JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2011, **44,** 145–150 NUMBER 1 (SPRING 2011)

A FEEDBACK INTERVENTION TO INCREASE DIGITAL AND PAPER CHECKLIST PERFORMANCE IN TECHNICALLY ADVANCED AIRCRAFT SIMULATION

WILLIAM G. RANTZ AND RON VAN HOUTEN

WESTERN MICHIGAN UNIVERSITY

This study examined whether pilots operating a flight simulator completed digital or paper flight checklists more accurately after receiving postflight graphic and verbal feedback. The dependent variable was the number of checklist items completed correctly per flight. Following treatment, checklist completion with paper and digital checklists increased from 38% and 39%, respectively, to nearly 100% and remained close to 100% after feedback and praise for improvement were withdrawn. Performance was maintained at or near 100% during follow-up probes.

Key words: checklist use, feedback, flight simulators, flight training, paper versus electronic checklists, rule-governed behavior, simulator use

Managing risks on the flight deck of any aircraft demands that pilots perform tasks in a timely, sequential, and correct manner. Checklists provide a foundation for verifying aircraft configuration, coordinating and enhancing crew performance, and maintaining quality control (Degani & Wiener, 1993). Even with the extensive Aviation Safety Reporting System (ASRA), flight deck observations, and pilot questionnaires, the incorrect use of flight checklists is often cited as the probable cause of or a contributing factor in a large number of crashes (Degani, 1992, 2002; Degani & Wiener, 1990, 1993; Diez, Boehm-Davis, & Holt, 2003).

Checklist formats vary and are comprised of paper, laminated paper or card, scroll paper, electromechanical, vocal, and computer-aided or electronic versions. In the last two decades, electronic or digital checklists have appeared on many regional and major airline flight decks and on some technically advanced general aviation aircraft. These digital checklists have been integrated into the aircraft panel by the manufacturer with software that eliminates many paper checklist errors observed in past studies (Arkell, 2006; Boorman, 2001a, 2001b; Degani & Wiener, 1990).

One drawback of aviation checklist studies is that they have not used the checklist as a dependent variable while manipulating an independent variable. As an exception, Rantz, Dickinson, Sinclair, and Van Houten (2009) demonstrated that the presentation of feedback increased the appropriate use of paper flight checklists from 53% of items completed correctly during baseline to 98% correct. This report is a replication of that study. In addition, a comparison between a modern digital checklist display and a traditional paper checklist was conducted using a technically advanced aircraft simulator.

METHOD

Participants

Six undergraduate students served as participants in this study. To be included, participants had to have a private pilot certificate, an instrument rating, and a minimum of 5 hr of experience using the flight simulator. Flight instructors did not serve as either data collectors or participants.

We thank the College of Aviation for providing support for this study.

Address correspondence to William Rantz, College of Aviation, Western Michigan University, Kalamazoo, Michigan 49008 (e-mail: william.rantz@wmich.edu).

doi: 10.1901/jaba.2011.44-145

Apparatus

The simulator was a Frasca 241, Cirrus SR20 FTD equipped with the Avidyne Entegra EX5000C flight display. The FTD had software capable of recording flight parameters that depicted how well participants flew the designated flight patterns vertically and horizontally.

All flights were recorded and stored digitally with video cameras equipped with microphones for the purpose of data collection. One camera was mounted midway on the glare shield of the instrument panel approximately 61 cm in front of the participant to capture facial position and eye movements. A second camera was positioned 90 cm behind and slightly above the participant (over-the-shoulder view) to observe the participant's arm positions and hand interactions with the flight panel. A third camera was positioned to record the instructor station, thereby ensuring documentation of the visual flight path flown on each trial.

Flight patterns. Six different flight patterns were programmed to be flown during simulation sessions (Rantz et al., 2009). Each pattern had segments that corresponded to the eight checklist segments used for each radar-vectored instrument approach flight. To simulate actual instrument approach realistically and to ensure that each trial was conducted in a consistent way across participants, the experimenter provided scripted air traffic control radar vectors throughout each flight trial. These scripts ensured that turns, altitudes, and headings were consistently assigned at appropriate positions in each flight. Sessions lasted approximately 2 hr, and participants flew four different flight patterns (trials) per session.

Flight checklists. The digital and paper checklists each contained 70 identical items divided into sections that corresponded to each of the eight flight segments. The digital checklist was an integrated Avidyne Entegra EX5000C used in technologically advanced Cirrus SR20 aircraft. The paper checklist was a spiral-bound booklet provided for use in the Cirrus Design SR20 (Pilot's Checklist Cirrus SR20, 2002).

Response Measurement and Reliability

Trained observers scored the number of paper or digital checklist items completed correctly per flight. All observed behaviors were compared to the criteria outlined in the observer's checklist behavior protocol (Appendix). At least 25% of the paper and digital sessions were scored by a second observer for each participant. Interobserver agreement was determined for the total number of checklist items completed correctly and was calculated by dividing the number of agreements by the number of agreements plus disagreements, multiplied by 100%. Mean interobserver agreement for total correct and incorrect item errors was 95% (range, 79% to 100%).

Design

The efficacy of feedback and praise for improvement was evaluated using a multiple baseline design across participants. A multielement design was used to compare digital and paper checklists. A maintenance probe was conducted 60 to 90 days after the completion of the study.

Procedure

Baseline. Participants were informed via scripts that the Frasca 241 FTD simulator was programmed for normal flight and that each flight pattern was a radar-vectored (assigned headings, air speeds, and altitudes) instrument flight, with an instrument landing system (ILS) approach to a full-stop landing. Participants also were informed that their behavior during each flight trial would be observed and recorded using the video cameras. In addition, participants were told that the experimenter would provide them with some technical postflight feedback after each trial and that it would take the experimenter about 5 min to prepare that material. After the participant completed a flight, the experimenter printed out a technical diagram of the flight pattern flown by the participant. These diagrams were automatically created by the simulator software and displayed the lateral and vertical deviations of flight paths from established reference points along the diagram. The experimenter gave the diagram to the participant and discussed the technical merits of the flight, praising adequate flight skill performance. This protocol was repeated for each flight during the baseline phase. Initial use of a digital or paper checklist was randomly assigned for the first trial of each session, and the remaining trials of that session alternated between paper and digital checklists.

Intervention. In addition to providing participants the technical diagram feedback that depicted critical flight parameters after each flight, the experimenter provided feedback on the use of the flight checklist. After each flight, the experimenter immediately calculated the number of checklist items completed correctly, entered it into the computer, and printed a line graph that displayed the number of correctly completed items for each trial, including all previous baseline performances. In another graph, the experimenter entered the number of checklist items completed correctly and incorrectly or omitted for each of the eight flight segments for that particular flight, and printed a bar graph that displayed those data. The experimenter first showed the technical flight diagram to the participant and praised, if appropriate, the technical merits of the flight. He then showed the cumulative line graph depicting only the correct number of checklist items completed for each trial to date. He then presented the bar graph that displayed only the number of correct, incorrect, and omitted items for the previous trial. Participants received praise for any improvements in checklist performance. However, no detailed vocal feedback was given to the participant, such as identifying which particular checklist items were performed incorrectly or omitted or recommending prescriptive behavior to improve checklist use.

Withdrawal. This phase was identical to the baseline phase. Participants received only technical flight diagrams and praise for adequate technical flight skill performance.

Probe. A probe was conducted 60 to 90 days after completion of the study to determine the extent to which performance improvement was maintained. Four alternating trials were completed using both paper and digital checklists. Only technical flight skill feedback was given.

RESULTS AND DISCUSSION

Figure 1 displays the total number of paper and digital checklist items completed correctly per trial by each participant. Baseline performance averaged 38% and 39%, respectively, for paper and digital checklists. Following the introduction of graphic feedback and praise for improvement relative to baseline, checklist performance increased to nearly 100% accuracy with both types of checklists, and these changes were maintained after the treatment was withdrawn. Results of the follow-up probe suggested the potential of the feedback intervention to maintain appropriate checklist behavior at nearly full strength for both presentation methods, over a period of 60 to 90 days.

The present study replicated the findings of Rantz et al. (2009), which showed that feedback increased paper checklist use. It also extended these findings by demonstrating that feedback increased both paper and digital checklist use and that the effects persisted over time after the treatment was terminated. Unfortunately, the current study did not attempt to measure covert behaviors; it is possible that participants formed new rules after receiving the graphic feedback intervention. Future research should examine this possibility.

Paper and digital checklists produced similar levels of performance, a finding that runs counter to industry opinions claiming superiority of the digital format over paper (Boorman, 2001a, 2001b). Given the industry trend to install digital instrumentation with checklists in newly manufactured aircraft, the underlying contingencies of behavior for using those checklists correctly must still be addressed.

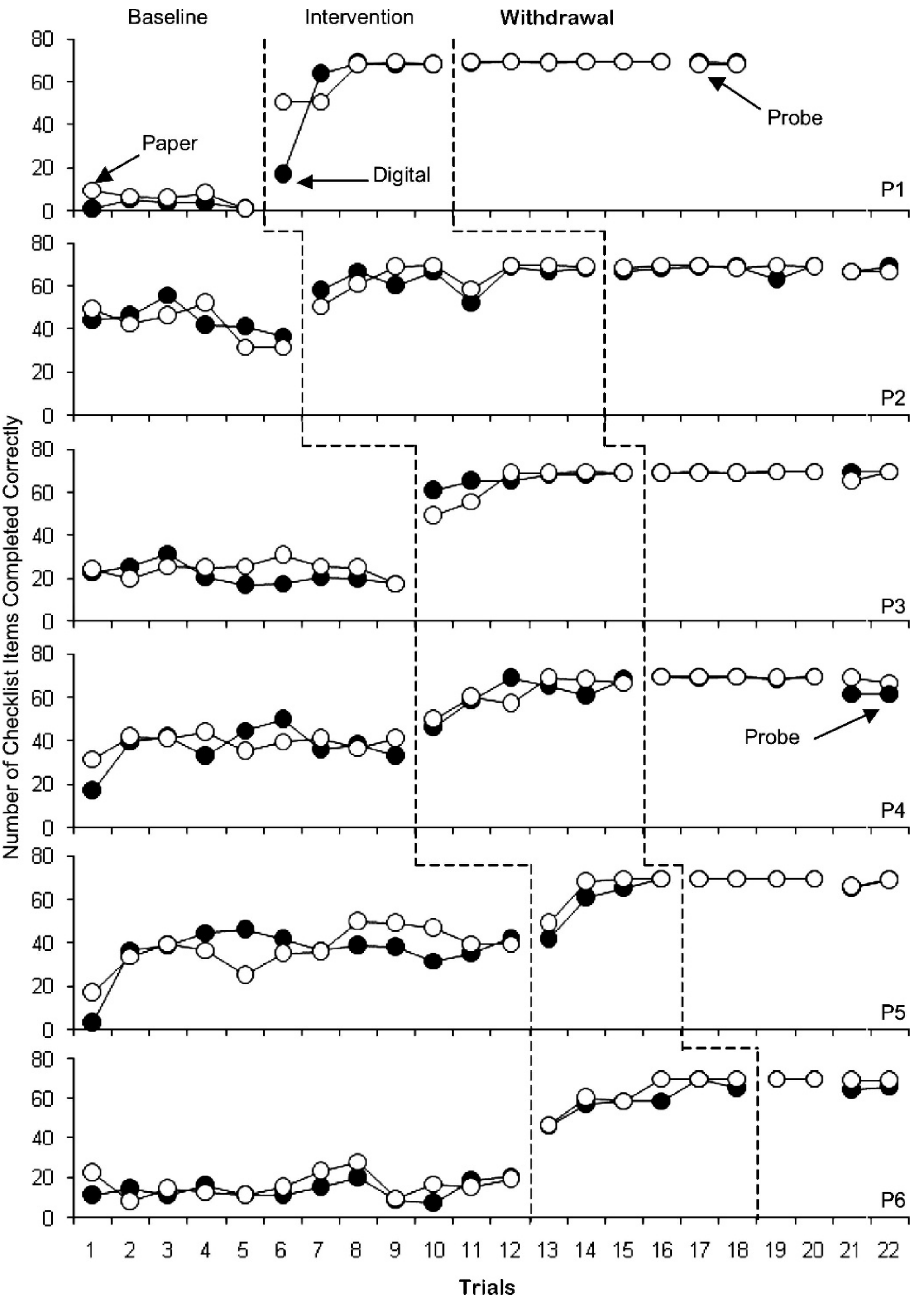

Figure 1. Total number of checklist items completed correctly by each participant. Each trial denotes the sequential trial number for each paper and digital checklist.

Although both checklists have their strengths, their inherent weaknesses include (a) the lack of effective rules for the consistent and proper use of either type of checklist, (b) the lack of salient stimuli to prompt the initiation of the checklist sequence, and (c) the lack of effective reinforcers to increase and maintain checklist use.

There are several possibilities for future research. Those most directly related to the current study include (a) replicating the current procedures and ascertaining whether checklist compliance transfers to actual flight, (b) training pilots to match the start of a checklist sequence with a particular salient stimulus, and (c) examining whether teaching rule statements can improve checklist use.

REFERENCES

Arkell, D. (2006). From safe to safer: Boeing's electronic checklist marks 10 years of enhanced safety for pilots and travelers. *Boeing Frontiers*, *4*(11). Retrieved from http://www.boeing.com/news/frontiers/archive/2006/april/i_ca3.html

Boorman, D. (2001a). Safety benefits of electronic checklists: An analysis of commercial transport accidents. *Proceedings of the 11th International Symposium on Aviation Psychology* (pp. 1–6). Columbus: The Ohio State University.

Boorman, D. (2001b). Today's electronic checklists reduce likelihood of crew errors and help prevent mishaps. *ICAO Journal*, *1*, 17–21.

Degani, A. (1992). *On the typography of flight-deck documentation* (NASA Contractor Rep. 177605). Moffett Field, CA: NASA Ames Research Center.

Degani, A. (2002, April). *Pilot error in the 90s: Still alive and kicking*. Paper presented at the meeting of the Flight Safety Foundation of the National Business Aviation Association, Cincinnati, OH.

Degani, A., & Wiener, E. L. (1990). *Human factors of flight-deck checklists: The normal checklist* (NASA Contractor Rep. 177549). Moffett Field, CA: NASA Ames Research Center.

Degani, A., & Wiener, E. L. (1993). *Cockpit checklists: Concepts, design, and use. Human Factors*, *35*(2), 28–43.

Diez, M., Boehm-Davis, D., & Holt, R. (2003). Checklist performance on the commercial flight-deck. *Proceedings of the 12th International Symposium on Aviation Psychology* (pp. 323–328). Columbus: The Ohio State University.

Pilot's Checklist Cirrus SR20. (2002). P/N 11935-003. Duluth, MN: Cirrus Design Corporation.

Rantz, W. G., Dickinson, A. M., Sinclair, G. A., & Van Houten, R. (2009). The effect of feedback on the accuracy of checklist completion during instrument flight training. *Journal of Applied Behavior Analysis*, *42*, 497–509.

Received September 1, 2009
Final acceptance April 2, 2010
Action Editor, David Wilder

APPENDIX

Sample of Observer's Checklist and Behavior Protocol

1. Doors: tactual contact or pushing door; verbal "latched"
2. CAPS handle: tactual contact to verify pin removed
3. Seat belts and shoulder harnesses: tactual contact and secure
4. Fuel quantity: tactual contact or pointing gesture quarter tank minimum; verbal "confirm"
5. Fuel selector: tactual contact to lever; verbal "fullest tank"
6. Fuel pump: tactual contact to switch; verbal "on"
7. Flaps: tactual contact flap handle; verbal "set 50% and check"
8. Transponder: tactual contact code set to 1200; verbal "set"
9. Autopilot: tactual contact; verbal "check"
10. Navigation radios: eye contact; tactual contact set comm freq. _____ tower; set nav freq. _____ILS OBS inbound course aligned; verbal "set for takeoff"
11. Cabin heat/defrost: tactual contact; verbal "as required"
12. Brakes: tactual contact with feet; verbal "hold"
13. Power lever: tactual contact; verbal "1700 rpm"
14. Alternator: tactual contact; verbal "check"

15. Pilot heat: tactual contact pilot switch; verbal "on"
16. Navigation lights: tactual contact light switch; verbal "on"
17. Landing light: tactual contact light switch; verbal "on"
18. Annunciator lights: tactual contact; verbal "check"

Eye contact: Refers to the participant's behavior of looking in the direction of a discriminative stimulus such as an instrument, lever, switch, or object.

Tactual contact: Refers to the participant's behavior of moving a finger or hand to touch a discriminative stimulus.

Pointing gesture: Refers to the participant's behavior to extend a directed finger at a discriminative stimulus.

Verbal: Refers to the participant's vocal behavior directed at tacting the condition or state of the discriminative stimulus.

Correct response: A response that uses the behavior of looking at the correct checklist item; moving an arm, hand, or finger to touch or point to the correct item; vocal behavior identifying the correct item and tacting the correct condition or state of the item.

Incorrect response: A response that uses the behavior of looking at the incorrect checklist item; moving an arm, hand, or finger to touch or point to the incorrect item; vocal behavior identifying an incorrect item or tacting the incorrect condition or state of the item; or any correct response not accompanied by correct vocal behavior.

Omitted response: An absence of responding.