Strategies for Controlling Checklist Reading Behavior: A Literature Review

William G. Rantz

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2005

Part of the Other Psychiatry and Psychology Commons

Repository Citation

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 - 2005 by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu, library-corescholar@wright.edu.
One of the highest frequencies of errors recorded by recent Line Oriented Safety Audits (LOSA) is within the category of intentional non-compliance of which checklists use is included. These errors have led to serious lapses in risk management and many well-documented cases of aircraft accidents. This paper reviews the literature of both organizational behavior management and applied behavior analysis where checklist use is an independent variable. This report presents various methods and technologies from other settings which may prove useful in the flight-training environment. Also included is a proposed study that will be conducted at a major flight training facility using undergraduate participants involved in checklist use while undergoing instrument flight training. This study applies various treatments to the participants to measure the effectiveness of checklist reading behavior and performance. Measures examine both, short term and long term effects of treatment, as well as any generalization of checklist reading performance to more advanced training environments.

Checking Checklist Performance

Upon reviewing the checklist literature it becomes apparent that much has been documented regarding checklist design, checklist importance (Degani and Weiner 1990, FAA, 1995), and accidents and incidents resulting from the misuse of checklists (NTSB, 1969, 1975, 1982, 1988a, 1988b, 1989, 1990, 1997). Yet there seems to be a lack of studies regarding effective technologies that will improve a crew’s use of checklists in flight operations. Using the Line Oriented Safety Audit (LOSA), Helmreich reported over 50% of all in flight errors were intentional non-compliance. Non-compliance errors were, "conscious violations of standard operating procedures (SOP) or regulations. Examples include omitting required briefings or checklists" (Helmreich 2000).

Checklist performance may vary widely between operators. (Diez, Boehm-Davis, and Holt, 2003). Two methods of checklist design are the challenge-do-verify (CDV) and the do-verify (DV). These methods can be paired with text/paper, mechanical, or electronic aids to ensure the crew is not relying only on memory (FAA 1995). Degani and Weiner (1990) identified similar checklist methods using challenge and response, memory-guided, and short-cutting or "chunking" the checklists. Chunking comprised calling a limited number of challenge items from the checklist, then checking those tasks by reading from the checklist.

Examining the checklist accident data, evidence would suggest that proper checklist use is vital in performing sequential tasks at the appropriate time in a potentially highly distractive environment. One question should be: How can we train and bring checklist behavior to a level of consistency that SOPs require? Another question might ask; Is there other variables within the training environment which may be paired with checklist reading behavior which might increase or maintain the frequency of checklist use?

Diez, et al (2003) identified significant differences in compliance with crews of one airline using a memory-based checklists. Verbal annunciation of checklist items was a required SOP for the airline of interest. Of the expected 100% compliance, the crew only vocalized approximately 78% of the checklist items. These differences were observed across aircraft and between checklists. Searching for checklist studies, such as the one above, where the checklist is the independent variable in aviation literature, is limited.

Experiments in other disciplines, which use checklists as a component of the independent variable, may provide important clues relating to any changes in performance resulting from manipulations of those variables. Through experimental design, performance in checklist use can be measured. Therefore the resulting intervention, and subsequent change in performance, may provide application in a flight training setting.

Literature Review

The purpose of current review is to (a) examine the literature for checklist use in organizational performance change, (b) determine those studies where checklist use was most effective, (c) discuss effectiveness of interventions where checklists were paired with other variables, and finally recommendations are made for future research in the area of checklist use in flight training.
Method

Articles were reviewed from the *Journal of Organizational Behavior Management* between the years 1977 and 2003 and the *Journal for Applied Behavior Analysis* between the years 1968 and 2003. Particular keywords, listed below, were used to search in the PsycInfo 1887 database for relevant articles.

Inclusion and Exclusion Criteria

Inclusion criteria were the keywords, (a) checklist, (b) job aid, (c) task analysis, (d) task clarification, (e) prompt, (f) task performance, (g) self-monitoring, and, (h) task list. Both journal titles were used concurrently with each keyword to generate search results.

Only studies in organizational contexts were used with the intervention being applied to real-world tasks. If it could not be determined that the participants were employed in the target organizational environment and given some form of a written checklist to use during performance of the task, the article was excluded. Articles in which target participants were mentally impaired or in school environments were excluded. This provided inclusion of studies where responses were prompted, which resulted in, a product, service or some other measurable accomplishment.

Under these criteria nine articles were discovered for review. One article of the nine contained three separate studies, which increased the total review to eleven studies of checklist use as the independent variable. Three other articles, from other journal sources were discovered, due to their contribution to the subject of checklist use, they will be discussed. These articles are not included in the final review and will only provide supporting discussion.

Interobserver Agreement

A trained rater independently coded each article in the review. Inter-observer agreement was calculated using this formula: (Disagreements/Agreements + Disagreements) x 100. Reliability was 100%.

Results

The literature review revealed multiple settings and tasks where checklists were used as a stand alone component or in combination with other independent variables. Two articles used a manufacturing environment as a setting. The first, Moses, Stahelski, and Knapp, 2000, used control charts and a check sheet process as a strategy to control reducing the size of metal tubes. The second study used posted written set-up procedures to reduce set-up times on a die extruder machine, Wikoff, Rowan, and Poling, 1990. Four studies involved the hotel industry. Three of the four studies were contained in one article by Anderson, Crowell, Sponsel, Clarke, and Brence, 1983. Checklists were used with various cleaning, room preparing, and support tasks. The forth article by LaFleur and Hyten, 1995 used checklist for preparing banquet rooms for hotel customers. The only study to use checklists in an office setting was by, Bacon, Fulton, and Malott, 1982. This study examined the tasks of record keeping, grading, lesson completion, and system maintenance. One study by Porterfield, Evans, and Blunden, 1985, demonstrated checklist use in the form of a distributed leaflet, to improve performance of health care workers with developely disabled individuals. Another interesting study used a student managed bar to conduct an intervention of checklist posting and feedback to improve cleaning tasks, Anderson, Crowell, Hantula, and Siroky, 1988. For a down home approach, Altus, Welsh, and Miller, 1991, used checklists in a student housing cooperative to improve task performance in a domestic setting. Customer service tasks were improved by including a posted task list for bank tellers in a study by Crowell, Anderson, Abel, and Sergio, 1988. For a complete comparison of the review studies, Table 1 summarizes the findings of the review.

Review Analysis

Checklist medium. All studies used a written checklist format. Some examples were provided within the studies to compare the compositions (Anderson et al. 1988; LaFleur et al. 1995; Moses et al. 2000). Eighty two percent of the studies described the checklist as some type of written document that could be posted or carried by the participant. As previously mentioned in the Crowell, Anderson, Abel, and Sergio (1988) study, the checklist could be as simple as a memo listing behavior prompts or in the Porterfield, Evans, and Bluden (1977) article which described the staff roles as a leaflet to be used at the discretion of the participant. Other studies demanded more involvement of the participants with the checklist, such as office task work by Bacon, Fulton, and Malott, (1982), which required participants to mark items on the checklist as the task progressed.

Checklist pairings. All the studies in this review, except one, used checklists as one component part of an intervention. Only the Bacon, Fulton, and Malott...
(1982) study used the checklist alone as an independent variable. The remaining ten studies paired the checklist with various forms of feedback, tokens, goal setting, or a punisher in the form of a fine. With all studies, the dependent variable measurably changed in the desired direction. However, one drawback is that a component analysis was not conducted in any of the studies with checklist pairing. Therefore it is difficult to conclude that the results produced by the interventions were due to paired checklist use or checklist use alone. It is assumed observed effects are from the pairing of checklists in combination with various forms of feedback. Many times it was difficult to determine from the study how consistently the checklist was used and therefore was it or was it not consistently paired during the intervention.

One manufacturing study may have paired a checklist with an attribute control chart for feedback (Moses et al. 2000). This particular chart tracks specific control limit events using statistical process control. It is not clear from the study if the participants who used the checklist to sequence the inspection process actually viewed the attribute control charts. This leaves a potential confound in the study with regard to whether pairing intervention components actually occurred.

The second manufacturing study paired the checklist of set up procedures with two feedback methods, observation audits and video feedback (Wikoff et al. 1990). Prior to the intervention, the experimenters conducted a task analysis and listed sequential steps in the set-up procedure, which would yield optimum performance. Copies of the written set-up procedures were given to each participant and one copy was posted on each machine. It is not conclusive from this study that the participants actually used the set-up procedures checklist each time they set up their machine. At least once each week, for four months, each participant was video taped and feedback provided regarding performance of set up times. After the four-month time period a trained supervisor conducted an operational audit at least once each week, for three months. Verbal feedback was provided by supervisor regarding the participant’s performance. Set up time did decrease during the intervention yet without controlling for consistent checklist use or conducting a component analysis, it is difficult to judge the effectiveness of the checklist or with the checklist in either feedback combination.

The six service studies used a variety of feedback methods paired with checklists. Three of those studies conducted by Anderson, Crowell, Sponsel, Clarke, and Brence (1983) used room cleaning, housemen, and doormen checklists paired with weekly posted charts of completed checklist items. This pairing was followed by a period of several weeks into the intervention with the same checklists paired with the same-posted chart and adding tokens awarded for criterion or better performances. Again, it is difficult to tease apart the effectiveness of each component in these interventions. There was a desirable directional change in many of the performances, however no definitive conclusions can be drawn as to which individual component or combinations of components may yield similar results.

The study by Porterfield, Evans, Blunden, 1985 investigated health care workers. During this four-phase experiment, the experimenters added the leaflet checklist and the daily observable feedback to the participants during the same phase. This procedure missed an opportunity to measure each component separately within the study. From the pairing of both the leaflet and the vocal feedback, it is evident that desirable behavior increased as illustrated in the study results.

The study of banquet set-up tasks was conducted by LaFleur and Hyten 1995. During this study task checklists were explained in detail to the staff. This study also used the checklist as a response sheet to record when each task was complete by signing their name beside each completed task. The checklists were later collected by the participants’ supervisor. Checklist use was simultaneously paired with publicly displayed, daily, setup completion percentages graphs, goal setting, and monetary bonuses. Results indicate that setup completion percentages increased from between 40 percent variable to 100 percent to nearly 100 percent consistently. This provides supporting evidence to the effectiveness of the total treatment package. However conclusions can only be made that the checklist usage in food setup tasks may be effective when used with some type of supporting reinforcing intervention.

The student operated bar paired task clarification with posted checklists in an experimental design that isolated this portion of the intervention from the feedback portion (Anderson et al. 1988). Feedback was provided in the form of publicly displayed line graphs. This study does show evidence that antecedent prompting of task clarification and posted checklists can make an immediate change in desired behaviors. The study states a sample of an unscored checklist was posted for continuous viewing. The study indicated that during the portion of antecedent-like treatment, behaviors increase modestly. Behaviors increased again after the line graph feedback treatment was added to the intervention.
This study provides some evidence of behavior change prior to the addition of feedback. However, the treatment protocol is vague regarding the consistent use or viewing of the checklist by the participants during the first phase of the intervention. Yet the evidence suggests that some type of effect occurred as a result using paired task clarification and posted checklists.

A similar intervention was used in the bank teller study by Crowell et al. (1988). In this study task clarification was explained followed by a “clarification” memo given to all participants explaining the behavioral categories and description of the point system for scoring performance. The study also paired the task clarification with a graph of mean transaction quality points, verbal feedback, and praise. The authors report, “performance change produced by clarification emerged quickly and remained relatively consistent throughout the phase” (p. 69). The study also reported, “present effects of task clarification are noteworthy because they are consistent with prior evidence showing that knowledge of task relevant behaviors can facilitate work performance, even in the absence of explicit feedback” (p. 70). This statement suggests that the participants gained knowledge of the tasks to consistently perform to a level higher than baseline. The duration of the task clarification phase intervention was 35 days. During this time the task knowledge was either drawn from memory of the initial task clarification briefing or from frequent review of the task clarification memo. This study does illustrate changes in desired behaviors during each phase of the experiment. Yet the study is unclear with regard to the experimental protocol concerning use of checklists to prompt the complex tasks of social interaction of bank tellers.

The Altus et al. (1991) study examines mean percentage of task completion of household duties by following a written checklist of those tasks. The checklist was paired with tokens for adequately completed tasks as well as fines for a specific level of uncompleted tasks. This is the only study in the review, which used an obvious punisher. Fines were assessed for behavior other than on task, paired with a checklist during the intervention. Behavior changed in a desirable direction using the paired intervention of written checklists and tokens and fines. The researchers concluded that the participants managed the tasks very effectively after the introduction of the checklists. The study suggests the resulting increase in task completion and the decrease in fines and complaints support the notion that the pairing of written checklists with rewards and punishers may have merit. However the study did not go into detail regarding potential confounds of using a punishment technique in the study. Counter control issues related to punishment or emotional bursts were never discussed.

The study, which isolated the checklist as an intervention, was the Bacon, Fulton, and Malott (1982) study. However this study did use three specific conditions from a study conducted by Brethower (1970), which required participants to understand and interact with the checklists. Except for the LaFleur and Hyten (1995) study, it was not evident that any other studies used all of these three elements as did Bacon et al. 1982. The first required element was the need to view the checklist daily (i.e., this assumes the required frequency of the task is daily). In doing so, the checklist requires evidence of completing the tasks. The second required element is recording, on the checklist, a specific time of task completion or amount of work that has been completed. The last requirement, to maintain checklist usage, is periodic review of the checklist by the participant’s supervisor. It was clear from the study that the participants would not receive any punishing consequences resulting from an incomplete checklist. Each week the participants would receive a new checklist without any mention from the researcher of the previous checklist. Therefore no feedback was given to the participants. The results of the study seem to demonstrate a desired change in task completed behaviors by using the checklists alone. The study also discussed the potential for reinforcing effects from using the checklist alone. Some examples of contingencies for continued checklist use might be the result of rule-governed behavior. The potential punishing consequences of the supervisor reviewing an incomplete checklist may support continued checklist use. The viewing of the checklist could also provide a reinforcing effect for actually working on the tasks by seeing the task checked off the list.

As with most of the studies reviewed, the pairing of checklist use was done without a component analysis to determine if on task behavior would result from checklist use alone. Perhaps many of the effects demonstrated in these studies are under the influence of multiple contingencies that are yet to be isolated and control of those contingencies demonstrated.

Discussion

The review articles claim that checklists have been used in capacities to prompt specific behavioral performance. Many of these studies use
interventions with multiple components, which are paired with checklists, to produce behavior change. These paired components ranged from task clarification of the environment and task knowledge, to incentives, graphic feedback, and punishment. Gilbert (1978) supported the notion that for workers to perform well they require environmental information, knowledge, and response supports. As one of the intervention components, checklists may provide response supports, which increase the efficacy of complex task completion.

Most checklists within the review articles were developed from the need to document specific behavioral response chains. This effort requires a detailed job task analysis of the desired knowledge, skills, and abilities of the specific task (Gatewood, and Field, 1998). One approach to developing a checklist is to conduct a thorough job task analysis, which should identify the present skills, and abilities that demonstrate a particular level of performance. Using information from the job task analysis, a checklist can be constructed that will maximize performance sequences and provide the worker task clarification (Anderson et al., 1988, Degani and Weiner, 1990).

Checklists can be used to prompt specific behavior to occur. Prompting of behaviors As such, training, job aids, task clarifications, and checklists can be antecedents for behavior. While many of the studies demonstrated behavioral changes due to the use of checklists, only the Bacon et al. study in the literature review used checklists alone to effect behavior change. All other studies eventually linked checklists with some form of reinforcer or punisher as a total intervention package. This leaves to question the behavioral function of the checklist and what contingencies may support continued checklist use.

One article not included in this review is by Shier, L., Rae, C., and Austin, J. (2003) which used five checklists to improve performance in a grocery store environment. The checklist were signed and returned to the researchers who then graphed the results and used that data for public posting. Again this intervention used task clarification, checklists, and feedback to demonstrate performance change. The most unique element of this study was the authors attempt to conduct an organizational functional assessment (Austin, Carr, and Agnew, 1999) to determine cause of the poor performance.

Due to the lack of empirical research on specific, stand alone, checklist systems, the exact behavioral functions of checklist use remain for future research.

The issue of checklist use and distraction, in a particular setting, is not addressed to any extent in most of the studies. No study addressed a treatment protocol for checklist handling. Assumptions were made that participants would use the checklist as needed perhaps depending on the strength of other contingencies for not using the checklist (LaFleur et al. 1995, Bacon et al. 1982, Moses et al. 2000, Anderson et al. 1988). No study examined the consistency of the environment where the checklist was used. Future research should examine checklist use in changing environments where the potential for error in checklist reading could be high and the penalty for error would be costly such as flight training.

Task clarification and checklists seem to produce a rapid change in behavior immediately and consistently after introduction (Bacon et al. 1982 and Crowell et al. 1988). As an effective, inexpensive method of improving and maintaining performance why is there not more dedicated research focused on the use of checklists alone? Only the Bacon et al. 1982 study tried to examine the use of checklists without pairing them with other interventions. Future research should investigate refinements of checklist construction. Studies could examine how the checklist interacts with the users. There should be a validation process to determine if some type of supervisory contingency must be created to maintain checklist use or can an interactive contingency be created with the checklist and the user alone.

**Conclusion**

This review compared the use of checklists in the literature to determine areas of commonality in the field of applied organizational studies. It is evident that service tasks have been the focus of many researchers. These tasks and settings present potential for checklist interventions by the nature of the somewhat transient populations, repetitive nature of the tasks, pay scale of the jobs, and the level of detail required completing the tasks. All of the studies were in applied setting which present unique challenges regarding experimental control. However each study did demonstrate some type of behavioral change through the use of an intervention. Unfortunately most of the interventions used some type of checklist pairing procedure without examining checklist effects independently prior to pairing with another component.
It is very likely that using similar checklist pairing procedures may result in performance changes in flight training performance. Checklist strategies have long been paired with other interventions. Perhaps it is time to examine in finer detail the changing technologies in which checklist use can be created, monitored, and refined. Generalizing from the performance successes reported in the literature review, future checklist research in flight training will isolate and test the antecedent, consequence, and motivating operations associated with checklist use in stable, threatening, and changing environments. Dependent variables will consist of observable behaviors in checklist reading i.e. fluency, frequency, latency, ratios of items performed per segment and ratios of items performed correctly over time. The manipulation of the independent checklist variable will consist of pairing checklist use with and without graphic feedback of the dependent variables. The research protocol will use PC-based flight trails while conducting an instrument approach. It would be of great benefit to confirm the reliability and validity of checklist pairings with various settings and tasks.

Using checklists has a long history in many settings and for many tasks. It is time to look closer at the checklist and determine if there isn’t more to using the checklist than already exists in the literature.

References


Washington, DC: National Transportation Safety Board.


