described as paper, laminated paper/card, scroll paper, electromechanical, vocal, and computer-aided/electronic. The most common method of presentation for checklists is the laminated paper/card (Degani & Wiener, 1994; Turner & Huntley, 1991). While this statement may be true for all general aviation aircraft manufactured in the last one hundred years, the rise of lower cost computing hardware and software is rapidly changing how newer aircraft present checklists. (Boorman, 2001a, 2001b).

Within the last two decades electronic or digital checklists have appeared on many regional and major airline flight decks, and some general aviation aircraft. These digital checklists are integrated into the new aircraft panel by the manufacturer with software designed to exclude many paper checklist errors observed in past studies (Arkell, 2006; Boorman, 2001a, 2001b). As avionics prices continue to decline, it is very likely more digital checklists will be installed on smaller general aviation aircraft, thereby expanding the demographics of the pilot users from airline professional to recreational novice. The future challenge is not that pilots understand that the checklist is a presentation method by which flight deck safety is enhanced.
The challenge seems to be recognizing the absence of stimulus control in a varying flight environment which may result in unpredictable checklist use.

An extensive review of the checklist literature has found many interesting areas outside of aviation where checklists are employed (Rantz, 2005). From the accident reports, errors using checklists have and continue to plague the aviation industry in particular. Given the number of aviation studies devoted to checklist use and how tasks are conducted on the flight deck, an extensive search of the aviation checklist literature revealed only one study that has examined whether the traditional paper checklist could be a) used as a dependent variable and b) whether behavioral interventions, could increase the appropriate use of flight checklists (Rantz, Dickinson, Sinclair & Van Houten, In Press, p. 20. The purpose of the present study is to compare and if possible improve the accuracy of both the traditional paper and standard digital checklists.

**Method**

**Experimental Design.** An alternating treatment, multiple baseline design with reversal plus an over sixty day probe across pairs of participants was used to compare paper and digital checklists and evaluate the effect of feedback on checklist use. There were four phases of the experiment, baseline, intervention, reversal, and probe. Sessions lasted approximately two hours and 6 participants flew four different flight patterns per session using the Frasca 241, Cirrus SR20 flight training device. Each flight was considered a trial, and checklist performance was scored and graphed separately for each trial. Each flight lasted approximately 20-25 minutes. There were six different flight patterns. The order of exposure to the flight patterns was randomized in blocks of six for each participant. This procedure insured no two patterns were repeated during one session. Paper or digital checklists were randomly assigned for the initial trial at the beginning of each session. For the remaining three sessions, paper or digital checklists were alternately assigned.

**The flight checklists.** The digital and paper checklists each contained 70 identical checklist items divided into sections that corresponded to each of the eight flight segments. The digital checklist was an integrated function of a multifunctional display (MFD) produced by Avidyne. The MFD model was the Entegra EX5000C used in Cirrus SR20 aircraft. The paper checklist was a spiral bound booklet provided for use in the Cirrus SR20 (Pilot’s Checklist Cirrus SR20, 2002), both the digital and paper checklists are used in the colleges’ flight training curriculum. The digital checklist display, when used, was in a fixed position ahead and slightly to the right of the pilot’s central view. The paper checklist, when used, was positioned on the right leg or lap of the participant and when not used usually remained on the seat beside the participant.

**Dependent Variables.** The primary dependent variable consisted of the number of paper or digital checklist items completed correctly per flight.

**Independent Variable.** There were two independent conditions during this study, using paper or digital checklists. The independent variable was the presence or absence of post-flight (a) graphic feedback on the total number of checklist items completed correctly per flight, (b) graphic feedback on the number of items completed correctly, incorrectly, and omitted for each of the eight flight segments per flight, and (c) praise for improvement in the number of checklist items completed correctly.
Inter-observer Agreement (IOA). A second observer watched randomly selected recordings of the flights and scored performance using the checklist observation form. This process was repeated for each participant. This ensured that (a) at least 25% of the sessions were rescored for each participant, and (b) the trials that were rescored were randomly selected. Inter-observer agreement was determined for the total number of checklist items completed correctly. Inter-observer agreement was calculated as follows: number of agreements divided by the number of agreements plus disagreements, multiplied by 100. Inter-observer agreement for correct and incorrect item errors was an average of 95% with a range of 79% to 100%.

Results

Figure 1 displays the total number of paper checklist items completed correctly (open circles) together with the total number of digital checklist items (closed circles) completed correctly for each participant per trial. All participants increased paper and digital checklist performance accuracy over baseline when post-flight graphic feedback was provided and those improvements remained during the withdrawal phase and during a delayed probe. Baseline paper and digital checklist performance varied considerably across participants with participant 1 showing the lowest level of performance in both paper (average 87% error) and digital (average 89% error). Participant 1 had a mean average of 3.37 correct for digital checklist items and 6.11 correct for paper. Both participant 2 and 4 showed the highest level in paper checklist (average 43% error) and both participant scored a mean average of 39.67 correct. However participant 2 averaged 33% error for the highest performance in digital checklist use with a mean average of 44.28 correct. Baseline trends were fairly stable over time for four participants (P1, P3, P4, and P5), with the exception of participant 2, who despite overall high mean average scores, performed one high peak in digital and one high peak in paper performance and showed a steady, overall decline in both paper and digital accuracy from the first trial and participant 6 who showed wide variability between paper (38.92 mean average correct) and digital (37.89 mean average correct) performance. Overall paper checklist baseline performance averaged 62% errors (27.42 mean average) for all participants while digital baseline performance averaged 61% errors (26.57 mean average).

Overall performance in both paper and digital checklist accuracy increased for all participants after the intervention was introduced. There was a dramatic intervention effect using both paper and digital checklists for both individual participants and cumulative across all participants. Two participants (P1 & P3) showed an abrupt level change of over 50% improvement in the first trial, following the introduction of the treatment and then continued an increasing trend. Participant 1 initially increased paper checklist accuracy 71% after the intervention, improving total correct checklist items from 1 item correct out of 70 items to 51 items correct. Participants 3 had the highest initial performance increase across both the digital and paper checklists, increasing level change by 61% for digital checklist items done correctly and 44% for paper checklist items. Participant 5 experienced an initial increase in level change of 36% improvement for digital and 40% improvement for paper checklist items performed correctly. Two participants showed an increasing level change, for both digital and paper, followed by an increasing trend (P2, P4). Only Participant 6, while initially increasing 13% in paper checklist performance, demonstrated an initial single trial decrease of 1% in digital accuracy followed by an increasing digital trend.
Performance criteria for the reversal phase was established as three consecutive trials in either paper or digital where participant’s checklist performance met or exceeded 95% correct on checklist items. All participants reached reversal criteria during paper checklist trials.

Overall across all participants, the average percentage of paper checklist items completed correctly increased from 38% items correct during the baseline phase to 90% items correct during the intervention phase. The average percentage of digital checklist items increased from 39% items completed correctly to 89% items correct during intervention. Improvement continued to near perfect levels for participants during the reversal phase with 100% paper checklist items correct and 99% digital items correct. The average percentage of paper checklist performance declined 3% between a 60 and 90 day delay. The average percentage of digital checklist performance declined 4% during that same time period.

Data contained in Figure 1 were used in the inferential analysis given the model, $Y_t = \beta_0 + \beta_1d_1 + \beta_2d_2 + \beta_3d_3 + \beta_4d_4 + \beta_5d_5 + \phi_1\gamma + \varepsilon_t$. The parameters of this model were estimated for each participant using the bootstrap based time-series regression method described in McKnight, McKean, and Huitema (2000). Results were statistically significant for each individual’s intervention effect in both paper and digital checklist use. Performance was generally not significant once optimum performance levels were reached during each following phase.

After parameter estimates for each participant were computed they were used as dependent variable scores in the group level analysis. The purpose of this analysis was to provide an overall evaluation of the effects of the interventions for the group of six pilots. The next stage of the group analysis consisted of conventional one-sample $t$-tests to evaluate the hypothesis that each intervention and phase-change parameter value is equal to zero. Once again, results were statistically significant for the overall intervention effect in both paper and digital checklist use. As for each individual’s results above, performance was generally not significant once optimum performance levels were reached during each following phase.

The third aspect of the analysis involved computing the difference in performance under the digital and paper conditions at each observation point and testing the difference between the digital and paper means. Once again, the double bootstrap method of McKnight, McKean, and Huitema (2000) was used to estimate the parameters of a time-series model developed to evaluate the hypothesis of zero difference between digital and paper feedback; this is a model that contains only an intercept and an autoregressive parameter. The difference between paper and digital checklist performance was found not to be statistically significant ($t = 1.78, p = .08$).

Discussion

This research is a follow up to the study by Rantz, Dickinson, Sinclair & Van Houten (in press) which evaluated the effects of feedback and praise on the use of a simple personal computer aviation training device and a paper checklist. The present study confirmed the findings of the former study, while using a much higher level of simulation. The current study additionally included comparing pilot’s performance using both paper and electronic checklists during all phases of the experiment. The results of the present study also suggest using graphic feedback and praise can simultaneously improve checklist reading performance in both traditional paper and modern digital presentation modes. The results also indicated, contrary to common opinion, that the use of a digital checklist did not lead to a reduction in errors compared to the traditional paper checklist in a normal workload environment. This study also suggests a
pilot’s checklist performance, regardless of presentation method, may be influenced by common underlying rule-based behaviors (learning history), structured feedback, and particular salient environmental prompts.

References


Figure 1. Total Number of Paper and Digital Checklist Items Completed Correctly.