

# A Study on the Evaluation Method of Employee's Motivation to Participate in Safety Activities Using Daily Work Report-Competency Evaluation of Reporters Based on Content Analysis of Information Noticed at Work Sites

Nanako Yokozawa and Yusaku Okada

Keio University, Science for Open and Environmental Systems, Yokohama, Kanagawa, Japan

## ABSTRACT

To improve the quality of safety activities, many companies are aiming to increase employees' understanding of safety activities and motivate them to participate in them. The main method to measure the willingness to participate is through questionnaires, but it takes time and effort. Therefore, we focused on information on "Awareness" events reported by on-site employees in daily work report. We call this information "Information on Preparing for the Unexpected". The purpose of this study was to develop a method to evaluate employees' understanding of and willingness to participate in safety activities using "Information on Preparing for the Unexpected". Through experiments, trouble-related data were collected at several workplaces, and text mining was conducted. We worked on the construction and validation of the system.

**Keywords:** Safety activities, Safety management, Text mining, Human reliability

## INTRODUCTION

Gathering information on "Awareness" events in daily operations is also a part of safety activities. Seven characteristics of human resources capable of such "Awareness" were identified. The more such employee, the better the contents of "Information on Preparing for the Unexpected" will be, and the better the quality of the information collected will be.

[1] "Information gathering and information sharing.

[2] "Have a moderate sense of anxiety/comfort.

[3] "Thoroughness of basics."

[4] "Positive expressions."

[5] "The idea of everyone helping each other."

[6] "Create 'safety and security' together."

[7] "Invest in the future."

**Table 1.** Details of good "Awareness" employee rating group.

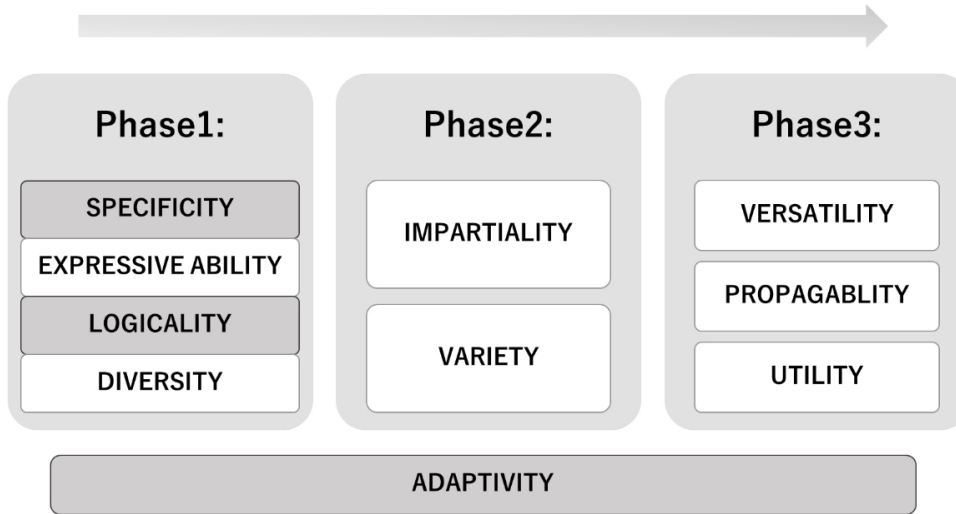
	Details
IMPARTIALITY	The index to measure whether the amount of each content is evenly distributed without bias.
SPECIFICITY	The index that considers a wide range of related PSFs (Performance Shaping Factors) from multiple perspectives, such as "Psychological," "Work," "Environment," and "Information. In addition to human factors, organizational and management factors are included.
EXPRESSIVE ABILITY	The index to measure whether the image clearly convey to others. The issues organized, detailed, and the background of the issues are explained, and it take into account the characteristics of the reader.
LOGICALITY	A wide range of background factors of the issue described are widely assumed and many possibilities are described.
DIVERSITY	The index to measure whether differences in skills, skills, knowledge, personalities, and other characteristics are assumed, and whether the opinions and ideas of various people are incorporated.
VARIETY	The index to measure whether that measures the number of types of topics covered.
VERSATILITY	The index to measure whether the content is useful for your department. It should be easy to understand so that members of your department can refer to it.
ADAPTIVITY	The index to measure whether a wide range of employees is targeted.
PROPAGABLITY	The index that consider whether it contains details of the adverse impact of the identified problem on the Covered System; that is, the characteristics of the risk, the extent of the impact if the risk materializes, and the factors that control the extent of the impact.
UTILITY	The index for weighting the extracted background factors by their importance.

### Proposal

Based on the results of the characteristic analysis of text and past case studies on human error factors, the following 10 characteristics were evaluated. We call these characteristics "Good 'Awareness' Employee Rating Group". These employee characteristics are the ones that we are trying to improve through safety activities, and we believe that they are causally related to the ability to discover "Awareness" in the workplace.

The ten characteristics shall be graded.

- Phase1: SPECIFICITY, EXPRESSIVE ABILITY, LOGICALITY, DIVERSITY
  - Measures whether a wide range of factors are taken into account.



**Figure 1:** Good “Awareness” employee rating group.

- Phase2: IMPARTIALITY, VARIETY
  - Measures whether the report has quantity and quality.
- Phase3: VERSATILITY, ADAPTIVITY, PROPAGABILITY, UTILITY
  - Measure whether the report considers not only the factors but also the direction and measures beyond the factors.

These evaluation phases were designed in accordance with the following policy. Characteristics are related to each other as shown in Fig. 1.

- Measuring the former phase enables the evaluation of the next phase.
- ADAPTIVITY can also be measured in the Phase1 in a simplified manner.
- Each of the 10 characteristics is scored on a 10-point scale and treated equally.

Therefore, of the four items in the Phase1, EXPRESSIVE ABILITY was excluded from this study based on the expectation that there would be no individual differences in Japanese writing. DIVERSITY is a characteristic that is close to a person’s own values, such as “proficiency, skill, and personality”. On the other hand, SPECIFICITY and LOGICALITY, are suitable as introductory training because they give employees a new perspective and change easily. In other words, both EXPRESSIVE ABILITY and DIVERSITY were excluded because they are expected to grow over a long period of time. ADAPTIVITY, which is not in the Phase1, can be measured in a simplified manner. Therefore, in this study, we evaluated the Phase1, SPECIFICITY and LOGICALITY, as well as ADAPTIVITY.

Following this we are discussing the three characteristics to be measured. The ability to capture various perspectives on the site leads to SPECIFICITY. In addition, the willingness to explain things to others in an easy-to-understand manner and to delve deeper into the background factors will lead to LOGICALITY. To assume to a wide range of employees will lead

to ADAPTIVITY. Therefore, these characteristics are the target of this study with the understanding that they will lead to the improvement of “The idea of everyone helping each other.”, one of the seven characteristics of human resources who can “Awareness” as mentioned above. It means “willingness to participate in safety activities”.

In analyzing the textual data, we decided to attempt to extract subjects by taking the intent of the text as the subject matter. Based on past accident reports and studies, the topics of the reports were classified into 18 themes. For each of these subjects, we understood the semantic content and prepared corresponding keywords. In addition, example sentences were assumed for each keyword and were made into subcategories. Next, synonyms were selected from the keywords based on word-to-word connections. The keywords were filtered as a result of feature analysis, considering the experimental data as the training data (Fig. 2).

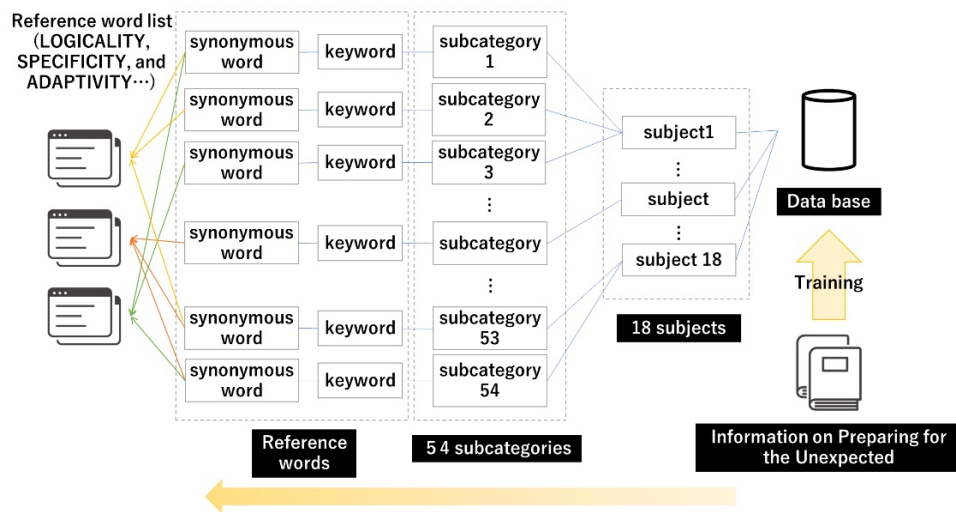


Figure 2: Design of the database.

Table 2. Examples of subcategories and the number of subcategories.

	subcategory	keywords	synonyms	The number of subcategories
ADAPTIVITY	5.2 Cannot talk casually with organization members.	carefree	unfriendly or unkind	34
LOGICALITY	18.2. Actively participating in safety activities is a waste of time and a hassle.	waste or hassle or time	spare or burden or extra or bothersome or time-consuming	20
SPECIFICITY	9.1 There is a difference in perception between generations.	Generation gap	generational or obsolete	14

The Table 2 and The Table 3 show keywords and synonyms for an example of an item that captures the three characteristics (LOGICALITY, SPECIFICITY, and ADAPTIVITY). This is the so-called reference word list. Subcategories were created for a total of 54 items: LOGICALITY (20 items), SPECIFICITY (14 items), and ADAPTIVITY (34 items), corresponding to the example sentences. We thought that the meaning could be aggregated by categorizing the subcategories in stages of abstraction, such as from word list to subcategory and from subcategory to subject matter.

The purpose of the system is not the scoring itself, but rather to support the development of the employee's competency. Therefore, the system provides

**Table 3.** Examples of reference word list of LOGICALITY (in Japanese).

subcategory	keywords	synonyms
1.1 Skip the pre-work meeting.	meeting	conference
2.2. Supervisors make safety a top priority as an organizational policy, but do not review team goals and their plans according to the situation.	rework or policy or plan or adjust	near (plan - adjust) or review or fix or quota or replace or arrange
4.1. Poor communication	say or stupid	deny
1.2. Instructions are vague	directive or ambiguous or vague	work instructions
1.3 Lack of guidance	guidance	instruction or training or parenting
11.6 Unfair treatment	unfairness	insecurity
11.2 Not working passionately and not energizing colleagues	passionate or energizing	enthusiastic negative or alienating or enlivening
11.3 If there is a problem, he/she will be held responsible for it.	near (responsibility - take) or shift responsibility or leading	party consciousness or anxiousness
11.4. An atmosphere of "shame for failure"	embarrassing or public	shame
11.7 Colleagues have a bad image of you	image	impression or negative or negative evaluation or condescending or misunderstanding
10.2. It does not appear to be growing.	growth	reaction
18.2. Actively participating in safety activities is a waste of time and a hassle.	waste or hassle or time	spare or burden or extra or bothersome or time-consuming
13.3. Doesn't want to inconvenience others or other departments.	near (others - annoyance)	adaptation or accumulation or near (experience - how many times)
14.3. Overconfidence in one's own abilities	overconfidence	prideful or proud or perfect
10.1. Effects, changes are not visible.	effect or change	
7.1 Not immediately reported and shared within the team	share	(information and sharing) or information exchange
9.4 Does not work with team members to learn and work on issues toward a common goal	cooperative	cooperation
6.1 No feedback received.	feedback or response or action	
11.1. There is no atmosphere of full commitment when difficult things have to be done	atmosphere	customs or environment
9.3 They do not always pay attention to the status and progress of other members' work.	progress or near (attention - pay)	trust or leave or union

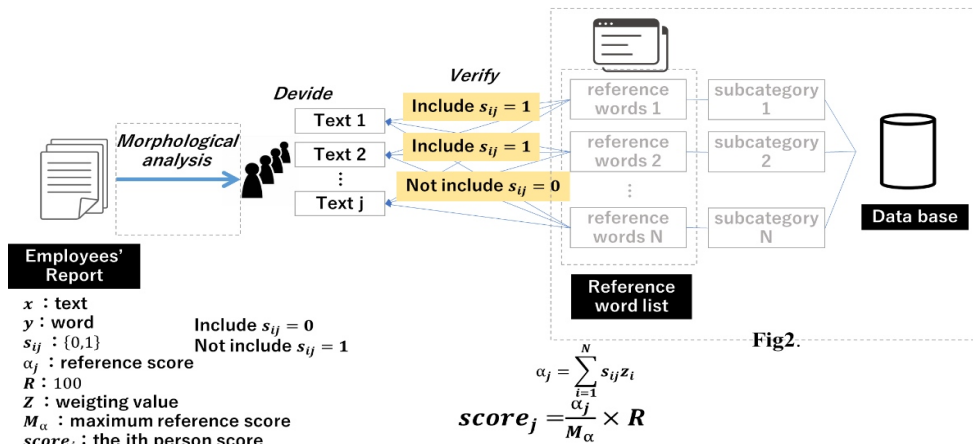


Figure 3: Overall view of the system.

advice to employee to deepen their understanding of safety activities. We prepared the advice data as a function that can be applied to the safety education of employee. The advice comments are based on standards developed over a 10-year period at one business by a human factors expert. They have been used to provide feedback to employees after Safety Lesson and have been well received by employers. The items have been revised and adapted based on the situation of the employees and previous studies. Each score is classified into six levels and used as a guideline for providing feedback. The following considerations were taken into account in the design.

- Ideal employee:  $40 > score \geq 32$
- Employees at the passing line:  $score = 24$
- The lower the score, the more conscious the comments are of the level above.
- Each item provides a guideline for future action.

In addition, extra points were allocated to employees with a particularly high level of skill, giving them a maximum of 100 points.

### Experiment

Outline: An experiment was conducted to simulate the collection of “Information on Preparing for the Unexpected” in a medical setting.

Purpose: Measurement of employees’ (Participants’) ability to “Awareness”.

Participants: Twenty-nine employees took the 90-minute Safety Lesson for the first time.

Experiment Design:

- The aim of lesson is improvement of SPECIFICITY/ LOGICALITY/ ADAPTIVITY.
- Ask students to describe hierarchically the issue of gathering information related to safety in the workplace.
- The human factors expert who conducted a Safety Lesson evaluated each employee.

- The response time was two hours.

Data from 29 employees, consisting of 703 sentences from the experiment, as described earlier. The mean was 24.24 sentences with a standard deviation of 3.54 sentences. The maximum character range was 4837 and the minimum 634. These collected experimental data are then used to explain the system built.

### Analysis

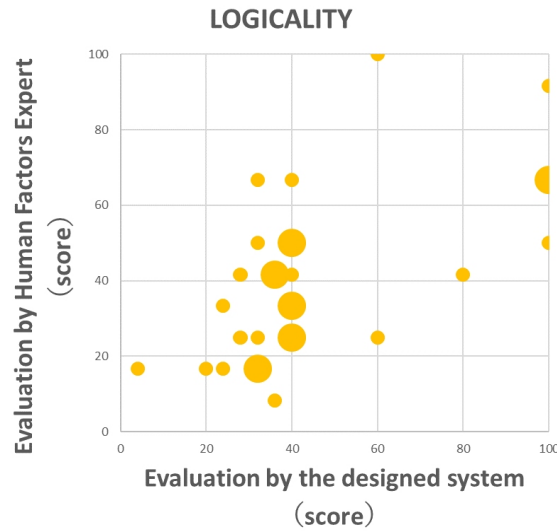
The experimental data in this study were text data. This was analysed using the morphological analysis function of KHcoder3 (Higuchi 2014, 2017), which is capable of text mining.

The filtered documents are matched against the reference word list. Due to the small size of the experimental data, we did not consider the effect of word level frequency and employed 0/1. Each sentence is matched against reference words in each of the three subcategories (LOGICALITY, SPECIFICITY, and ADAPTIVITY). The cross-tabulation function of KHcoder was used.

$$\alpha_j = \sum_{i=1}^N s_{ij} z_i \dots (1), \text{score}_j = \frac{\alpha_j}{M_\alpha} \times R \dots (1)$$

Then, using the reference score  $\alpha_{ij}$ , the reference score for the  $j$ -th person in the  $i$ -th subcategory is expressed in (1).

The basic and final scores are calculated using the model equations in (1) and (2). Since the number of subcategories differs for each evaluation dimension,  $M_\alpha$  is used to align the highest score among the evaluation dimensions. In addition,  $R$  is the highest score for scoring, which in this case was set to 100.



**Figure 4:** Comparison of evaluation scores in LOGICALITY.

## RESULTS

The scores calculated by this system are compared with those evaluated by the human factors experts who conducted the Safety Lesson and are shown in Fig. 4. The coefficients of determination R are 0.61, 0.55, and 0.62 for LOGICALITY, SPECIFICITY, and ADAPTIVITY, respectively, indicating that there is a correlation. The performance of this system was confirmed.

## DISCUSSION

It can be said that the scores evaluated by the human factors experts and the scores calculated by this system generally correspond to each other. The results suggest that it is possible to measure the "Awareness" of employees. In addition, the results of the analysis clearly show the variation in the scores of the 29 participant, which are distributed from high to low scores. Thus, a quantitative evaluation could be made by ascertaining words from the reference word list.

Finally, we demonstrated how the system we have built can extract features not from experimental data, but from actual accident reports. We used three years of potential incident reports from 2019 to 2021 for 31 employees who have been taking Safety Lesson for 10 years. Certain changes could be observed. Relatively reliable results were obtained in actual practice.

## CONCLUSION

In this study, a system was constructed to measure the willingness to participate in safety activities. Through experiments, a correlation was confirmed between the evaluations of the human factor experts and the results measured by this system. On the other hand, the advice to reporters was found to be inadequate in that its content was too abstract and could easily be taken as team-level content, indicating the need for further improvement for practical application. In the future, it is possible that better accuracy could be obtained with various options. For example, the discussion could be extended by the application of the weighting value Z. This time, three characteristics were evaluated as an introduction, but by expanding the number of evaluation items (Employee Rating Group) to 10, it is possible to evaluate and foster "Awareness" about safety that cannot be achieved by a questionnaire. By expanding the number of evaluation items (Employee Rating Group) to 10, it is possible to evaluate and foster "Awareness" regarding safety, which is not possible with a questionnaire.

## REFERENCES

- Ayumu Osawa, Tsubasa Takagi, Miwa Nakanishi, (2020). "An Attempt to Extract Good Jobs from Safety Reports Using Text Mining -Analysis of Voluntary Information Contributory to Enhancement of the Safety (VOICES) Data".
- Erik Hollnagel, David D Woods, Nancy Leveson, (2006). "Resilience Engineering: Concepts and Precepts", Crc Press.
- Haruna Yamazaki, Kana Kimura, Yusaku Okada, (2019). "Evaluation method for safety management characteristic of medical workers".



- 
- Hollnagel E. (1998). *Cognitive Reliability and Error Analysis Method (CREAM)*, Elsevier.
- K Okudaira, C Kojima, S HEMMI, T Takemoto. (2020) “A Study of Identify the Occurrence Factors for Prevention of Seafarer’s Disasters” (in Japanese), *The Journal of Japan Institute of Navigation*, Vol. 142, pp. 45–52.
- Masaomi Kimura, Michiko Ohkura and Fumito Tsuchiya, (2006). “Text Mining Analysis of Medical Near-miss Reports”.
- Masato Kuratani, Jun Sakaniwa, Ken Kusukami and Akinori Komatsubara, (2020). “Development of a Railway Traffic Controllers Competency Extraction Method to Ensure Safety During Restricted Operations”.
- Shimada T., Okada Y., (2019). “A study on the quality of information in potential incident report: *Advances in Intelligent Systems and Computing*”, Volume. 791, pp. 229–239.
- Takiko Imai, Miyuki Takase, Kenichi Satoh, (2018). “Usefulness of Mixed Analysis Method in Combination with Text Mining in Qualitative Data: Analysis of Free-Answer Comments about Turnover Factors of Newly Graduated Nurses”.