INL/CON-13-28524
PREPRINT

Computer–Based Procedures for Nuclear Power Plant Field Workers: Preliminary Results from Two Evaluation Studies

2013 International Annual Meeting of the Human Factors and Ergonomics Society

Katya L. Le Blanc
Johanna H. Oxstrand

October 2013

The INL is a
U.S. Department of Energy
National Laboratory
operated by
Battelle Energy Alliance

Computer–Based Procedures for Nuclear Power Plant Field Workers: Preliminary Results from Two Evaluation Studies

Katya L. Le Blanc and Johanna H. Oxstrand
Idaho National Laboratory

The Idaho National Laboratory and participants from the U.S. nuclear industry are collaborating on a research effort aimed to augment the existing guidance on computer-based procedure (CBP) design with specific guidance on how to design CBP user interfaces such that they support procedure execution in ways that exceed the capabilities of paper-based procedures (PBPs) without introducing new errors. Researchers are employing an iterative process where the human factors issues and interface design principles related to CBP usage are systematically addressed and evaluated in realistic settings. This paper describes the process of developing a CBP prototype and the two studies conducted to evaluate the prototype. The results indicate that CBPs may improve performance by reducing errors, but may also increase the time it takes to complete procedural tasks.

INTRODUCTION

The majority of work activities conducted by personnel in a nuclear power plant (NPP) are guided by procedures, which are typically presented on paper. While the current paper-based process has a demonstrated history of ensuring safe operation at NPPs, researchers argue that leveraging the capabilities of digital technology in the procedure process by transferring the existing paper-based procedures (PBPs) to computer-based procedures (CBPs) may further enhance safety (Fink et al., 2009; Le Blanc & Oxstrand, 2012). CBPs are recognized as a way to potentially boost safety and efficiency in NPP operations by assisting the operator with place keeping, evaluating step logic, gathering process information and a variety of other tasks (Fink et al., 2009).

Though the concept of CBPs has been around for over 30 years, the United States (US) nuclear power industry has not yet adopted a CBP process. Potential reasons for this include lack of operating experience with CBPs, concerns about regulatory acceptance, and the cost associated with migrating to CBPs.

Most research has focused on implementing CBPs in the control room; however using CBPs in the field may prove more fruitful as a first step. Field workers typically conduct their work in the plant, and the ability to have access to work instructions on a single handheld device paired with the ability to communicate in real-time with supervisors has potential to greatly enhance efficiency. Given the significant investment associated with migrating to CBPs and the uncertainty associated with regulatory acceptance of CBPs, utilities want to be confident that they will see performance advantages with CBPs before they pursue conversion of their PBPs to CBPs.

With the goal of defining a path to transition from PBPs to CBPs, researchers at Idaho national Laboratory, in close collaboration with the US nuclear industry, are working to develop and demonstrate CBP concepts. The primary purpose of the research is to define how to design the user interaction with CBPs to encompass all aspects of the technology procedure process while optimizing the benefits of digital without introducing new errors.

Researchers have designed and evaluated two successive CBP prototypes to identify desirable interactions with CBPs. The evaluation studies were carried out at two different NPPs. The CBP prototypes were designed based on requirements identified in a previous study (Le Blanc & Oxstrand 2012). The main features of the CBP prototype are:

- Paper-based procedures are designed to cover a wide variety of tasks and conditions; therefore much of the information is irrelevant for a particular situation. The CBP prototype is designed to only present procedure paths and instructions that are relevant to the current task, plant mode, and equipment status (referred to as context sensitivity).
- The CBP simplifies step logic where possible. Rather than presenting operators with complex conditional statements, the CBPs prototype prompts the operator to input the condition(s) into the procedure system, and presents the instructions based on that condition (or set of conditions).
- The CBP is designed to make important information stand out based on the current situation. This included automatic place keeping, providing dynamic cues when conditions need to be monitored, and providing warnings when conditions are violated.
- The CBP automatically performs calculations and verifies input.

EVALUATION STUDY 1

To evaluate performance using the CBP, researchers compared performance on a realistic procedural task using the CBP prototype and the traditional PBP.

Method

Participants. The participants in the evaluation study included 13 technicians at an operating NPP. The technicians came from varied disciplines within the plant including two electricians, two mechanics, three I&C technicians, one chemistry technician, two procedure writers, one IT expert and two others. All of the participants were male. The average age of participants was 48 years (*SD* = 12.95 years).

Procedure Selection. Researchers worked with plant personnel to identify a procedure that would ensure that the functionality of the prototype could be demonstrated during the evaluation scenario. The procedure selected was a plant procedure for “racking out a breaker” adapted for use in the plant’s electrical training lab. The use of a training lab in the evaluation study allowed for a realistic setting, in which technicians could take actions on the equipment without affecting the plant.

Researchers prepared a paper-based version of the procedure that conformed to the plant’s procedure-writing guide, and a computer-based version using the prototype CBP software. Based on previous research (LeBlanc & Oxstrand 2012), the researchers determined that field operators need a device small enough to put in their pockets (due to the fact that they frequently climb ladders and work in cramped spaces).Therefore, the computer-based version was presented on an Apple iPod touch.

Surveys. To assess the workload associated with using the CBP prototype compared with the traditional PBPs, researchers used the NASA Task Load Index (TLX) (Hart, 2006). The NASA TLX was administered using paper and pencil. Researchers developed their own usability survey to assess the interface of the CBP prototype. The 8-item survey targeted the availability of information, ease of navigation, and ease of use of the CBP interface. Researchers also developed a 6-item usability survey to assess the usability of the device. Researchers also developed a debrief questionnaire to gain more detailed feedback on the design of the user interface and the overall experience using the CBP. The questions on this survey were open-ended.

Design. The researchers used a 2-factor within-subjects design. The independent variable was procedure presentation type (i.e., CBP or PBP). Participants were assigned to either PBP-first or CBP-first order. The order was counterbalanced across participants.

Experimental Protocol. When participants arrived, they filled out an informed consent form. They then completed a pre-job brief that included a review of the procedure, a discussion of the conditions that would be encountered in the scenario, as well as a discussion of the potential safety issues associated with the scenario. This pre-job brief served as the pre-job brief for both conditions (PBP and CBP) and was executed with the PBP. If the assigned order was PBP-first, the participant completed a two-minute drill using the paper-based procedure; the two-minute drill occurred at the job-site and included a brief overview of the expected initial conditions and the potential safety hazards Participants were instructed to complete the procedure scenario to the best of their ability and at their own pace. A researcher and a qualified electrician observed each scenario. Researchers were trained to recognize deviations from the optimal procedure path. They followed the scenario closely and recorded any deviations. Additionally, the qualified electrician observed the scenario and was instructed to note any deviations and share them with the researchers after each participant completed the scenario. The researcher started a stopwatch at the initiation of the first step and stopped the stopwatch once the final step was completed. When the scenario was complete, the participant was given the first NASA TLX and then he completed the scenario with the CBP. Before the CBP scenario, the participant was given a 5-minute training session on how to use the CBP interface. The training provided was minimal, but was expected to be sufficient to allow the participants execute the procedure. The rest of the scenario occurred exactly as the PBP scenario, except that the procedure was executed using the CBP prototype. Once the scenario was completed, the participant was given another NASA TLX form and the usability survey. If the participant was assigned to the CBP-first scenario, the sequence of events was the same except that the participant completed the drill using the computer-based procedure and executed the scenarios in reverse order.
Once both scenarios were complete, the participant was given the debrief questionnaire.

Results and Discussion

Deviation from optimal procedure path. Researchers recorded no deviations in the optimal procedure path across all participants and conditions. Because there were no verified deviations in the study, a comparison of performance between the paper and computer-based versions of the procedures was not possible. The lack of deviations is likely due to the fact that the procedure and the scenario were relatively simple to execute.

Completion time. The completion time was defined as the time elapsed between the start of the first step of the procedure and the successful execution of the last step in the procedure. Researchers used a stopwatch to measure the time in seconds. The completion time data was subjected to a 2 X 2 (Order X presentation style) mixed analysis of variance. The within-subjects factor was procedure style (CBP or PBP) and the between-subjects factor was order (PBP-first or CBP-first). The results indicated that it took longer to execute the scenario with the CBP (M =447, SD = 103) than with the PBP (M = 332, SD = 111), the main effect of presentation style was significant $F(1, 11) = 37.595$, $p < .001$. There was a significant interaction between procedure style and order $F(1,11) = 14.087$, $p < .01$; if the CBP scenario was completed first, then the difference between CBP and PBP was greater than if the PBP was completed first. Essentially, the effect of practicing the actual procedure had a greater effect on the CBP than the PBP. There was not a significant main effect of order.

There are several possible explanations for why it took longer to execute the scenario with the CBP than with the PBP. The first is that the participants were using the CBP for the very first time with minimal training, while most the participants were used to using the PBPs on a daily basis. Another potential explanation is that the participants would often stop and comment on aspects of the CBP during the scenario execution even though they were instructed not to, increasing the overall completion time of the scenario.

Subjective Workload. Subjective workload, as measured by the NASA TLX scores, was compared between the scenario execution with the CBP and scenario execution with the PBP. The raw TLX (Hart, 2006) score was computed for each participant and compared between procedure styles. A paired samples t-test revealed that there was not a significant difference between subjective workload scores for the CBP (M = 2.23, SD = 2.04) and the PBP (M = 1.73, SD = 1.63).

The workload scores were relatively low across conditions, indicating that participants found the scenario to be relatively easy to execute. A more difficult task may have yielded larger differences in workload. Importantly, the workload was similar for both the procedure execution with the CBP and the PBP. This indicates that managing the CBP interface did not add significant workload for the participants. In a task in which the overall workload is higher, it might be possible to detect an advantage in workload for CBPs.

Usability Survey. The usability survey was designed to assess the overall usability of CBP interface by inquiring about the ease of navigation, the availability of information, and other common usability dimensions. The overall usability score was computed by averaging the scores across all of the questions. Participants rated the CBP interface as moderately usable; the mean overall usability score was 3.5 on a 6-point scale. The lowest scores were reported for items related to the navigation of the user interface.

Debrief Questionnaire. Before scoring the debrief questionnaire, researchers developed a coding scheme for the open-ended responses. Any responses that did not fit into the *a priori* coding scheme were marked as "other" during coding. The results are presented on a question-by-question basis below.

Procedure style preference. More participants preferred the CBP to the PBP. Seven participants indicated that they preferred the CBP; two said they preferred the PBP, and four indicated that they had no preference. The most common reason for favoring the PBP (or having no preference) was that the CBP provided no way of looking ahead. This is an interesting result, because the CBP prototype had multiple ways of looking ahead, including an overview and the capability to preview the next steps. These features were pointed out during training. This indicates that either the participants were not trained extensively enough on these features, that they did not remember the look-ahead functions, or that they did not find them useful. Other reasons for preferring the PBP (or having no preference) were the fact that the procedure pages took too long to load, and that the iPod touch was too small. The most common reason for preferring the CBP was that it reduced the opportunity for errors compared with PBP.

Context-sensitivity. All participants reported that only seeing the relevant steps in the CBP was an improvement over the paper-based procedure. This result indicates that participants did not find it confusing or disorienting to have portions of the procedure "hidden" because they were irrelevant to the current task. Additionally, participants unanimously preferred the simplified step logic of the conditional statements. This indicates that context-sensitivity and simplification of step logic are highly desirable features of CBPs.

In addition to the results presented here, the researchers elicited specific feedback on many of the features of the user interface. They incorporated that feedback into a revised prototype evaluated in a second study.

EVALUATION STUDY 2

The researchers modified the CBP prototype based on the results and feedback from evaluation study 1 and conducted a second study at a different operating NPP.

Method

Participants. Ten employees at an operating nuclear power plant (NPP) participated in this study. Nine were Nuclear Equipment Operators (NEO) and one was a training manager. The mean age of the participants was 40 years (*SD* = 11 years). The participants had an average of 10 years of experience (*SD* = 8 years) in their current role. All participants were male.

Procedure Selection. Results from evaluation study 1 indicated that it may be difficult to detect performance advantages with CBPs if the procedure and scenario are too simple. Therefore, the researchers worked with plant personnel to select a more complex procedure and scenario for this evaluation study. The procedure utilized was an existing procedure used to train field operators in a functioning flow loop training facility; the specific scenario involved initiating the cold water system and the control loop system. The features that contributed to the increased complexity in the scenario chosen for this study include the presence of multiple conditional statements and branching to other procedures. These features of the procedure and scenario allowed the research team to better assess the impact of context sensitivity in the CBP prototype. Identical versions of the procedure were prepared for paper and the CBP prototype (which was presented on an Apple iPod touch).

Surveys. Researchers used the same surveys described in evaluation study 1 with a one exception. The debrief survey was modified to reflect changes made to the CBP prototype.

Design. This evaluation study was designed to determine whether the CBP prototype offers performance advantages over the PBP. Therefore, the main factor in this study was the procedure presentation style (PBP or CBP) and it was manipulated within participants. The main dependent variables in the study were completion time and deviations in the optimal procedure path. In order to investigate the CBP effect on performance time, researchers measured the completion time of each scenario via a stopwatch. The timer was started when the participant indicated that they were entering the first section of the procedure, and stopped when the participant completed the final step in the procedure.

The researchers also recorded the number of deviations from the optimal procedure path. Researchers worked with plant personnel to identify possible deviations prior to conducting the study. Deviations were broken down into two categories: recovered deviations and non-recovered deviations. Recovered deviations were defined as situations in which operators were not following the optimal path, but ultimately recovered. For example, when an operator walked to the wrong location or attempted to verify the wrong component (but ultimately found the right component or location), it was recorded as a recovered deviation. Non-recovered deviations were defined as deviations in which the operators failed to take an action specified in the procedure or took the wrong action. The researchers chose to classify the deviations separately because they wanted to capture deviations that indicate confusion or misunderstanding (especially with respect to the CBP prototype, but that would not typically be considered deviations in the procedure). The non-recovered deviations would have a greater impact on system performance and safety than the recovered ones.

Experimental Protocol. Participation in the evaluation study was conducted in two sessions. During the first session, the participants were familiarized with the task and trained on how to use the CBP prototype. During the second session, participants executed the scenario with the CBP and with the PBP. Upon arrival at the first session, participants signed an informed consent form. Participants were then randomly assigned to complete the scenario with the CBP first or with the PBP first (order was counterbalanced). Participants then filled out a brief demographics survey. Following the completion of the demographics survey, participants were instructed on how to use the CBP. The researchers trained the participants to navigate through the CBP, and then completed a simulated walk-through of the procedure to demonstrate how to use features such as barcode scanning. The training took approximately 30 minutes.

The participants returned the following day to execute the procedure using both the PBP and the CBP. Participants were given a pre-job brief which included a discussion of the task to be completed, the potential hazards, and other important safety concerns. If the participant was assigned to the PBP-first condition, then he completed the procedure with the PBP. When the participant indicated that he was starting the first section of the procedure (i.e., reviewing the prerequisites), the researcher started the stopwatch. The researcher watched the scenario closely and recorded any deviations. Once the final step was completed, the researchers stopped the stopwatch and recorded the time. The participant then filled out the NASA-TLX for the PBP condition. The participant then completed the same scenario using the CBP in the same manner. Following the execution of the scenario with the CBP, the participant completed the NASA-TLX for the CBP task, the CBP usability survey, and the debrief survey. If the participant was assigned to execute the task with the CBP first, the experiment proceeded in the same manner except they executed the scenario with the CBP before the PBP.
The researcher had to make one potentially important modification to the above protocol because equipment in the flow loop facility was malfunctioning. When the first participant executed a portion of the scenario, it was discovered that an air operated valve was malfunctioning. Maintenance personnel were not able to repair the valve during the visit, so the researchers instructed the participants to simulate the portion of the procedure that relied on that valve. For that portion of the procedure, the participants ran through the procedure as though they were conducting it, but did not actually manipulate the equipment as they did for the remainder of the procedure. This modification was the same for both the CBP and PBP conditions.

Results and Discussion

Scenario completion time. The time to complete the scenario was measured in seconds using a stopwatch. A 2 X 2 (order by presentation style) mixed analysis of variance indicated that there was a significant main effect of presentation style (CBP or PBP) on the scenario completion time $F(1,7) = 7.12$, $p = .032$. It took longer to complete the scenario with the CBP (M = 1724 seconds, SD = 358) than the PBP (M = 1231 seconds, SD = 274). There was no significant main effect of order and there were no significant interactions with order. The finding that it took longer to execute the scenario with the CBP than the PBP is consistent with study 1. Though the researchers provided more training in study 2, the difference in completion times may still be related to learning to use the new interface. The operators will need more time to become familiar with the CBP interface in order to eliminate familiarity as an explanation for differences in completion time.

Recorded Deviations. Recorded deviations were reviewed by the research team and a trainer from the plant. Any error that did not fit straightforwardly into the predetermined coding scheme was discussed until the team came to a consensus as to how (and if) the error should be coded. Five recorded errors were eliminated by this process (four in the PBP condition and one in the CBP condition). Analysis of the recorded deviations was conducted using non-parametric statistical tests because the data violated the assumption of normality. A Friedman test revealed a marginally significant effect of presentation style based on type of deviation ($\chi_f^2 = 7.73$, df = 3, $p = .052$). There were more non-recovered errors with the PBP than the CBP. These results indicate that the CBP may be effective in preventing non-recovered errors. The CBP helped operators catch potential errors (i.e., recovered errors such as scanning the wrong barcode), and correct them before them became unrecoverable.

Usability. Participants generally found both the device and the interface to be usable. The mean usability rating for the device was 3.9 (SD = .68) on a six-point scale. One of the specific factors that concerned operators related to the iPod was the fact that it was likely not rugged enough for work in some areas of the plant (e.g., radiation areas). The mean usability rating of the CBP interface was 4.7 (SD = .69) on a six-point scale, indicating that the operators felt the interface was usable.

*Workload Scores**.*** For the workload scores, the simple averages (or raw TLX scores) were computed for each condition. There were no significant differences in subjective workload between the PBP condition (M = 2.15, SD = 1), and the CBP condition (M = 1.67, SD = 1). This indicates that the CBP does not increase or decrease workload compared with the PBP.

*Open-Ended Questions on the Debrief Questionnair**e.*** Before scoring the debrief questionnaire, researchers developed a coding scheme for the open-ended responses. Any responses that did not fit into the *a priori* coding scheme were marked as “other” during coding. The results are presented on a question-by-question basis below.

Procedure Style preference. Eighty percent of the participants preferred the CBP over the PBP. Most participants indicated that they preferred the CBP because it eliminated irrelevant steps and information and because it provided a more reliable means of correct component verification through barcode scanning. The participants that indicated they preferred the PBP, cited familiarity with the PBP process as the reason for preferring the CBP. This indicates that given a usable interface, even those performers who are resistant to switching to CBPs may shift their preference as the using the CBP becomes more familiar. These results highlight the fact that context-sensitivity is one of the more desirable advancements that can be achieved with CBPs.

Context-Sensitivity. Participants unanimously preferred the context-sensitive CBP presentation compared to the static PBP presentation. Most operators indicated that only being presented the steps relevant to the current task and conditions greatly streamlined the process, and prevented them from spending time and effort evaluating which conditions were relevant while they were out in the field.

CONCLUSION

CBPs may be effective in enhancing human performance. Though the scenario used in study 1 was too simple to detect a performance advantage with CBPs, study 2 provided some evidence that CBPs may help prevent errors. In study 2, there were fewer overall deviations when the procedure was executed using the CBP than with the PBP. Though there were more recovered deviations with the CBPs, there were a greater number of non-recovered deviations when the procedure was executed using the PBP. These findings show that the CBP may help operators to catch potential errors and prevent them. The CBP required operators to scan the barcode on the equipment to verify that they were on the correct equipment before they took any action, and thus provided a salient cue if they were at the wrong equipment. Therefore, the potential deviations were easily recovered from with the CBP.

The most common non-recovered deviation in the scenario occurred when operators were using a type of controller that most of them were unfamiliar with. Though all the participants had received training on this controller, many of them had not used one since that training. The deviation occurred when the operators were required to verify a tank level at another location. Many operators mistakenly believed that the level indicated on the controller was the tank level they needed to verify. However, the controller was simply displaying the set-point, not the actual level. Therefore in many cases, the performers failed to verify the tank level. The CBP required the operator to scan the barcode on the tank, and the operator was warned that he was viewing the wrong component if he scanned the controller. Therefore the CBP prevented the deviation.

Though the CBP may be effective in preventing errors, it appears that it may be a tradeoff between reduction of errors and time to complete the task. In both studies it took longer to perform the task with the CBP than the PBP. It is possible that the slower completion times are simply a result of being unfamiliar with conducting procedures with CBPs. However, it may be the case that there is an enduring effect on completion times, even when operators are familiar with using the CBPs. In that event, utilities will have to weigh whether reduced errors are worth the extra time it takes to execute procedures with CBPs.

The results of both studies also indicate that operators are likely to readily accept CBPs. The majority of participants reported that they preferred the CBP over PBPs, they rated the CBP as highly usable and the unanimously preferred the dynamic context-sensitivity of CBPs to static PBPs.

The researchers plan to modify the prototype one more time and conduct a final evaluation study to verify the concepts related to operator interaction with CBPs.

Although this research provides some promising evidence that CBPs may enhance performance, there are several limitations that should be considered when interpreting the results. First, the limited availability of NPP field operators meant that the studies had to be conducted with small samples, and that we had a limited amount of time with the participants. This resulted in limited the statistical power for our analyses in addition to causing some methodological challenges (e.g., limited time for training, and the necessity to use participants who did not quite fit the desired roles). Additionally, the simplicity of the evaluation scenarios may not accurately reflect the complexity of real-world procedural activities. The consequences of that are two-fold. The first is that the small differences in the number of errors and completion times may not exist or be important in real-world settings. The second is that these simple scenarios may not demonstrate the full potential of CBPs to increase efficiency and reduce errors.

DISCLAIMER

REFERENCES

Fink, R., Killian, C., Hanes, L., & Naser, J. (2009). Guidelines for the design and implementation of computerized procedures. Nuclear News. VOL 52 (3), 85-88, 90.

Hart, S.G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later, *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting,* Santa Monica, CA: Human Factors& Ergonomics Society

Le Blanc, K., & Oxstrand, J. (2012). A Model of Operator Interaction with Field Procedures: Insights for Computer-Based Procedures. 56th Annual Meeting of the Human Factors and Ergonomics Society, Boston, Massachusetts.