*Transactions of the Korean Nuclear Society Spring Meeting*
*Jeju, Korea, May 19-20, 2022*

# Policy Statements and Basic Principles Proposed for Effective Feedback of Human Errors and Related Events in Nuclear

Yong Hee Lee
*Accident Monitoring and Mitigation Dept., Korea Atomic Energy Research Institute*
*898-111 Daeduk-Daero., Yuseong-Gu, Daejeon,*
**Corresponding author: yhlee@kaeri.re.kr*

*Keywords: human error policy, violation, accident investigation, Human Error 3.0, countermeasure*

## 1. Introduction

Various controversies over human errors are growing due to rapid changes in new technologies and up-roaring sensitivity to safety. In the field of high reliability safety, such as nuclear power, the extent to be considered also acts as the biggest cause of threatening security and social acceptance due to controversy and sensitivity. Technical efforts and progresses to prevent human errors in advance are steady, but effective feedback of human error cases occurring in the actual operation is still the most essential process for safety.

This paper discusses further considerations for a more effective feedback of human error-related cases and proposes basic principles and policy declarations. Technical changes and requirements were summarized through previous studies that reviewed the controversies and cases related to human error in nuclear and occupational safety. And a new paradigm of *Human Error 3.0* was discussed to cope with these recent changes and demanding on human error studies. Based on them, a supplemented classification analysis requirement was proposed to more deliberately classify the boundaries related to the first responsibility and criticism. Five basic starting principles and policy statement declarations (drafts) on human error analysis necessary to support this in practice were proposed.

## 2. Discussions on Previous Studies and Approaches to Human Errors and their Countermeasures

### *2.1 Approaches to Human Error and Human Error Events*

In human factors engineering, the characteristics and countermeasures of human errors have been continuously studied [2, 3, 6, 8, 9, 18]. Human errors are also classified into *human error 1.0 and 2.0* reflecting the fundamental conversion of the subject of human error research from human itself to system factors related to human [7]. In coping with human errors, the 3E principle established for the priority of countermeasures has been proposed in the order of *Enforcement, Education and Engineering*. [14, 15]

The most common key task in human error research is retrospection on the causes according to causal relations to the consequences [4]. Although long-standing psychological studies on internal causes as well as the psychological type and cognitive structure of human error have been continued, the results available for engineering applications are not yet insufficient [5, 6, 16]. The fundamental diversity of errors as well as incomplete interpretation of human internal psycho-logical mechanisms may be an inevitable limitation [18].

For engineering perspective, human error is fundamentally far from the internal psychological process of humans. This is because it is realistic that human error is treated as a combination of external influencing factors rather than an internal psychological process. Human error can be defined by external phenomena (*situational characteristics* of human error definition), and controllable details are explained in a multi-level structure of system elements (*chain structure characteristics* of human error) that sometimes had been latent under the surface of external events[7, 8, 10].

Quantitative approaches to human errors have been active under the name Human Reliability Analysis (HRA). There may also be a fundamental mis-understanding of directly estimating the human error probability of workers and their task behaviors. It should be carefully established as a sub-process of probabilistic risk assessment(PRA) in that it is a relative evaluation of human factors related to the job task functions of workers rather than works themselves.

According to the early theorem of safety engineers such as Heinrich, the coincidence law of loss consequences, the causality law of event sequences, the law of eventual preventability, and the priority law of selecting countermeasures are similarly applied to the analysis of human error events. Very similar experiences and observations can only be practically presented. However, a few studies have been discussed that there are the big discrepancies between the basic theorem and the experienced [7, 8, 12] despite the development in human factors engineering. Various new types of human errors such as violations have not been included within the fundamental scope of human error, and treated respectively case-by-case or unconditional safety culture subject to a controlling approach such as regulation and punishment. [1, 2, 3, 5]

In the era of high reliability, a new perspective has been proposed that can include violations such as *Safety II/Resilience and Human Error 3.0* as technical targets [2, 7, 8, 9, 10]. It might be more comprehensive considerations on ultimate human responsibility for all negative consequences of the system according to new paradigms such as *Normal Accidents and Risk Societies*. It is intended to overcome the negative problems caused by new technologies and aspects in advance and achieve a more stable era of high reliability, but it looks still remained conceptual [.

On the other hand, automation that completely excludes humans was pursued as an engineering countermeasure against violation-type human error. This is expected to be visualized with the recent development of robot and artificial intelligence technology. However, as seen in Uber Taxi deaths and B-737 Max accidents, automation is feared to add the possibility of human errors that have never been experienced in areas such as supervision and maintenance, or raise new issues such as security, priority, and control.

### *2.2 Perspectives on Human Errors and Human Error 3.0 Paradigm*

There are many ways to define human error. In cognitive psychology, it is also defined as slip/lapse and mistake or blunder or cognitive under-specification. According to Reason's proposal, which presented this as a brief classification system, it can be defined by type depending on whether there was an internal intention for the result. It is divided into slips that occur in vain due to their own variability without intention and laps that occur due to the limitation of cognitive resources. In addition, human error involving intention is classified into mistake and violation according to the appropriateness of intention. In violation, malicious cases in which intention itself is an attack are sometimes classified again as sabotage.

However, in ergonomics, human error is defined not by an internal phenomenon, but by the negative aspects of its effects and consequences. Human error can be defined by all human behaviors and related factors related to negative endings. Engineeringly, human error can be defined as 'error related to human' rather than 'error of human'. It is intended to deal with the possibility of error in human factors related to errors. Therefore, the notion of *Human Error 3.0* defines all accidents as human resources for the purpose of preparing ultimate responsibility and comprehensive future counter-measures for the negative characteristics (or possibilities) of the ending, even if there are no wrong intentions or defects in humans.

In cognitive psychology, violation is defined as a human error that includes intention, but it is slightly different in reality. When specified as a violation, it is often based only on whether it conforms to the relevant rules (and standards). In other words, most human errors as well as inconsistencies or nonconformities with rules and standards are often concluded as violations. Since there are many cases where the discussion of responsibility and punishment takes precedence over technical understanding simply because of the term violation, more detailed term choices are needed.

Depending on the level of rules, which are the criteria for violations, there are very large variations in working methods and procedures, general restrictions and bylaws, laws, and ethics. Just as terms such as slips, lapses, errors, mistakes, misunderstandings, misconceptions, accidental accidents, and negligence are confusingly used in relation to human error, violations may have different actual meanings. In order to effectively deal with violations in ergonomics, they need to be used through a clearer definition. Therefore, a new perspective on *Human Error 3.0* was proposed for the safety of the large reliability system. Human error 3.0 does not necessarily assume a particular defect or failure. Therefore, it is a proactive view that finds room for additional possible or necessary measures for the safety of the system rather than for the cause of the accident.

- Non-faulty/No-defect normal accident: human factors to human resources
- The ultimate/infinite responsibility premise of a dangerous society: the common destiny of future safety
- Possibility of countermeasures independent to the direct causes: Field safety practicality
- Trust-based human factors safety: mutual trust-based discovery process among stakeholders
- Participatory responsibility to safety: Continuous expansion of safety value through active participation
- Future-oriented proactive safety measures: Focus on future possible creative measures

### *2.3 Categorization of Human Errors including Violations*

Criteria for classifying human error include various methods and criteria such as *behavioral type, cause, result (loss) consequence, job and function, cognitive level, emotional type, internal/external occurrence process, related function, individual and organization*. In addition, it is also possible to classify by general safety information such as *occurrence time, related system/place, target object, responsibility and repetition*. The detailed classification of human error can be fundamentally selected in various ways according to the purpose of dealing with human error and fundamental perspective on human errors. [2, 7, 8, 12, 18] Therefore, it is a very important premise in human error studies that absolute types and objective classification analysis methods do not exist [12]. It may not mandatory to investigate the scientific facts on human error events and human errors. [18] Even in the case of violations, for example, various types have been reported and observed, such as empirically classifying characteristics or classifying cognitive psychological processes and levels and causes. [3,5,6,12]

- Types: Convenience/Individual/Situational/ Permissive Daily Violation, Exceptional, Transcendence, Repeated, Interest Violations
- Ways: Conduct/Active, Intervention/Interruption, Negligence/Observation, By-Standing, Ignorance/Indifference, Repeated/Exceptional, Superior/Distraction, Test(Trial)/Curiosity Interest, Counter-attack, Prohibition
- Causes: Rule defects, situational elements, organizational elements, mis-understanding/lack of knowledge, malicious violation (sabotage), violation for examination, etc.

There are various attempts to classify the types of violations by cause as well as by the revealed type, but they are generally not effective nor sufficient in practice compared to the classification by psychological level proposed by Reason, etc.[1]. This is because it is difficult to determine the incontrovertible cause like other human errors. Therefore, it is effective to subdivide the types of violations into possible countermeasures by characterizing the detailed type based on detailed control factors affecting them.

According to a cognitive detailed analysis of serious disaster (death)-related violations [16, 17], it was found that few types of violations are frequent such as In-sufficiently-Recognized and Risk-taking-by-inertia even after recognition. In addition, in the basic survey of the types of violations observed and experienced, it was analyzed that there was no difference in the types and priority of violations along with traffic safety, every-day and job-related [15].

## 3. Suggestions for More Effective Feedback of Human Errors and their Experiences

In order to effectively deal with human errors and related events, a supplementary approach focused to counter-measure which may be a more practical target, was proposed rather than a cause-oriented approach. The following section 3.1 summarizes an improved analysis method as an example of classifying with a simple *haddon matrix* [13, 16, 17], and taxonomy of countermeasures proposed for investigation of human errors.

And five basic principles and policy statements have been proposed as a new direction for effectively dealing with human errors through the proposed approach by applying a new safety paradigm of *Human Error 3.0*.

### *3.1 Overcome the Limit of Causal Investigation and Backward Reasoning for Human Error Studies*

The analysis of the cause of human error is instinctive. The demand for objective causes is based on realistic needs along with curiosity. However, it seems to be an important theorem related to human error that objective causes with clear scientific validity do not exist or are not possible to grasp [18]. In addition, it seems that the argument that effective measures should take precedence over objective causes is practically recommendable and valid for the engineering purpose of human error analysis [12,13,14].

Most of the causes of human errors are derived through backward reasoning analysis from the results. In particular, elements related to this are selected for the negative aspects of the results, and it is unclear whether relevance necessarily means absolute causality. Reverse inference is based only on the probability of causality and is not exclusive and objective, so the absoluteness of the results is not guaranteed. Nevertheless, human factors related to negative results are judged to be causal factors (i.e., causes) of the results.

In the analysis of actual accidents, most of the human factors of certain contents and levels can be considered. Although this may be only a behind-sight, it is frequently highlighted as a cause in causal cause analysis due to a confirmation bias error. Human factors related to accidents are regarded as human errors of related persons and are considered to be responsible at the same time. In particular, when various types of standards exist, human error can be immediately considered a violation. This is because due to the uncertainty of the term and concept of violation, even implied matters that can be generally expected from obligations specified and promised by law can be adopted as criteria for violation.

Discussing compensation, disciplinary action, and punishment through analysis of the responsibility perspective for violations is essential for specific purposes such as judicial administration. However, as a result, the technical access to violations as a human error was blocked through parallel analysis. To this end, a multi-layered analysis system capable of multi-perspective and multi-purpose analysis was proposed [13]. It is a plan to conduct analysis in parallel to prepare technical countermeasures against violation human error through an engineering approach. A three-dimensional multi-layered analysis scheme was illustrated from the perspective of function, behavior, and responsibility applied to the re-analysis of violation-type human error [10, 13, 14, 16].


### *3.2 Starting Basic Principles for Human Error Studies and Policy Statements Proposed*

Violations frequently demand direct blaming to the intention before the any investigation on technical enhancements. Proposed method is an example approach to maintain plausible technical investigation in parallel, at least, with blaming interrogation. For the practical implementation of any enhanced approach to human error studies, especially investigation on human error events including violations, can be supported by a set of basic principles and policy on human errors. A set of starting principles revised from previous study [15] were suggested as followings to cope with the practical demanding during human error investigation process, especially for considering effective feedback to violations rather technically.

- *Principle of Unintentional and Good Faith*: The basic assumption in the investigative analysis of an event is that the worker (or interest groups) is good faith and not malicious in the detailed functions and roles of the event concerned. It is assumed that the worker tried to do his/her best to the function and assigned role expected at the time of job/task performance. Even if it is a violation, it is assumed that there is at least no malice (ill-intention). Therefore, sabotage initiated with malicious intent is not the subject of technical analysis. If a technical approach to a malic is required, detailed analysis treats the parties' intentions separately through a separate confirmation logic.
- *Principles of Controllable Evidence*: A comprehensive list to identify all relevant influencing factors and identify root causes to identify factors that caused human error. At this time, in order to determine the causal factors, their controllability by the worker must be firmly verified. If the worker's own control is not likely or insufficient, other evidence should be adopted.
- *Principle of Countermeasure Independence to Cause*: Countermeasures of an event can be independent to causes. Eliminating the cause may be chosen as one of the important candidates for countermeasures, but it is not the only counter-measure. Countermeasures can be constructed regardless of the cause, and can be developed in various ways through adopting new emerging technologies and creative proposals. In particular, regardless of the cause, the regret from the hind-sight might be an important starting point for development of effective countermeasures.
- *Principle of Practical Effectiveness over Causality*: Measures should be practical and chosen in two dimensions. It is a relative review of the resources and efforts required to implement the measures and the results of the measures. This refers to the process of frequent cost-effectiveness analysis in engineering decisions, but it should be a safety decision and not be based on simple efficiency.
- *Principles of Liability Limitation and Proportionality*: In order to be judged to have been violated, it is necessary to go through the test logic of the constituent requirements of the violation. However, even if the violation turns out to be true, its responsibility still requires additional judgment steps. Responsibility cannot be defined beyond the limits of ergonomic capabilities as well as the scope of authority granted. The size of responsibility cannot exceed the effective size of authority assigned to the worker. As recently raised in the issue of 'Organized Irresponsibility', the legitimate responsibility related to the violation, even if the assigned responsibility is agreed, shall not exceed the scope of the competence of the parties and the authorities granted. In the case of multiple party responsibilities, personal responsibility should be established in proportion to the relative size of the authority.

Above set of five principles is proposed to change the current basis on human error investigation including violations. It may help to start more focused investigation on technical aspects of human error events, since it leads to more practical results with countermeasures rather than causes and their responsibilities. It may include conflicting descriptions on human error itself.

In addition, a policy declaration linked to safety and health management system or safety culture is needed as a comprehensive basis for the enhanced approach. Followings may the example set of the proposed human error policy declarations (draft) for the nuclear field, since we may appreciate more countermeasures rather than causes from the human errors investigations.

- Nuclear workers and stakeholders may always commit human error despite their utmost efforts to ensure nuclear safety.
- Human error causes may come inadvertently from design, facilities and settings, underlying human characteristics and their systematic limitations, and unknown elements that are yet not identified.
- The scope and importance of human errors are relatively increased with the development of technology, the maturity of the system as well as the safety. It may change with different perceptions of stakeholders, so continuous feedback and different kinds of efforts are indispensable.
- Human error is an independent event with a new perspective rather than just one cause of incident, so stick to the basic laws of safety management and take practical best measures based on the latest technologies such as ergonomics.
- Human error has the best value of preventing future risks or minimizing inevitable damage and losses by applying different perspectives on past risks to interpret and cooperate independently.

## 4. Conclusions and Discussions

This paper discussed ways to more effectively feedback the experience of human error-related events, especially including violations. Technical changes and analysis requirements for human errors including violations, were summarized through previous studies that reviewed the controversies and cases related to human error and suggested *Human Error 3.0* paradigm.

It was discussed that the ultimate purpose of human error feedback is not limited to identify the causes or responsibility, but to further extended to secure the ultimate safety by finding out plausible countermeasures. To this end, there is a problem of responsibility and blaming criticism that comes first. A supplemented human error classification analysis scheme was proposed as an example to more delicately distinguish the boundaries related to responsibility blaming. And a set of supplementary requirements for investigation and feedback was proposed based on new paradigm. The more practical considerations to support human error feedback were reviewed through previous studies. Five starting principles and policy statements on human error analysis (draft) is suggested to be declared in nuclear.

Further works might be expected to implement the requirements proposed in this paper for a more effective human error feedback in nuclear. They may include new design basis and guidelines for autonomous features such as AI as well as a new event investigation scheme and classification taxonomy on human errors and events.

## ACKNOWLEDGMENT

This paper is supported by the Nuclear Safety Research Program grant funded by Nuclear Security and Safety Commission (NSSC) and KOFONS (No. 2003010).

## REFERENCES

[1] Govaarts, C., Establishment of Just Culture Principles in ATM Safety Data Reporting &Assessment, Eurocontrol, 2006
[2] Hollnagel, E., Safety I & Safety II, Elsevier. Ashgate, 2009
[3] Hudson, P. et al., Bending the Rules: Managing Violation in the Workplace, Society of Petroleum Eng. Int. Conf. on Health, Safety & Env. in Oil & Gas Exploration, 1998
[4] INPO, Human Performance Evaluation System, 1987
[5] Jung, Y.H. et al., Current Status of Just Culture in Aviation and Its Applications to Nuclear Power Plants (in Korean), KINS/RR-2011, 2020
[6] Kang, B. et. al., Conceptual Models of Violation Errors in a NPP, *J. Korean Society of Safety*, 31(1), pp.126-131, 2016
[7] Lee, Y.H., Human Error 3.0 Concept for High Reliability Era, Proc. ESK-2015-Fall, 2015
[8] Lee, Y.H., Human Factors Engineering Approach to Safety Culture, Proc. ESK-2016 Fall, 2016
[9] Lee, Y.H., A Revisit to Technical Issues & Approaches for Human Error Event Investigations, ESK-2019-Fall, 2019
[10] Lee, Y. H., How to Treat Violation Errors during Human Error Investigations in Nuclear, KNS2019-Autumn, 2019.
[11] Lee, Y.H., A Preliminary Study on the Culpability of Violation Errors in Nuclear Events and their Investigations, KNS-2020 Autumn, 2020
[12] Lee, Y.H., Discussions on the Post Human Error Activities for 21-st Century - Focused to Effective Coping with Violations, ESK-2021 Spring, 2021
[13] Lee, Y.H., A Scrutinized Step Proposed to More Effective Human Error Investigations for including Violations and their Countermeasures, KNS-2021 Spring, 2021
[14] Lee, Y.H., A Preliminary Survey on the Counter-measures to Violation Errors for the Establishment of Human Error Policy in Nuclear, ESK-2021 Fall, 2021
[15] Lee, Y.H., Development of Starting Principles and a Culpability Logic Suggested for more Effective Analysis on Violations in Nuclear Events, KNS-2021 Autumn, 2021
[16] Lim, H.K. et al., Reanalysis of Fatal Industrial Accident Cases from the Viewpoint of Violation Type Human Errors, (NSTAR-20NS41-135), 2020.
[17] Lim, H.K., et al., Development of An Objective Judgement Procedure for Determining Involvement of Violation-Type Unsafe Acts Causing Industrial Accidents, *J. Korean Society of Safety*, vol. 39, 2021
[18] Rasmussen, J., Concept of Human Error: Is it Useful for the Design of Safety Systems? *Safety Science Monitor*, 3(1), 1999