

# Development of Interactive Educational Games about Human Error for Railway Personnel

Midori Inaba<sup>a</sup>, Ryo Nakamura<sup>a\*</sup>, Takashi Imaizumi<sup>a</sup>, Osamu Onodera<sup>a</sup>,  
Yuichi Takeda<sup>a</sup>, Ken Kusukami<sup>a</sup>, and Shigeru Haga<sup>b</sup>

<sup>a</sup> Department of Industrial Engineering and Management Systems  
East Japan Railway Company  
Nisshin-cho, Saitama 331-8513, Japan

<sup>b</sup> Department of Psychology  
Rikkyo University  
Niza, Saitama 352-8558, Japan

## ABSTRACT

One approach to prevent accidents caused by human errors is to promote understanding of errors by those who are required to be safety-conscious in their work. We report the process of developing a serious game, in which people working for the railway learn a slip, a human error of an action that is carried out incorrectly. To help players understand that anyone can make a slip, we have been developing games to give them with a chance to experience their own error. In one of the games, the player's task is to classify (or categorize) each of the presented images into one of two categories under time pressure. After making the basic features of the game, we carried out experimental trials to examine impacts of factors related to the setting. We plan to reflect the findings of these trials on the setting of the game to increase the likelihood of the players' slip. We continue to develop this game and the additional short commentary about a slip tailored towards railway personnel. This game will support people who work in railway companies to comprehend a human error and operate safely, resulting in reduction of accidents caused by a human error.

**Keywords:** Serious Game, Human Error, Slip, Railway Personnel, Education

## INTRODUCTION

Competencies of personnel who work in safety-sensitive operations involve understanding of human factors and human errors (Hetherington, Flin, and Mearns, 2006; O'Neill, Goffin, and Gellatly, 2012; Yule, Flin, Paterson-Brown, and Maran, 2005). In order to serve the safe transportation system, the public railway companies are responsible for supporting these personnel to comprehend the human error. For this purpose, we have developed educational programs, in which people working for the railway learn the human error. Considering variety of working conditions and jobs in railway companies, the programs needs to fulfill the conditions as below:

- Personnel can use the program individually on a computer using their time not engaged in work.

Safety Management (2019)

- They can learn a human error regardless of their age, job, or area.
- The program provides them fun along with contents to be learnt.

\* The current affiliation is Railway Technical Research Institute, Kokubunji, Tokyo 185-8540, Japan. To meet these requirements, we determined to apply the format of a game to our learning programs. The learning program, although not one solely for enjoyment, is called a serious game (Pourabdollahian, Taisch, Kerga, 2012). Recently, the approach of serious games has been introduced in safety management (Haga, Onodera, Yamakawa, Oishi, Takeda, and Kusukami, 2013; Tesei, Barbieri, and Kessel, 2012).

We have created five serious games to teach different types of a human error. This report shows the process of developing one of our games. This game is aimed at having the player understand not only how a human error occurs, but that anyone can make a human error. Because it is indicated that results or feedbacks to player's performance in the game affect the players' learning (Garris, Ahlers, and Driskell, 2002), we have especially elaborated the game design to provide the player with a chance to experience their own error in the game. In the current report, we first show the basic features of our game. Then, we report investigations of impacts of factors related to the setting of the game that may increase the players' chance to make the error.

## FREMEWORK FOR DEVELOPING OUR GAME TO TEACH A SLIP

A human error that is taught in the game we report in this study is a slip. A slip is an error when the appropriate action is carried out, but incorrectly (Reason, 1990). Rasmussen (1983) categorized the slip as a skill-based error when the routine for actions toward one stimulus is activated toward the other stimulus. Such an error is often considered to occur when responding to infrequent stimulus among frequent stimuli (Norman, 1981). In addition, it is suggested that a slip is sometimes a reflection of the speed-accuracy tradeoff, so that increasing a speed in responses contributes to the slip (Baars, 1992).

According to these implications, we decided the basic features of the game to cause participants' slips in the game as follows.

### Basic Features

The education program about a slip is consisted of two parts: the game named "Detect aliens!" and the short commentary on a slip. In this game, participants are hypothesized to be an officer at the security checkpoint to detect a few aliens who have invaded and tried to conquer the earth. The participants' task is to match the presented image against the list of the peculiar characters of aliens' looking, decide whether the image is an alien or a human, and press the button corresponding to the alien or the human.

In detail, the game includes a set of trials, followed by instructions including practice trials. In each trial, an image of a face of an alien or a human is presented (Figure 1). We reduced the frequency of trials that presents images of the alien compared to those of the human. It was predicted that such setting of the frequency result in participants becoming accustomed to responses to the images of humans. The participants were expected to need more control on their action to respond to the images of the alien correctly. Moreover, they might unintentionally downgrade retaining their control on their action because the participants press the same button continuously. To promote this attitude and induce participants' slips, we set a time limit in each trial. Following the game, the short commentary included explanation about a slip and instances of accidents or incidents caused by a slip in three types of the job: station staff/operation command, driver/conductor, and maintenance staff. These instances are like the following: an incorrect action in repetitive pressing of buttons to set the track closure using a device in the tablet form.

Safety Management (2019)

<https://openaccess.cms-conferences.org/#/publications/book/9781145812100-5>



## EXPERIMENTAL TRIALS TO ELUCIDATE FACTORS INDUCING A SLIP

The design of the game has been further developed by manipulating factors that may increase the possibility of making a slip in players. We focus on three factors.

### a) Categorization rule

Categorization rules to decide the image as an alien or a human are concerned to four parts of a face of the presented image: eyes, nose, mouth, and hair (Figure 2). Less discriminability of the aliens' peculiar parts from the humans' facial parts is considered to increase participants' chance to make an error. It is also possible that the complicated rules are predicted to increase many incorrect responses. Considering of our goal, however, the design of the game is desirable to induce only slips, not other errors such as mistakes caused by participants' confusion. We have to balance the discriminability and complexity in setting the categorization rules.

### b) Intervals of presenting images of an alien

A greater interval between trials in which an alien appear is considered to make it difficult for the player to respond to the aliens correctly. In our experimental trials, we examined impacts of its interval while the frequency of presenting these images is the same. Thus, the set of trials that have the greater intervals of aliens' appearance include the smaller interval at the same time.

### c) Length of the time limit in each trial













Severe time pressure is expected to lead participants to emphasize on speed in their responses more than accuracy to detect aliens, resulting in a higher chance that a slip occurs.

## Pilot Study

We carried out a pilot study to check the basic features of our game. The features of the game in this study were as described below:

- The game included a set of 300 trials. Of these 300 trials, 30 images of an alien were provided every 10 trials (Table 1).
- Categorization rules were: a) up-angled eyes to be an alien, b) pig nose to be an alien, c) pouting mouth to be a human even with a) or b), d) long hair to be a human even with a) or b).
- Time limit in each trial was 3 seconds.

Participants were 58 undergraduate students. They individually played the game using a computer. Results showed 80% of the participants made an error that categorized the images of an alien as a human at least once. We considered the setting in the pilot study to be sufficient to cause slips in participants. Note that, however, participants

	Alien	Human
Eyes		 
Nose		 
Mouth		 
Hair		 

required approximately 20 to 30 minutes playing the game. Because it seemed to be too long for personnel in our company to willingly play the game and learn a slip, we cut 230 trials and left 70 trials including 7 images of an alien in the next version of the game.

### Experimental Trial 1

#### 1) Aims of this experimental trial

Experimental trial 1 was performed with adults to confirm the reliability of the setting used in the pilot study. This experimental trial was also aimed at testing if 70 trials were sufficient for participants to have an experience of a miss to detect aliens. In addition, we compared the error rate in two types of the length of the time limit.

#### 2) Methods

Participants were 87 adults in total. Of 87 participants, 48 participants were required to respond in each trial within 3 sec. For the rest of the 39 participants, the time limit was 2.5 sec. Participants played the game including 70 trials. Methods other than mentioned above were the same as the pilot study.

#### 3) Results & Discussion

Results were shown in Table 2. 67 % of participants in the 3-sec time limit group categorized an alien as a human at least once. However, the averaged percentage of correct responses was considerably low even in this condition (64%). This percentage further decreased to 44% in the 2.5-sec time limit group. The percentage of the participants who committed a miss to detect an alien was lower in this group (28 %). Instead, time-out errors increased.

Our goal is to cause participants to respond to humans correctly but to aliens incorrectly. The low percentage of correct responses and the high percentage of time-out errors indicated that it was hard for participants to respond even to humans correctly. Many participants gave remarks that the categorization rules were too complicated to categorize the image in the short duration. Moreover, the shortened time limit appeared to confuse and freeze the participants out. Additionally, it seemed to be still a little long for personnel to play the game using their time not engaged in work. A few participants took over 15 minutes in playing the game with 70 trials.

Table 1. Interval of aliens' appearance. The image of a human was presented in trials between images of aliens (depicted by arrows).  
Exp. 1: Experiment 1; Exp.2: Experiment 2

	Exp. 1	Exp.2		
		Interval of aliens' appearance		
		Condition A	B	C
0	Start	Start	Start	Start
10	Alien	Alien	Alien	Alien
20	Alien	Alien	Alien	Alien
30	Alien	Alien	Alien	Alien
40	Alien	Alien	Alien	Alien
50	Alien	Alien	Alien	Alien
60	Alien	Alien End	Alien End	Alien End
70	Alien End			

Correct response (of all trials)	Time-out error (of all trials)	Miss (of trials with aliens)	False Alarm (of trials with humans)	Participants who missed at least in one trial
<b>Experimental Trial 1</b> Table 2. Average percentages of responses. Miss is an error that participants categorized an alien as human. False alarm is an error that they categorized a human as an alien Time limit in each trial 3 seconds				
64	10	0.5	16	67 (32/48 participants)
2.5 seconds				
44	43	58	56	28 (11/39 participants)
<b>Experimental Trial 2</b> Interval of aliens' appearance				
Condition A				
97	0.009	4	1.7	20 (6/32 participants)
Condition B				
96	0.02	3	2	10 (3/30 participants)
Condition C				
97	0.02	0.6	0.8	3 (1/31 participants)

On the basis of these results, we have decided the setting of the game in the next experimental trial as below.

- Categorization rules should be simpler.
- The length of time limit in each trial should be 3 sec.
- The number of trials in total of the game should be 60.

## Experimental Trial 2

### 1) Aims of this experimental trial

There were three aims of this trial. One was to check the percentage of participants who categorized an alien as a human incorrectly using the revised version with 60 trials. The other was to check the modified categorization rules in order to result in participants responding to humans easily. The final aim of this experimental trial was to examine an impact of the interval of aliens' appearance on the error rate.

### 2) Methods

84 adults in total played the game with 60 trials including 5 trials in which an alien appeared. These participants were divided in 3 groups. Each group was allocated with one of three conditions of the interval setting (Table 1).

### 3) Results & Discussion

As shown in Table 2, 20 % of participants categorized the image of an alien as a human at least in one trial when aliens' appearance considerably constantly, that is, when the set of the game did not include the short interval of the aliens' appearance. The average rate of correct responses was 97% and the percentage of time-out errors was 0.009 %. We further compared the response time among correct or incorrect responses to an alien or a human (Table 3). Participants in all of 3 groups pressed the button most fast when they categorized aliens as humans incorrectly.

Safety Management (2019)

Modified categorization rules reduced participants' errors to categorize an alien as a human. However, the average RT in incorrect responses was shorter than the average RT in correct responses in trials with aliens' appearance although

General average	Alien		Human	
	Correct	Incorrect	Correct	Incorrect
Interval of appearing aliens				
Condition A				
1113	1087	809	1115	1936
Condition B				
1312	1249	1155	1317	2049
Condition C				
1170	1117	650	1177	1916

participants spent longer time when they made incorrect responses compared to correct responses. This finding indicated that participants required great cognitive processing to respond correctly to the alien. Such cognitive processing is considered to involve the attentional control over the low-frequent events. Without it, the participants might have pressed the button incorrectly towards aliens. This kind of incorrect responses seems to fit

Table 3. Average response time depending on the participants' responses to an alien or a human

several conditions of a slip that were suggested in previous research. Therefore, we saw the possibility that our modified categorization rules can cause players to make a slip. Our next goal is to further revise these rules and increase the percentage of participants' responses that were observed in this experimental trial.

## CONCLUSIONS

We have shown the process of developing our game to teach personnel about a slip. After we created the basic features of the game, we examined impacts of factors related to the setting of the game in order to make it sure for players to make slips in the game.

The development of our serious games that teach human errors is still continuing. As for "Detect aliens!," we need to find out optimal categorization rules. We further plan to develop the short commentary about human errors for railway personnel. Finally, we will elucidate how much an experience of an error in the game and the short commentary contribute to learning of human errors along with the player's subjective feeling of self-relatedness to the errors.

## REFERENCES

- Baars, B.J. (1992), "Experimental Slips and Human Error", New York: Plenum Press.
- Garris, R., Ahlers, R., Driskell, J.E. (2002), "Games, Motivation and Learning: A research and Practice Model", *Simulation Gaming* 33, pp.441-466.
- Haga, S., Onodera, O., Yamakawa, A., Oishi, A., Takeda, Y., Kusakami, K., Kikkawa, T. (2013), "Training of resilience skills for safer railways: developing a new training program on the basis of lessons from tsunami disaster", *Proceedings of the 5<sup>th</sup> Symposium of the Resilience Engineering Association*.
- Hetherington, C., Flin, R., Mearns, K. (2006), "Safety in shipping: The human element", *Journal of Safety Research* 37, pp.401-411.
- Norman, D.A. (1981), "Categorization of Action Slips", *Psychological Review* 88, pp.1-15.
- O'neill, T.A., Goffin, R.D., Gellatly, I.R. (2012), "The knowledge, Skill, and Ability Requirements for Teamwork: Revising the Teamwork-KSAT Test' validity", *International Journal of Selection and Assessment* 20, pp.36-52.
- Pourabdollahian, B., Taisch, M., Kerga, E. (2012), "Serious Games in Manufacturing Education: Evaluation of Learners' Engagement", *Procedia Computer Science* 15. pp.256-265.
- Rasmussen, J. (1983), "Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models", *IEEE Transactions on Systems, Man and Cybernetics SMC-13*, pp.257-266.
- Reason, J. (1990), "Human Error", Cambridge: Cambridge University Press.
- Tesei, A., Barbieri, A., Kessel, R. (2012), "Survey on serious games applied to security, safety, and crisis management", *Procedia Computer Science* 15. pp.320-321.
- Yule, S., Flin, R., Paterson-Brown, S., Maran, N. (2005), "Non-technical skills for surgeons in the operating room: A review of the literature", *Surgery* 139, pp.140-149.