

EXPERIMENTAL STUDY ON THE EFFECT OF A PROCEDURE UNDER UNEXPECTED SITUATIONS

Makoto Takahashi¹

Daisuke Karikawa¹

Genta Sawasato¹

Yoshitaka Hoshii¹

¹⁾ Tohoku University, Japan

Abstract

Preparing procedures for a wide range of situations is considered important and has been confirmed to be effective for realizing reliable operations and a higher level of safety in complex socio-technical systems. However, once a situation drifts outside the scope of prepared situations, the effectiveness of using a pre-defined procedure may degrade, and operators have to make decisions without relying on procedures. In such cases, stronger operators are dependent on procedures, and the more unreliable operators' decision may come from the possible lack of a deeper understanding of the objective systems. The present study experimentally confirms that a strong dependence on the procedure may deteriorate the performance under unexpected situations outside the range of procedures. A Smart Grid Simulation environment, in which dynamic decision making is required to deal with the given situations to avoid blackout, is adopted as the experimental testbed.

The scenarios for the expected situation presented herein are constructed on the basis of the events that have not been presented to the subjects in advance. The subjects are divided into two groups. The subjects in Group A are not given a written procedure that enables an effective operation for the predefined situations, whereas subjects in Group B are instructed to strictly follow the procedure. The results imply that the task performance evaluated on the basis of the number of causing a blackout degrades for the subjects following the procedure when facing unexpected situations.

Keywords: Procedure, unexpected situations, task performance, cognitive experiments

1. INTRODUCTION

In achieving a higher level of safety in huge complex socio-technical systems, such as nuclear power plants, aviation systems, and medical systems, among others, the operators' potentials to deal with severe situations, which are not supposed to happen, must be enhanced in advance [1]. The accident at the Fukushima Daiichi nuclear power plant owned and operated by the Tokyo Electric Power Company after the 2011 Great East Japan earthquake caused a huge and tragic impact on the people living not only in Fukushima but also in nearby prefectures [2]. The

accident is attributed to the huge tsunami, followed by the station blackout, which led to the loss of the total heat sink that was not totally unexpected and not virtually supposed to happen [3]. A new regulation based on the lessons learned from the disaster was issued. This regulation requires a nuclear power plant to be prepared to manage highly severe situations by improving not only hardware systems but also human performance based on revising procedures and training protocols [4,5].

Preparing procedures for a wide range of situations is considered important, and this has been confirmed effective for realizing reliable operations and a higher level of safety in complex socio-technical systems. Using procedures, excluding personal differences in the work to be done and contributing to the enhancement of the level of work efficiency and safety are expected. A procedure is effective in reducing human errors caused by the lack of knowledge. Troubles originating from flaws in procedures frequently happen. A procedure is generally considered as quite important in working environments.

Although using a procedure is considered quite important, depending too much on this procedure may have a negative effect on dealing with situations outside the predefined ones. Once a situation drifts outside the scope of prepared situations, the effectiveness of using a predefined procedure may degrade, and operators have to make decisions without relying on procedures. In such cases, stronger operators are dependent on procedures, whereas the more unreliable operators' decision may be caused by the possible lack of a deeper understanding of the objective systems. This situation is one example of the so-called Irony of Resilience [6].

A training process directing too much on following procedures may lead trainees to a state, called “fallacy of predetermination” [7]. This state is a biased expectation that the future unfolds as forecast and foreseen brought by the training process depending too much on the procedure, which may lead to the failure of considering unexpected situations. Although preparing procedures covering a wider range of possible situations may reduce the uncertainty in operations, it may also degrade human creativity, autonomy, and insights, leading to a degraded performance in dealing with unexpected situations.

Our group performed a cognitive experiment studying the difference of the operators' performances of dealing with expected and unexpected situations depending on the directions given by the two types of procedures [8]. One type of procedure provides information only on the chronological task elements to be done. The other provides background knowledge as to why. The results of the cognitive experiments using a realistic simulation environment showed that providing additional information had a positive effect on dealing with unexpected situations.

The present study focuses on the negative aspect of emphasizing the requirement of following the given procedures when facing situations that have not explicitly been mentioned in the procedures. The present study aims to experimentally examine whether or not the strong dependence on a procedure may deteriorate the operator's performance under unexpected situations outside the range of procedures.

2. METHOD

2.1. Definition of an “unexpected event”

The present study focuses on the human performance of dealing with “unexpected” situations. The definition of “unexpected situations” should be clarified herein because of a variety of

interpretations for the term “unexpected.” Unexpected events can be categorized into the three following categories:

- events that had not been imagined,
- events that had been imagined but excluded from consideration because of its extremely low probability and high uncertainty, and
- events that had been expected but optimism prevented from taking counteractions because the amount of required resources had become too huge to consider.

Although the socio-technical viewpoint of the abovementioned unexpectedness is important to consider, unexpected events should be defined in the cognitive experiments in which a student takes part as a subject. In this study, the unexpected events are defined from the viewpoint of the existence of past experiences and the level of difficulty. The three kinds of situation are as follows:

- Normal: normal operations following procedures,
- Expected abnormal: abnormal events experienced in the practice runs and can be handled just by following the given procedure, and
- Unexpected abnormal: abnormal events not experienced in the practice runs and cannot be handled just by following the given procedure.

The cognitive experiments are performed herein using complex task environments. The task performances in terms of dealing with expected and unexpected abnormal events are compared according to the level of dependency on the given procedures.

2.2. Experimental environment

A dynamic task simulating a Smart Grid (SG) system was adopted as the experimental environment, in which the subjects were required to control the system according to dynamically changing situations. The operation aimed to stabilize a grid voltage, which was subject to change by the balance between the supplied power and demand, using four power sources. The subjects were required to understand the characteristics of each power source and avoid a blackout caused by the excessive unbalance between supply and demand. Real-time operations following the given procedure were necessary to achieve the goal. The grid voltage goes up when supply exceeds demand. In contrast, the grid voltage goes down when demand exceeds supply. Figure 1 shows the main operation screen of the SG system.

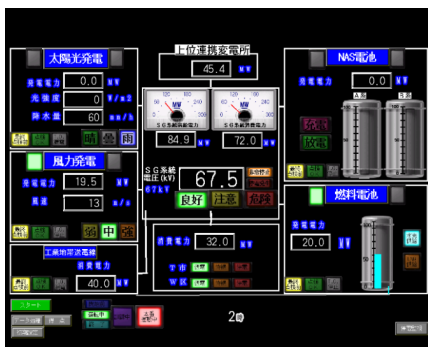


Figure 1 Main operation screen of the SG system



Figure 2 Operation window of wind power generation

The level of power supply and demand is shown in the middle with the SG grid voltage determined by the balance of supply and demand. The state of the SG grid voltage is indicated by the “Voltage quality indicator”: Good (green), Caution (yellow), and Danger (red).

3. The state of four power sources (i.e., wind power, solar power, Sodium Sulfur(NAS) battery, and fuel cell) is displayed surrounding the middle part of the display. An additional sub-window for controlling each power source opens by pointing the region of each power source. Figure 2 shows an example of the windows of the power source (wind power).

Each power source consists of two trains (lines A and B). The output power of the wind power generation is determined by the wind velocity, whereas that of the solar power generation is determined by the level of the luminous intensity of the sun. The NAS battery can be charged and discharged by operators. The fuel cell has a hydrogen tank and two supply lines. Hydrogen is automatically supplied to the fuel cell depending on the level of remaining hydrogen. It takes 5 s to start the wind power and 8 s to start the fuel cell. The solar power and the NAS battery can be immediately started.

Operators must perform a daily inspection for each power source and line. The inspection can be performed on the operation interface for each power source and line by pushing the inspection button after shutting down. Inspection was allowed during 0:00 and 18:00 in system time. If the inspections for all power sources cannot be completed until 18:00, the power source expires, and troubles may occur in the SG system. When the power source is out of order, operators are notified when an inspection is done. Operators are also notified of a possible failure in the coming 24 h through the “Sign of Trouble” indicator.

As described earlier, the task environment adopted herein can simulate realistic and dynamic situations requiring real-time decision making and operations considering the different characteristics of each power source.

3.1. Experimental task

The operators were directed to perform two tasks. One was a grid voltage stabilization task, and the other was an inspection task. The SG grid voltage should be controlled at approximately 67 [kV]. The situation escalates to a “Danger” when it deviates from more than ± 2 [kV]. The operators must control each power source to avoid the “Danger” situation from continuing. The SG grid voltage was determined on the basis of the amount of power supply and demand. As the demand changes according to daily consumption patterns, the operators must control each power source depending on the situations to avoid a blackout. The voltage stability was one of the evaluation criteria herein. The closer to 67 [kV] the voltage was controlled, the higher the score of the voltage stability. Although the voltage stability was adopted to evaluate the task performance, the operators were instructed that avoiding blackout is of the highest priority.

The task scenarios were constructed to intentionally provide a tradeoff between the voltage stabilizing and inspection tasks and realize a realistic task situation. When an inspection is performed, a power source must be shut down. Consequently, the total power supply decreases, resulting in a lower grid voltage. An inspection that is not performed in a specified time period may result in the failure of power sources, which also results in undesirable situations. The operators are required to control the SG grid voltage considering the possible grid voltage change while performing an inspection task.

The operation was evaluated on a daily basis with the final operation goal of completing a 1 day operation without causing a blackout. A blackout occurs when the grid voltage continuously deviates from 67.0 [kV] by more than 12 s in system time.

3.2. Procedures provided

Two types of operation manuals were prepared. One is a *basic manual*, and the other is a *procedure manual*. The basics of the SG grid system and the operation method necessary in performing the required task were described in the basic manual (e.g., rated power of each power source, required time to reach full power, and conditions for emergency stop). Meanwhile, an effective procedure to perform the required tasks was described in the procedure manual. This manual consisted of two parts: 1) procedure for voltage recovery and 2) procedure for inspection. Figure 3 shows an example of the procedure for voltage recovery, which must be referred to when the state of the grid voltage changes to “Caution” and provides information on the effective procedure to avoid blackout in the form of a flowchart. Figure 4 presents an example of a procedure for effective inspection. The appropriate time for inspection of each power source and line was explicitly provided to achieve effective inspections.

The basic manual provided minimum necessary information for performing the task, whereas the procedure manual provided information that will help operators in effectively performing a task without a deeper understanding of the SG grid system as long as the situation remains within the predefined (expected) ones.

The present study aimed to evaluate the effect of procedural information additionally given to operators. In the cognitive experiments, the operators (i.e., subjects) were divided into two groups:

Group A: Subjects are given only a basic manual.

Group B: Subjects are given both basic and procedure manuals.

3.3. Experimental scenarios

This section describes the task scenarios constructed for the cognitive experiments, which simulated the normal operation and expected abnormal and unexpected abnormal situations. The required conditions for each situation are as follows:

- Normal and expected abnormal situation: The scenarios already experienced in the practice runs and blackout can be avoided by just following the given procedures.

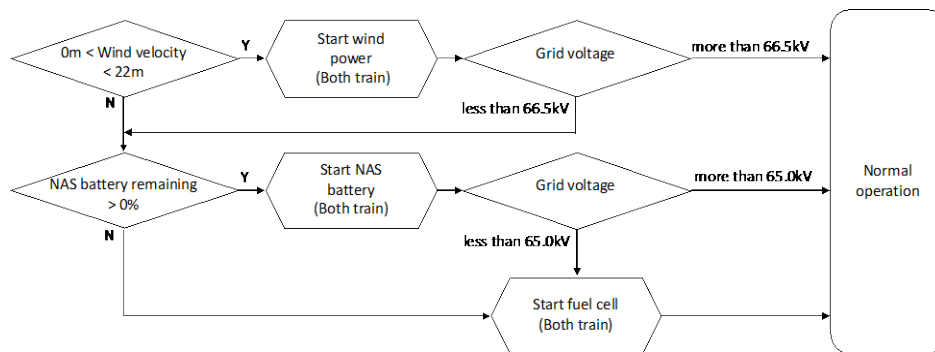


Figure 3 Example of the procedure for voltage recovery

Objectives	Time for inspection
NAS battery	0-1 o'clock
Fuel cell	2-4 o'clock
Wind power	5-7 o'clock
Power line	8-9 o'clock

Figure 4 Example of the procedure for inspection

- Unexpected abnormal situations: The scenarios not experienced in the practice runs and blackout cannot be avoided by just simply following the procedures and require improvisational operations.

Table 1 shows the scenarios for the expected abnormal situations. Table 2 presents the scenarios for the unexpected abnormal situations. These unexpected scenarios were constructed such that the operators cannot deal with the situations based on the knowledge obtained in the other unexpected scenarios, and all scenarios were highly independent of each other.

3.4. Timeline of the experiment

The cognitive experiments consisted of “Explanation,” “Practice runs,” and “Evaluation runs.” During the practice runs, the subjects performed training tasks using the normal and expected abnormal scenarios. The subjects in Group A performed the required operations based on their own understanding of the system without a procedure manual. In contrast, the subjects in Group B were instructed to perform the required operations by following the procedure manual. Here, it was emphasized for the subjects in Group B that following the procedure manual is important. The following are the actual steps in the cognitive experiments:

(1) *Explanation* (10 min)

A basic manual was given to the subjects in both groups, and a basic operation was instructed.

(2) *Practice for basic operation* (15 min)

The subjects in both groups performed practice runs based only on the basic manual.

(3) *Performance test* (15 min)

Scenario	Description
Heavy rain	Rain continues all day and no power obtained from Solar system
Wind power failure	Both train of Wind power fails in the morning and remain unavailable hereafter
NAS failure	Both train of NAS battery fails in the morning and remain unavailable hereafter
Failure of hydrogen supply	Hydrogen supply to fuel cell fails in the morning and the Wind power becomes unavailable because of strong wind velocity
Failure of all power sources in one train	One of the trains of all power sources fail in the morning

Table 1 Scenarios for the expected abnormal situations.

Scenario	Description
Grid voltage unknown	Indication of both SG grid voltage and Voltage quality indicator lost
Upstream power unstable	The incoming power from upstream power stations fluctuates widely
Typhoon	Heavy rain and strong wind continue long time and no power from Wind and Solar is available
NAS battery and Solar power unavailable	Both trains of NAS battery and Solar power fail

Table 2 Scenarios for the unexpected abnormal situations.

The performance test, in which subjects must perform easy scenarios, was conducted for each subject to exclude those with an extremely degraded performance for operations. The subjects who caused a blackout of more than two times were excluded from the experiments.

(4) *Instruction of the procedure manual* (10 min)

The subjects in Group B were provided with the procedure manual and given an explanation of the procedure. They were expected to learn the operation using the procedure and understand the effectiveness of using this procedure.

(5) *Training run 1* (20 min)

The subjects in both groups performed practice runs using the scenarios of 4 days of normal operations. Although the subjects in Group A dealt with the situations based on their own judgments, those in Group B dealt with them following the procedure. An experimenter monitored if a subject followed the procedure; otherwise, a warning to follow the procedure was given to that subject.

(6) *Training run 2* (20–120 min)

The subjects in both groups performed practice runs using 5 days of expected abnormal scenarios. This training phase aimed to achieve a certain level of operation skill for all the subjects; hence, they were asked to repeat the trial runs to complete a 1 day operation without blackout up to five times. Training run 2 comes to an end when a subject successfully operates all the expected abnormal scenarios without a blackout. The subject who failed to fulfill this goal was excluded from further experiments. This screening procedure aimed to equalize the capability of a subject to deal with the expected abnormal situations.

(7) *Evaluation runs* (75 min)

The subjects in both groups performed 6 days of normal operation, 5 days of expected abnormal situations, and 4 days of unexpected abnormal scenarios. They were never notified of the occurrence of abnormal events, which they never experienced in the training periods. The evaluation criteria were recorded during this period and utilized for a statistical analysis to compare the performances of Groups A and B.

(8) *Mental workload evaluation* (10 min)

The mental workload was evaluated using NASA-TLX for the different experimental conditions. The mental workloads of the subjects during the normal operation, expected abnormal scenarios, and unexpected abnormal scenarios were compared for both groups.

3.5. Subjects and evaluation criteria

A total of 41 university students participated in the experiment. The subjects were grouped into Groups A and B based on a stratified randomization. All subjects had never participated in any experiment using SG grid simulators. The experimental protocol has been approved by the ethical committee of the School of Engineering of Tohoku University. Screening was conducted on the basis of the results of the training runs to reduce individual difference in operational capability during the final evaluation phase.

The performance of each subject was evaluated on the basis of the following criteria:

- number of blackouts,
- number of inspection time over,
- grid voltage stability,
- mental workload, and
- number of operations.

4. RESULTS

4.1. Screening result

No subject was screened out on the basis of the first screening process in the performance test. One of the subjects in Group A could not complete the operation without blackout in the wind power failure scenario in the five times of trial; hence, this subject was excluded from further experiments. The number of subjects and their personal data after the screening process are summarized as follows:

- Group A

Number of subjects: 20 (male: 13; female: 7)

Age range: 20–24 years

Average age: 22.2 years

- Group B

Number of subjects: 20 (male; 12; female: 8)

Age range: 21–25 years

Average age: 22.1 years

4.2. Evaluation results

The evaluation results in the evaluation runs were summarized below.

Figures 5 and 6 illustrate the average number of blackouts and the inspection time over. Figure 7 presents the average grid voltage stability score.

Meanwhile, Figure 5 shows that the average number of blackouts of Group B in the unexpected abnormal situations was significantly higher than that of Group A ($p = 0.000 < 0.01$). In contrast, the average number of blackouts of Group A in the expected abnormal situations was higher than that of Group B ($p = 0.081 < 0.1$). No significant difference in the two groups was found in the normal situation.

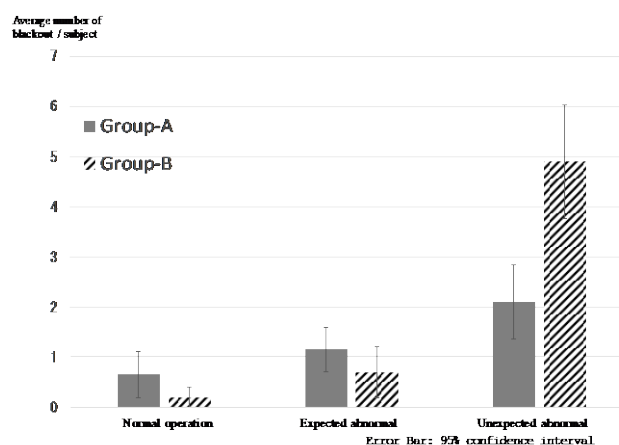


Figure 5 Average number of blackouts

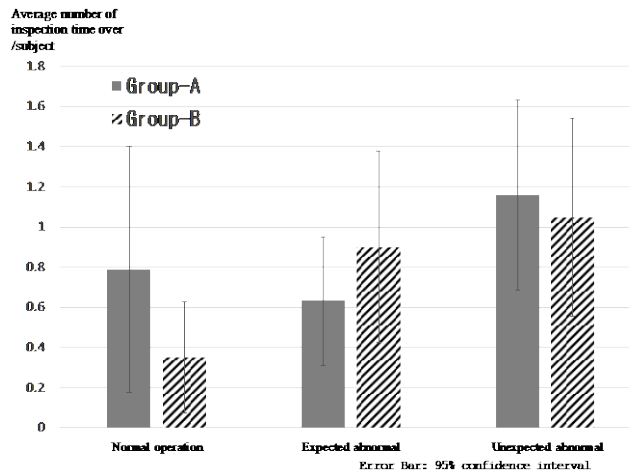


Figure 6 Average number of inspection time out

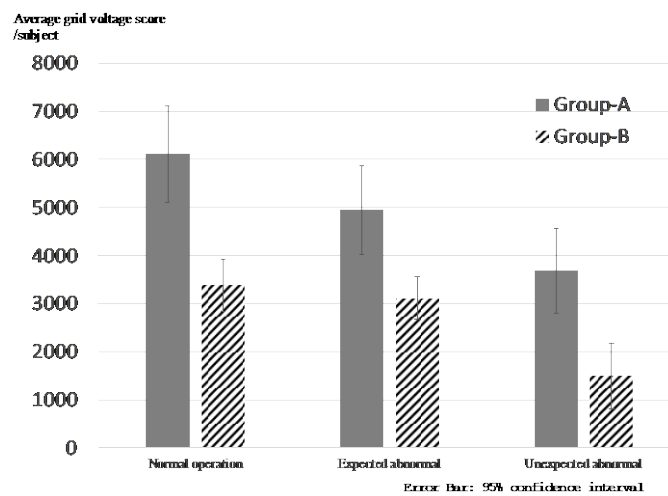


Figure 7 Average voltage stability points

Figure 6 exhibits no significant difference in the average number of inspection time out in the two groups for all situations.

Figure 7 depicts that the average voltage stability points of Group A are significantly larger than those of Group B for all situations (normal operation: $p = 0.000 < 0.01$; expected abnormal: $p = 0.000 < 0.01$; unexpected abnormal: $p = 0.000 < 0.01$). The standard deviations of Group B for the normal operation and the expected abnormal situation were significantly smaller than those of Group A (normal operation: $p = 0.048 < 0.05$, expected abnormal: $p = 0.046 < 0.05$).

Figure 8 shows the evaluation results of the mental workload. No significant difference in the WWL score was found in the two groups for all situations. Table 3 presents the evaluation result of each elemental scale in calculating the WWL score. The elemental score of “Frustration” for the expected abnormal situation and “Time pressure” for the unexpected abnormal situation showed a significant difference in the two groups:

(“Frustration” for the expected abnormal situation: $p = 0.027 < 0.05$. “Time pressure” for the unexpected abnormal situation: $p = 0.042 < 0.05$).

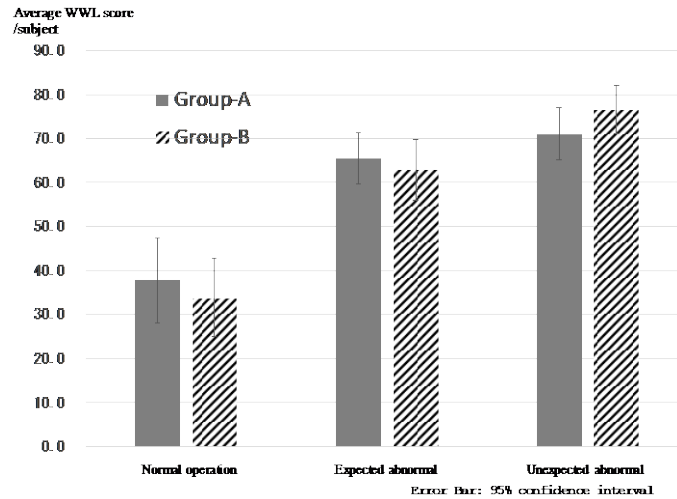


Figure 8 Evaluation results of the mental workload (WWL)

Title	Group	Normal operation	Expected abnormal	Unexpected abnormal
Mental demand	A	32	54	70
	B	30	62	75
Physical demand	A	16	35	42
	B	14	39	46
Temporal demand	A	40	67	66
	B	38	67	79
Effort	A	32	68	74
	B	33	60	71
Frustration level	A	42	72	64
	B	35	58	66
Performance	A	35	62	69
	B	35	65	80

* p<0.05

Table 3. Elemental scale for calculating the WWL score

4.3. Discussions

The results of the average number of blackouts for Groups A and B showed an opposite tendency in that less number of blackout for Group B was observed for the expected abnormal scenarios, whereas less number of blackout for Group A was observed for the unexpected abnormal scenarios. This result illustrates that the subjects in Group B, who were accustomed to performing an operation following the procedure, could avoid causing a blackout in expected situations, but caused a larger number of blackouts compared with those in Group A. In summary, the average performance of the subjects in Group B was better in the expected situations, including a normal operation, and was degraded in the unexpected situations.

The results of the average number of inspection time over showed an insignificant difference between Groups A and B. This result implies that the existence of a procedure manual did not influence the time available and affordable for inspection during operations.

The results of the grid voltage stability points showed that smaller points (lower stability) were observed for the subjects in Group B for all situations. The closer to 67 [kV] the grid voltage is maintained, the higher the voltage stability point per unit time is given. The subjects in Group B began the voltage recovery action following procedures only when the grid voltage exceeded the predefined threshold value. In contrast, the subjects in Group A performed voltage control actions based on their own judgment to control the voltage to 67 [kV] as close as possible, which may have resulted in the higher voltage stability points.

The difference of the variances in the voltage stability points for each group should be mentioned herein. The variance of the subjects in Group B was smaller, implying that using the procedure manual may have a positive influence in reducing individual variations among subjects.

As for the evaluation results of the mental workload, no significant difference in the WWL score was found among the two groups for all situations. This result implied that the addition of the procedure manual and imposing instructions to follow the procedure have not directly led to an increase of the mental workload. The elemental score of “Frustration” for an expected abnormal situation and “Time pressure” for an unexpected abnormal situation showed a significant difference among the two groups.

For the expected abnormal scenarios, the “Frustration” score of Group A was higher than that of Group B because of the differences in dealing with the situations. Although the subjects in Group A had to make a decision to mitigate the abnormal situations based on their judgments, those in Group B could rely on the procedure, which resulted in a lower stress level and a sense of unease as long as the situations stayed inside the expected and manageable range.

Meanwhile, the score of “Time pressure” showed a larger value for the subjects in Group B for the unexpected abnormal scenarios. When an abnormal event happens, the subjects in Group B first try to manage the situation by following the procedure and then begin to take remedial actions based on their own judgments, resulting in insufficient time available to avoid blackout. The wasted time trying to follow procedures may have led to the higher “Time pressure” score.

5. CONCLUDING REMARKS

The present study experimentally confirmed that strong dependence on a procedure may deteriorate performance under unexpected situations outside the range of procedures. An SG Simulation environment, in which dynamic decision making is required to deal with the given situations to avoid blackout, was adopted herein as the experimental testbed. The results of the cognitive experiments showed that preparing the procedures with an instruction to follow them may have a negative influence on the operators' performance in dealing with unexpected situations outside the coverage of the given procedures, although they have a positive influence in terms of reducing the variation of the operator dealing with normal and expected situations. The result of the present study is consistent with the expectation based on the empirical understanding of the work domain. However, the authors believe that it is a significant achievement to confirm it experimentally based on a realistic simulation environment. Further research is definitely required to develop an effective training strategy to enhance the capability of dealing with unexpected situations.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Number16H04625.

REFERENCES

- [1] Hollnagel, E., Woods, D. D., & Leveson, N. (2006). Resilience engineering: Concepts and precepts. Ashgate Publishing, Ltd.
- [2] Hatamura Y.(Ed) (2012) Final Report of Investigation Committee on the Accident at Fukushima Nuclear Power Station of Tokyo Electric Power Company, July 23.
- [3] Kitamura, M. (2014). Resilience engineering for safety of nuclear power plant with accountability. Resilience engineering in practice, 2, 47-62.
- [4] Nuclear Regulation Authority. (2013), New Regulatory Requirements for Light-Water Nuclear Power Plants-Outline-, <http://www.nsr.go.jp/data/000067212.pdf>. (accessed 19 January 2016)
- [5] BWR Operator Training Center Corporation. (2015),Reconsideration for Improvement of Operator Training Program to Take All-inclusive Events into Account. BTC Training Report, No.96, pp.7-8.
- [6] Hollnagel E., Paries J., Woods D.D., Wreathall J., (Eds.) (2011) Resilience Engineering in Practice, Ashgate Publishing Co.
- [7] Mintzberg: Rethinking Strategic Planning Part I Pitfalls and Fallacies; Long Range Planning, Vol.27, pp.12-21 (1994).
- [8] Horiuchi, Y., Nakanowatari, H and Takahashi, M. (2014), Experimental Study on the Effect of Rule Presentation on Task Performance. Journal of the Atomic Energy Society of Japan, Vol.13, No.4, pp.145-154.