Advances in Intelligent Systems and Computing 589

Ronald Laurids Boring *Editor*

Advances in Human Error, Reliability, Resilience, and Performance

Proceedings of the AHFE 2017 International Conference on Human Error, Reliability, Resilience, and Performance, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA

Springer

Advances in Intelligent Systems and Computing

Volume 589

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Chairman

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in

Members

Rafael Bello Perez, Universidad Central "Marta Abreu" de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu

Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchado@usal.es

Hani Hagras, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk

László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu

Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu

Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlin@mail.nctu.edu.tw

Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au

Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org

Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br

Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl

Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

More information about this series at http://www.springer.com/series/11156

Ronald Laurids Boring
Editor

Advances in Human Error, Reliability, Resilience, and Performance

Proceedings of the AHFE 2017 International Conference on Human Error, Reliability, Resilience, and Performance, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA

Springer

Editor
Ronald Laurids Boring
Idaho National Laboratory
Idaho Falls, ID
USA

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-60644-6 ISBN 978-3-319-60645-3 (eBook)
DOI 10.1007/978-3-319-60645-3

Library of Congress Control Number: 2017943026

© Springer International Publishing AG 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Advances in Human Factors and Ergonomics 2017

AHFE 2017 Series Editors

Tareq Z. Ahram, Florida, USA
Waldemar Karwowski, Florida, USA

8th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences

Proceedings of the AHFE 2017 International Conference on Human Error, Reliability, Resilience, and Performance, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA

Advances in Affective and Pleasurable Design	*WonJoon Chung and Cliff (Sungsoo) Shin*
Advances in Neuroergonomics and Cognitive Engineering	*Carryl Baldwin*
Advances in Design for Inclusion	*Giuseppe Di Bucchianico and Pete Kercher*
Advances in Ergonomics in Design	*Francisco Rebelo and Marcelo Soares*
Advances in Human Error, Reliability, Resilience, and Performance	*Ronald L. Boring*
Advances in Human Factors and Ergonomics in Healthcare and Medical Devices	*Vincent G. Duffy and Nancy Lightner*
Advances in Human Factors in Simulation and Modeling	*Daniel N. Cassenti*
Advances in Human Factors and System Interactions	*Isabel L. Nunes*
Advances in Human Factors in Cybersecurity	*Denise Nicholson*
Advances in Human Factors, Business Management and Leadership	*Jussi Kantola, Tibor Barath and Salman Nazir*
Advances in Human Factors in Robots and Unmanned Systems	*Jessie Chen*
Advances in Human Factors in Training, Education, and Learning Sciences	*Terence Andre*
Advances in Human Aspects of Transportation	*Neville A. Stanton*

(continued)

(continued)

Advances in Human Factors, Software, and Systems Engineering	*Tareq Z. Ahram and Waldemar Karwowski*
Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries	*Paul Fechtelkotter and Michael Legatt*
Advances in Human Factors, Sustainable Urban Planning and Infrastructure	*Jerzy Charytonowicz*
Advances in the Human Side of Service Engineering	*Louis E. Freund and Wojciech Cellary*
Advances in Physical Ergonomics and Human Factors	*Ravindra Goonetilleke and Waldemar Karwowski*
Advances in Human Factors in Sports, Injury Prevention and Outdoor Recreation	*Tareq Z. Ahram*
Advances in Safety Management and Human Factors	*Pedro Arezes*
Advances in Social & Occupational Ergonomics	*Richard Goossens*
Advances in Ergonomics of Manufacturing: Managing the Enterprise of the Future	*Stefan Trzcielinski*
Advances in Usability and User Experience	*Tareq Ahram and Christianne Falcão*
Advances in Human Factors in Wearable Technologies and Game Design	*Tareq Ahram and Christianne Falcão*
Advances in Communication of Design	*Amic G. Ho*
Advances in Cross-Cultural Decision Making	*Mark Hoffman*

Preface

To err is human, and human error is consistently implicated as a significant factor in safety incidents and accidents. Yet, as pervasive and important as human error is, the study of human error has been fragmented into many different fields. In fact, in many of these fields, the term "human error" is considered negative, and terms such as human variability or human failure are preferred. Across differences in terminology and approach, the common link remains an interest in how, why, and when humans make incorrect decisions or commit incorrect actions. Human error often has significant consequences, and a variety of approaches have emerged to identify, prevent, or mitigate it. These different approaches will find a unified home in this conference.

The 1st International Conference on Human Error, Reliability, Resilience, and Performance (HERRP) took place at The Westin Bonaventure Hotel, Los Angeles, California, USA, from July 17 to 21, 2017. The conference was organized within the framework of the International Conference on Applied Human Factors and Ergonomics (AHFE) as an Affiliated Conference. The purpose of the conference is to bring together researchers and practitioners in different fields who broadly share the study of human error. The HERRP conference is intended to serve as an umbrella for human error topics by providing an annual forum for otherwise disjoint research efforts. As such, the conference is intended to complement but not replace existing specialized forums on particular facets of human error. The HERRP conference is distinctly interdisciplinary, encouraging the submission of papers in focused technical domains that would benefit from interaction with a wide human factors audience. Additionally, the HERRP conference aims to provide a yearly, high-quality, archival collection of papers that may be readily accessed by the current and future research and practitioner community.

The HERRP scientific advisory board invited papers related to a broad range of topics on human error, including but not limited to:

Human performance
Human variability
Human reliability analysis

Human performance shaping factors
Root cause analysis
Accident investigation
Human resilience and resilience engineering
High reliability organizations
Safety management
Medical error
Driver error
Pilot error
Automation error
Defense in depth
Errors of commission and omission
Human error taxonomies and databases
Human performance improvement and training
Cognitive modeling of human error
Qualitative and quantitative risk assessment

Many of these topics and others are reflected in these proceedings. Contributions encompassed empirical research studies, original reviews, practical case studies, meta-analyses, technical guidelines, best practices, or methods. Papers encompassed traditional topics of human error such as found in the safety critical industries like energy, manufacturing, and medicine. We also encouraged innovative explorations of human error such as security, defense, new human–technology interactions, and beneficial uses of human error.

The sections of these proceedings are grouped as follows:

Part I *Human Reliability Analysis*
Part II *Bridging Human Factors and Human Error*
Part III *Human Error and Resilience in Safety-Critical Domains*
Part IV *Evaluating Expert Performance*
Part V *Advanced Analysis Techniques for Human Error*
Part VI *Human Reliability and Human Factors Research*

It has in my view been a very successful first installment of the HERRP conference, and I look forward to watching the evolution of this conference. I am grateful to the organizing committee of the 8th International Conference on Applied Human Factors and Ergonomics and the affiliated conferences for making this embedded conference possible. I also wish to thank the authors for their exceptional contributions to the conference and to scientific advisory board for encouraging strong submissions:

Harold Blackman, USA
Y.J. (James) Chang, USA
David Gertman, USA
Katrina Groth, USA
Xuhong He, Sweden
Stacey Hendrickson, USA

Yochan Kim, Korea
Barry Kirwan, France
Karin Laumann, Norway
Zhizhong Li, China
Peng Liu, China
Ron McLeod, UK
Naj Meshkati, USA
Aino Obenius-Mowitz, Sweden
Jinkyun Park, Korea
Manikam Pillay, Australia
Alice Salway, Canada
Carol Smidts, USA
Oliver Straeter, Germany
Claire Taylor, Norway
Patricia Trbovich, Canada
Matt Weinger, USA
A. Whaley, USA

It is my hope that these proceedings will prove a valuable and regular addition to the broad human error literature.

July 2017 Ronald Laurids Boring

Contents

Human Reliability Analysis

Bridging Human Factors and Human Error

Human Error and Resilience in Safety-Critical Domains

Evaluating Expert Performance

Advanced Analysis Techniques for Human Error

Human Reliability and Human Factors Research

Human Reliability Analysis

Evaluation and Consolidation of the HEART Human Reliability Assessment Principles

Julie L. Bell[1(✉)] and Jeremy C. Williams[2]

[1] Health and Safety Executive, Harpur Hill, Buxton SK17 9JN, UK
julie.bell@hse.gov.uk
[2] Jerry Williams Human Factors Ltd., Hale Barns WA15 0BW, UK

Abstract. Since its publication over 30 years ago, the Human Error Assessment and Reduction Technique (HEART) has been used as a risk assessment, accident investigation and design tool. The human factors literature has grown considerably since the method was first published and this has led to some concern about the age and relevance of the data that underpin HRA methods. The objective of the current research was to critically evaluate the data and evidence base underpinning HEART. This was achieved by analyzing research published between 1984 and 2015 to compare with HEART's existing Generic Task Types (GTTs), Error Producing Conditions (EPCs) and Human Error Probability (HEP) distributions. This scientific/knowledge based contribution provides confidence that the basic structure of the HEART method is well-supported by the literature of the last 30 years and shows that only minor adjustments are required to refine, support and extend its continued successful application.

Keywords: HEART+ · Human error · HRA · State-of-the-Art

1 Introduction

The Human Error Assessment and Reduction Technique (HEART) was first proposed in 1985 [1]. It is a general-purpose, pre-processed human reliability assessment technique developed to help risk analysts and designers identify the major influences on human performance and the likelihood of error, in a systematic and repeatable way. It is based on the general principle that for each task in life there is a basic probability of failure that can be assessed within probabilistic limits. Affecting each of these tasks are varying levels of Error Producing Conditions (EPCs) that can influence human reliability in systems operations. HEART is a relatively quick, additive factors method that is straightforward to use and is applicable to any industry where human reliability is important.

Nine generic task types (GTTs) and proposed nominal human unreliability values with suggested bounding values, together with thirty-eight EPCs were identified in the original method. Each EPC had an associated multiplier, which is the maximum amount by which the nominal unreliability might change depending on the proportion of the EPC applicable at the time of task execution. A method for computing the assessed likelihood of error and its bounding values is provided.

© Her Majesty the Queen in Right of United Kingdom 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_1

HEART has been in use for 30 years providing insight into the impact of human error in the healthcare, rail, aviation, nuclear, process, chemical and offshore engineering industries. The published literature includes examples of HEART being used to inform error potential associated with data entry in radiotherapy treatment [2], rail vehicle axle inspection [3], the transport of radioactive material [4] and performing a corrective task in a control room in order to prevent a hazardous event [5].

In addition, several validation studies have been undertaken [e.g. 6]. These have concluded that HEART enables rapid assessments that broadly correlate with known data and produce in sights that are equivalent to those derived by using more complicated methods.

Over the years, concerns have arisen about how to interpret some of the method specific terminology and, as time has elapsed, about the age of data underpinning established HRA methods, including HEART. This paper describes the work to consolidate and add to the data that underpin HEART by drawing on the last 30 years' Human Factors Literature, and to offer greater insight into the application of the method. In this paper, we report the results of work to consider the concept of EPCs, the appropriateness of each, and their associated multiplier. It provides insight to the GTTs, the characterization of Human Error Probability (HEP), HEP distributions and the underpinning additive factors method.

2 Methodology

2.1 Assessing the Error Producing Conditions

The aim of this element of the research was to aggregate studies that used conceptually similar EPCs in order to generate an estimate of their impact on human reliability. This involved an extensive literature review to identify studies that had a discernible EPC as factor in the experimental design.

The papers were scrutinized to establish if the experimental sample size was at least 10 per study condition and that the participants were representative of the working population. Importantly, the results had to be reported as error data. Studies using reaction time and other measures of performance were not included in the quantification because they require judgment and interpretation and therefore, introduce more uncertainty into the analysis. These strict criteria significantly reduced the number of papers that could be included in the research, but this degree of rigor provides greater confidence in the results.

The final sample of papers was generated by searching through some 35,000 published papers. The publication sources were Human Factors (~2,500 papers), Proceedings of the Annual Meetings of the Human Factors and Ergonomics Society (~8,000 papers), Ergonomics (~4,000 papers), Applied Ergonomics (~1,000 abstracts), Journal of Experimental Psychology: Learning, Memory & Cognition (~2,000 papers), Attention, Perception, & Psychophysics (~4,500 papers), Psychonomic Bulletin & Review (~2,500), Behaviour and Information Technology (~2,500 abstracts), and some 8,000 Internet, cognitive neuroscience, experimental and applied psychology sources.

Some 180 papers met the strict evidentiary selection criterion. Data points were extracted from each paper and collated into spreadsheets for each EPC. The proposed maximum multiplier for each EPC was derived by considering the data cluster and making an estimate of the worst case. The process involved taking proper account of apparently spurious data points and potential outliers, to form a justifiable position. This included, for example, making contact with authors to ask for clarification.

2.2 Assessing Generic Task Type Data

As part of the literature search for EPC data, evidence was also collected that related to GTTs (current and potential new ones), the verification, or otherwise, of the assumed log-normal distribution of the human error rates and the additive nature of how assessed proportions of EPC elements combine. Initially, 170 papers were identified as having potential to add to the GTT knowledge base. These papers also provided information about HEP distributions and additive factors but these findings are the subject of other papers [e.g. 7].

In order to check the validity of the GTT concept, papers that potentially contained meaningful task descriptions and HEPs were collated in a spreadsheet. The authors independently aligned the experimental task descriptions with GTTs using the method outlined in the HEART User Manual [8]. In most cases, both authors matched in their determination of the GTTs as identified from the task descriptions. The authors could not reach agreement regarding a suitable GTT in two experiments and so these were discarded, another four were discarded because experimenters had used a “staircasing”, error tracking algorithm or “titration” method in order to produce predictable error probabilities. In three other cases, experimenters had reported only error proportions rather than error probabilities, so these experiments were also discarded.

Once the 151 GTTs had been agreed, error probability data were extracted from each experiment. Only one data point was taken from each experiment to ensure that no experiment was over-represented and to provide a high degree of diversity of data sources.

3 Results

The analysis confirmed that 32 (out of 38) HEART original EPC concepts and multipliers appear fit for purpose [10]. In each case, there is either new supporting evidence for, or there is no new information to challenge, the original conclusions. There was sufficient evidence to recommend that six EPCs be revised slightly and for two new ones to be incorporated into HEART.

3.1 Revised EPCs

Six of the original 38 EPCs were modified based on the results of the review. The following sections provide a summary of the basis for the authors’ conclusions.

EPC 29. High Level Emotional Stress. Although this factor is frequently referred to in the human reliability assessment literature as being a major factor in some assessments, only a few quantitative studies were identified from the last 30 years. The studies applied a range of 'stressors' and the results showed a difference in performance of approximately 2, between stressed and non-stressed individuals. For example, a paper by Leon and Revelle [9] found that more anxious subjects had a higher mean error rate that was 2.2 times greater than less anxious subjects. Other studies reported findings in the range of 1.02 [11] through to 2 [12]. Overall, these figures suggest an increase in the multiplier from 1.3 to a factor of 2 would be prudent, until further evidence becomes available.

EPC 32. Inconsistency of Meaning of Displays and Procedures. Considering the importance of consistency in product and system design, and although much has been written, very few relevant human reliability data have been published over the last 30 years. In the literature, the emphasis is on automated techniques to detect and minimize inconsistency of human action. The few studies that report user error result in factor increases that are somewhat higher than was originally proposed. For example, Mendel [13] reported errors that were increased by a factor of 2.2 and concluded that consistency may be especially important for difficult tasks with high cognitive load. Overall, these studies suggest that a multiplier of 3 is more appropriate than the original 1.2 but this EPC would benefit from further investigation.

EPC 35. Disruption of Normal Work-Sleep Cycles. With few exceptions, most studies report that fatigue induced by sleep loss has been found to reduce human reliability by an amount proportionate to the amount of sleep lost. On average, although there is up to 5% human reliability improvement every afternoon, overall human reliability has been found to reduce by approximately 20% for every 24 h of sleep deprivation [14]. Based on a considerable amount of research over the last 30 years [e.g. 15], it is proposed that the multiplier be increased to × 1.2 (compound) for every 24 h' sleep lost to at least 72 h without sleep.

EPC 33. A Poor or Hostile Environment (Below 75% of Health or Life-Threatening Severity). Since the original publication of HEART, changes in regulation [16] have made it less likely that individuals in Europe will be exposed to very harsh environments. Despite this, there has been research over the last 30 years and studies suggest that whilst the general magnitude of the multiplier might be in roughly the right region, an increase from 1.15 to 2 is appropriate. Mäkinen et al. [17], for example, studied performance over a ten day period at a 25 degree ambient and a 10 degree Celsius experimental condition and reported significant changes in accuracy; the multipliers that emerge from this is study were 1.9 and 1.4 respectively.

EPC 37. Additional Team Members over and above Those Necessary to Perform Task Normally and Satisfactorily. Two studies informed the original multiplier of 1.03 per additional person. From the last 30 years' literature, only one study [18] was found that is relevant to this EPC. This study on testing individuals separately and in pairs showed that the likelihood of driving test failure increased by a factor of 1.29 when an extra person was present. Bearing in mind the original multiplier was derived from two studies and that only one from the last 30 years has been found, it is prudent to retain this EPC and revise the multiplier to 1.2. This EPC does probably affect human reliability and it might be better informed by further research in due course.

EPC 38. Age of Personnel Performing Recall, Recognition and Detection Tasks (Ages 25 to 85 Years and Mentally Competent). A considerable amount of research exploring the impact of age on performance has been published in the last 30 years. Data from the some 15 papers [e.g. 19–22], indicate a consistent reduction in reliability that is associated with age and this change is relatively uniform across a wide range of ages. It is likely that personnel involved in performing detection, recall and recognition tasks will be 15–20% less reliable for every ten years age and that, within very narrow limits, and this impact will be consistent over the range 25 to 85 years. While it might be argued that experience can 'eliminate' the impact of age on performance, Nunes and Kramer [23] report that, "high levels of experience do not broadly provide immunity against the detrimental effects of advancing age on basic cognitive functioning". Based on this information, this EPC should be adjusted upward from 1.02 to 1.16 for every 10 years for ages 25 to 85 years (assuming the individuals are not suffering from dementia).

3.2 New Error Producing Conditions

In addition to those EPCs that have been revised, the research identified two EPCs that can now be quantified.

- *Distraction/Task Interruption.* Distraction can be defined in a variety of ways, but the phenomenon is generally agreed to be about the drawing of attention away from something with the simultaneous possibility of confusion and increased potential for current task failure. Experiments that have been performed to investigate distraction have tended to be of two basic types, those that look at differences between interruptions and those that look at what happens to performance when attention is switched. Williams et al. [14] initially reviewed these two types of experiments as two distinct types of EPC, but it quickly became apparent that the impact on human reliability is largely the same, regardless of experimental type or method. Williams et al. concluded the impact of distraction ranges from about a factor of two up to a factor of four. For the purposes of assessment, it is suggested that the upper limit be regarded as a factor of four.
- *Time-of-Day (from Diurnal High Arousal to Diurnal Low Arousal).* Provisional investigations and assessment of the impact of time-of-day on human reliability suggest that the variation between the diurnal high and low of performance is about a factor of 2.4. The highest reliability is identified as being around 16:00 h and the lowest reliability, somewhat unsurprisingly, at around 03:00 h. The "neutral" point is at around 20:00 h. There is some evidence of a "post lunch dip" that contributes about a factor of 1.3 to overall human unreliability specifically at around 14:00–15:00 h. The data analyzed and consolidated come from a range of sources, including laboratory, field and accident studies for example Mittler et al. [24] and Folkard [25]. Sufficient is now known that a new EPC and associated multiplier of 2.4 will be included in the update of HEART, in order to capture its potential contribution to human failure with respect to the time-of-day.

When HEART was first developed, it was recognized that there were gaps in our knowledge and it was anticipated that more research would be published in due course as the need to quantify such effects became apparent. Evidently, the EPC literature is now much larger and more accessible than in the mid ‘80s, but our research shows there are substantial gaps in knowledge and areas where research with quantifiable results is lacking. Despite the lack of quantifiable data, the research does show that significant performance patterns are apparent in the human factors literature. These patterns provide assurance that the EPCs outlined in HEART are valid and remain relevant for managing human performance in modern industry. If there are questions regarding the EPCs they are not about whether they have an impact on performance, or the relative impact of one EPC to another, but rather the size of that impact.

While 32 of the 38 EPCs are considered fit for purpose and are accepted to have an impact on human reliability, it would be beneficial if more specifically quantifiable research were undertaken. The update of HEART will include all of these EPCs with their original multiplier because they are of relevance and some increasingly so, in contemporary life. For example, EPC 18 “A conflict of short term and long term objectives” has not been researched in a systematic way, yet the demands to achieve a ‘quick fix’ while putting off the main problem until a later date is a recognizable issue in commercial work. EPC 28 “Little or no intrinsic meaning in a task” needs research because some jobs are becoming increasingly boring due to technological advances that reduce the need for human resourcefulness and action. Currently, the Human Factors community does not have sufficient quantitative information to be informed about this possible impact on human reliability. Stress is a cause of significant health issues, but while researchers know a great deal about the impact of stress on health the lack of quantitative data means it is not measurable in an operational sense for risk assessments. In addition, as the production burden on industry increases, employees are being required to respond to increasingly pressurized situations. The body of qualitative research could suggest that EPC 29 “High level emotional stress” needs a higher multiplier, but how much higher is not evident. Similarly, research into a lack of exercise (EPC 22) suggests that whilst this may be important, although perhaps in the noise level, we do not know whether it affects human reliability, and, if so, by how much. Research in this area tends to focus on performance improvement with exercise rather than any detrimental impact due to inactivity.

The limitations of the available data notwithstanding, the basic structure and underpinning data of the HEART method are well-supported by the literature and only minor adjustments are required to refine, support and extend its continued successful application. There is considerable strength in the HEART method because it draws together information from a vast range of studies for each EPC and this means they can be applied confidently to a variety of situations that are not task, context or person specific.

3.3 Generic Task Type Findings

Tasks were classified using the HEART GTT descriptions and details associated with each experiment such as the number of trials and subjects were recorded on a spreadsheet. The spreadsheet was then interrogated by sorting against task type and characterizing each group of tasks by amalgamating the data from each experimental data point. To facilitate processing error probabilities were log-transformed and these data were then combined in order to derive summary statistics for each HEART GTT. The results are shown in Table 1.

Table 1. HEART generic task types and HEPs

GTT	Sources	GeoMean	5th%ile	95th%ile
A (1985)	1	0.55	0.35	0.97
A (2015)	7	0.41	0.19	0.85
B (1985)	2	0.26	0.14	0.42
B (2015)	15	0.17	0.02	1.0[a]
C (1985)	4	0.16	0.12	0.28
C (2015)	37	0.17	0.05	0.6
D (1985)	1	0.09	0.06	0.13
D (2015)	54	0.06	0.02	0.19
E (1985)	8	0.02	0.007	0.045
E (2015)	46	0.02	0.005	0.09
F (1985)	2	0.003	0.0008	0.007
F (2015)	6	0.001	0.00002	0.04
G (1985)	4	0.0004	0.00008	0.009[b]
G (2015)	8	0.002	0.0002	0.01
H (1985)	1	0.00002	0.000006	0.0009
H (2015)	2	0.00004	–	–

(1985) proposed nominal values in Williams (1985).
(2015) values derived from the last 30 years.
[a] These data may not be lognormally distributed.
[b] 0.007 in User Manual (Williams, 1992).

The correlation of all 175 data pairs is 0.87, which is significant at the 0.00001 level (Pearson 2-tailed). The correlation of each nominal HEART GTT against the geometric means derived for each task is 0.98, which is significant at the 0.00003 level (Pearson 2-tailed) and the correlation of the limit values from the method against the derived limit values is 0.94, which is significant at the 0.000001 level (Pearson 2-tailed).

These findings confirm that the nominal and limit values proposed in the original model are similar to the values derived from the literature of the last 30 years.

4 Discussion

It will be apparent that the majority of the originally proposed EPCs appear to be supported by the literature of the last thirty years, not only conceptually, but, in many cases, in a quantitative sense. Where the quantification of EPC impacts would benefit from amendment, the changes that the literature suggests might be merited are relatively slight and, in all cases, upwards.

For GTTs a similar story emerges, except that the amount of change to nominal values justified by the literature is minimal, but it, too, appears to be in an upwards direction for low values. The nominal values for GTTs "C", "D", "E" and "F" appear to be well-supported by the experimental literature of the last 30 years. However, it is notable that some GTT data that are urgently needed by the HRA community are currently unavailable in any sizeable quantity.

It would be completely feasible to collect the necessary data using experiments and observations de-signed specifically to furnish these data needs. For example, it should be noted that the 37 studies that provided data for "C" Task consolidation collectively required of the order of 440,000 experimental trials. For "D" task data consolidation, some 590,000 trials were required in 54 studies and, for the 47 studies that underpin the "E" Task HEP, some 510,000 trials were performed.

Assuming the GTT concept is generally valid, it can be seen that some of the mid-range GTT values line up fairly well and others, somewhat less so. Unsurprisingly, we are short of data in the high/low ends of the error probability spectrum – the "A", "B", "F", "G" and "H" sectors. Bearing in mind the fact that each mid-range GTT now has the support of 500,000 observations underpinning each estimate, it would be a relatively simple matter to direct future Human Factors and Psychological HEP research to elicit data that will inform the "F" and "G" GTT HEPs. For "A" and "B" GTTs, the problem is not to do with the low overall number of observations, but a lack of experimentation that is directed at these types of task that will require substantial numbers of "one trial" subjects. The data shortage problem consists of two very particular shortages, therefore, which can be described, pinpointed and addressed.

5 Conclusion

The body of research over the last 30 years has made it possible to consolidate HEART. This paper has summarized the findings of a significant literature search to confirm, revise and refine some of the EPC multipliers, to identify new EPCs and provide further evidence for GTT quantification. While it has highlighted some research needs to establish better estimates of some EPC multipliers and GTT HEPs, it provides substantial evidence that the method is fit for continued use as a general-purpose, quick, effective, valid and practical data-based HRA method. The knowledge from this research has been used to consolidate and update the HEART method. To reflect this extension of our knowledge, the expanded method will be referred to as HEART+.

References

1. Williams, J.C.: HEART- A proposed method for achieving high reliability in process operation by means of human factors engineering technology. In: Proceedings of a Symposium on the Achievement of Reliability in Operating Plant, Safety and Reliability Society, vol. 16, pp. 5/1–5/15 (1985)

2. Chadwick, L., Fallon, E.F.: Human reliability assessment of a critical nursing task in a radio-therapy treatment process. Appl. Ergon. **43**(1), 89–97 (2012)
3. Gibson, W.H.: Research into the Effects of Human Factors in Axle Inspection – Scoping Study Report. Rail Safety and Standards Board, London (2012)
4. IAEA: Input data for quantifying risks associated with the transport of radioactive material. International Atomic Energy Agency, IAEA-TECDOC-1346, Vienna (2003)
5. Nespoli, C., Ditali, S.: Human error probability estimation for process risk assessment with emphasis on control room operations. Chem. Eng. Trans. **19**, 219–224 (2010)
6. Bye, A., Lois, E., Dang, V.N., Parry, G., Forester, J., Massaiu, S., Boring, R., Braarud, P.Ø., Broberg, H., Julius, J., Männistö, I. Nelson, P.: International HRA Empirical Study – Phase 2 Report: Results from Comparing HRA Method Predictions to Simulator Data from SGTR Scenarios, NUREG/IA-0216, vol. 2. US Nuclear Regulatory Commission, Washington D.C. (2011)
7. Williams, J.C., Bell, J.L.: Consolidation of the generic task type database and concepts used in the Human Error Assessment and Reduction Technique (HEART). Saf. Reliab. 2016a. (in press)
8. Williams, J.C.: A user manual for the HEART human reliability assessment method. DNV Technica report C2547. Nuclear Electric plc, Gloucester, March 1992. (unpublished)
9. Leon, M.R., Revelle, W.: Effects of anxiety on analogical reasoning: a test of three theoretical models. J. Pers. Soc. Psychol. **49**(5), 1302 (1985)
10. Williams, J.C., Bell, J.L.: Consolidation of the error producing conditions used in the Human Error Assessment and Reduction Technique (HEART). Saf. Reliab. **35**(3), 26–76 (2016)
11. Shackman, A.J., Sarinopoulos, I., Maxwell, J.S., Lavric, A., Pizzagalli, D.A., Davidson, R. J.: Anxiety selectively disrupts visuospatial working memory. Emotion **6**(1), 40–61 (2006)
12. Brosch, T., Pourtois, G., Sander, D., Vuilleumier, P.: Additive effects of emotional, endogenous, and exogenous attention: behavioral and electrophysiological evidence. Neuropsychologia **49**(7), 1779–1787 (2011)
13. Mendel, J.: The effect of interface consistency and cognitive load on user performance in an information search task. Clemson University, South Carolina, MS thesis (2009)
14. Williams, J.C., Bell, J.L., Daniels, C.L.: The quantitative impact of fatigue, distraction and age on human reliability. In: Aven, Vinnem (eds.) Risk, Reliability and Societal Safety – Proceedings of ESREL 2007, pp. 2053–2058 (2007)
15. Wickens, C.D., Hutchins, S.D., Laux, L., Sebok, A.: The impact of sleep disruption on complex cognitive tasks: a meta-analysis. Hum. Factors **57**(6), 930–946 (2015)
16. Council of the European Union: Council Directive of 12 June 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work (89/391/EEC), Brussels (1989)
17. Mäkinen, T.M., Palinkas, L.A., Reeves, D.L., Pääkkönen, T., Rintamäki, H., Leppäluoto, J., Hassi, J.: Effect of repeated exposures to cold on cognitive performance in humans. Physiol. Behav. **87**(1), 166–176 (2006)
18. Rosenbloom, T., Shahar, A., Perlman, A., Estreich, D., Kirzner, E.: Success on a practical driver's license test with and without the presence of another testee. Accid. Anal. Prev. **39** (6), 1296–1301 (2007)
19. Ardila, A., Ostrosky-Solis, F., Rosselli, M., Gómez, C.: Age-related cognitive decline during normal aging: the complex effect of education. Arch. Clin. Neuropsychol. **15**(6), 495–513 (2000)
20. Kvavilashvili, L., Kornbrot, D.E., Mash, V., Cockburn, J., Milne, A.: Differential effects of age on prospective and retrospective memory tasks in young, young-old, and old-old adults. Memory **17**(2), 180–196 (2009)

21. Salthouse, T.A., Hambrick, D.Z., Lukas, K.E., Dell, T.C.: Determinants of adult age differences on synthetic work performance. J. Exp. Psychol. Appl. **2**(4), 305–329 (1996)
22. Whiting, W.L., Smith, A.D.: Differential age-related processing limitations in recall and recognition tasks. Psychol. Aging **12**, 216–224 (1997)
23. Nunes, A., Kramer, A.F.: Experience-based mitigation of age-related performance declines: evidence from air traffic control. J. Exp. Psychol. **15**(1), 12–24 (2009)
24. Mittler, M.M., Carskadon, M.A., Czeisler, C.A., Dement, W.C., Dinges, D.F., Graeber, R. C.: Catastrophes, sleep, and public policy: consensus report. Sleep **11**(1), 100–109 (1988)
25. Folkard, S.: Black times: temporal determinants of transport safety. Accid. Anal. Prev. **29**(4), 417–430 (1997)

Assessing Dependency in SPAR-H: Some Practical Considerations

Harold S. Blackman[1(✉)] and Ronald L. Boring[2]

[1] Boise State University, Boise, ID, USA
`haroldblackman@boisestate.edu`
[2] Idaho National Laboratory, Idaho Falls, ID, USA
`ronald.boring@inl.gov`

Abstract. Dependency, the relationship between a series of human errors, is an important factor to consider when conducting a human reliability analysis (HRA). The premise of this paper is that we should not expect dependency to be applied frequently. This paper presents guidance on the application of dependency in the SPAR-H HRA method. Overall guidance is provided to prevent overuse of dependency, and then eight specific insights are provided to ensure dependency is applied correctly. Without proper application of dependency, there is the risk that human error probabilities produced by HRA methods will be inaccurate. This paper seeks to ensure dependency does not lead to spurious quantification of errors.

Keywords: Human reliability analysis · Dependency · SPAR-H method

1 Concept of Dependency

In human reliability analysis (HRA), dependency (also simply called *dependence*) for human actions refers to the extent to which performance on one task impacts performance of a subsequent task. Typically, dependency models failure paths, which is the likelihood that an error on one task increases the likelihood of error on a subsequent task, i.e., error begets error. Dependency may also model success paths, which is the likelihood that a success on one task increases the likelihood of success on a subsequent task, i.e., success begets success. In practice, the mechanism of success dependency is not as well understood as its counterpart for failure dependency. As such, success dependency is not included in most HRA methods and applications.

The Technique for Human Error Rate Prediction (THERP) [1], the original HRA method, considered three types of dependency, as depicted in Fig. 1. Given two Tasks —A and B—in sequence, Task A may influence Task B in the following manners:

- *Independence*—in which the outcome of Task B occurs regardless of the outcome of Task A
- *Direct dependence*—in which the outcome of Task A directly affects the outcome of Task B
- *Indirect dependence*—in which a mediator variable like a performance shaping factor (PSF) affects both Task A and Task B and therefore links the two tasks

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_2

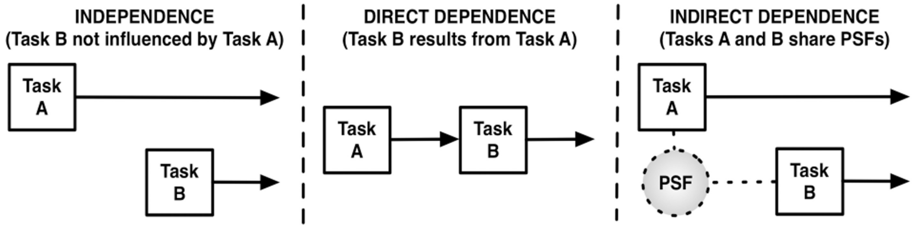


Fig. 1. Types of dependency in human reliability analysis.

Most HRA methods in contemporary practice do not explicitly model indirect dependence.

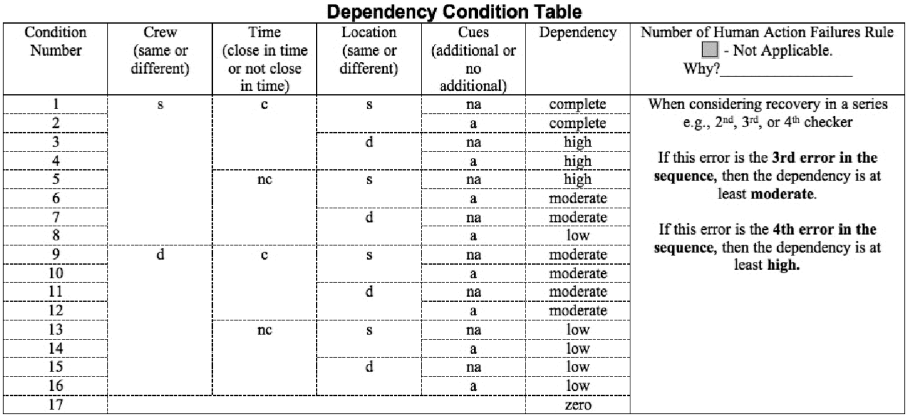
Dependency acts to increase the human error probability (HEP). HRA methods will therefore typically account for dependency by increasing the HEP proportionate to the level of dependency. The level of dependency is determined and influenced by a variety of factors. Classically, the kinds of factors that are considered include how close in time the tasks are performed, if the tasks are performed by the same individual or the same crew, or if the means of taking the action is closely located, such as two switches side by side. These are important considerations in that they describe a circumstance where little or no new information is provided that would cause a change in the tasking that is being carried out by the human. So, if an operator is opening a series of values on a train, and opens one, s/he is likely to open the second valve on the same train, whether correct or not.

There are a variety of methods to assess dependency; we will focus on THERP [1] and the Standardized Plant Analysis Risk-Human Reliability Analysis (SPAR-H) [2, 3] methods, as they are closely related. The dependency factors in THERP and SPAR-H include:

- *Time*—generally, the closer in time two tasks occur, the higher the dependency. Extended time allows forgetting and emptying of working memory, causing the human to rethink the situation.
- *Location*—generally, the more proximate two tasks occur, the higher the dependency. Changes in location introduce new information, potentially interrupting the operations script, allowing rethinking to occur.
- *Person or crew*—generally, if two or more tasks are performed by the same personnel, they are considered dependent. The same person of crew allows for mindset to develop with no new information being introduced.
- *Cues*—generally, if additional information cues appear between two tasks, the two tasks are considered independent. Cues can stimulate the human to think differently by focusing attention on an important, new piece of information.

The SPAR-H [3] method integrates these factors into a table that can used to determine the level of dependency (see Table 1 below), ranging from zero to complete dependency.

Table 1. SPAR-H [3] dependency level determination table.



Dependency Condition Table

Condition Number	Crew (same or different)	Time (close in time or not close in time)	Location (same or different)	Cues (additional or no additional)	Dependency	Number of Human Action Failures Rule ▨ - Not Applicable. Why?________________
1	s	c	s	na	complete	When considering recovery in a series e.g., 2nd, 3rd, or 4th checker If this error is the **3rd error in the sequence**, then the dependency is at least **moderate**. If this error is the **4th error in the sequence**, then the dependency is at least **high**.
2				a	complete	
3			d	na	high	
4				a	high	
5		nc	s	na	high	
6				a	moderate	
7			d	na	moderate	
8				a	low	
9	d	c	s	na	moderate	
10				a	moderate	
11			d	na	moderate	
12				a	moderate	
13		nc	s	na	low	
14				a	low	
15			d	na	low	
16				a	low	
17					zero	

Based on the assessment of the factors in Table 1 and the resulting determined level of dependency, a formula taken from THERP [1] is then used to modify the failure rate. In THERP's parlance, there are three stages of calculating HEPs:

- *Nominal HEP (NHEP)*—is the default HEP looked up for each task.
- *Basic HEP (BHEP)*—is the NHEP modified for context, which are typically PSFs.
- *Conditional HEP (CHEP)*—is the BHEP modified for dependency.

The dependency adjustment serves to anchor the HEP at a minimum value, namely 1/20 (0.05) for low dependency, 1/7 (0.14) for moderate dependency, 1/2 (0.5) for high dependency, and 1.0 for complete dependency. The formulas used for calculating the CHEP with dependency are found in Table 2. The approach represented in the tables is straightforward. The challenge is determining what level of dependence to apply and whether to apply any dependence modification at all.

Table 2. Dependency calculations derived from THERP [1].

Dependency level	Dependency adjustment	THERP equation
Zero	$\mathrm{CHEP} = \mathrm{BHEP}$	10–14
Low	$\mathrm{CHEP} = \frac{1 + 19\,\mathrm{BHEP}}{20}$	10–15
Moderate	$\mathrm{CHEP} = \frac{1 + 6\,\mathrm{BHEP}}{7}$	10–16
High	$\mathrm{CHEP} = \frac{1 + \mathrm{BHEP}}{2}$	10–17
Complete	$\mathrm{CHEP} = 1.0$	10–18

2 The Challenge of When to Apply Dependency

Dependency of one task upon another arises from the physical couplings (e.g., same crew performing the recovery actions) of the performer of the second task with respect to the occurrence and/or effect of the previous task. One interpretation of the human causal chain in error dependency is that it arises when there is a particular mindset about completing related tasks [4]. Unless there is something to disrupt the continuation of the mindset, e.g., new information cues, the completion of tasks may be considered as occurring in the same dependent mindset.

Mental models and mindsets are updated to coincide with experience and are, in turn, impacted by the other factors present in the environment that may influence performance. For example, cues such as alarms, indications from instrumentation, display-based alarm lists, and others are what the operator attempts to attach to their model of the situation. The more accurate these cues are and how well they match what the operator has stored in memory from training, the greater the tendency that he or she will make the correct diagnosis or take the correct action. Prior actions and errors can act as "current" cues and establish expectancies leading to propensities to look or not look for specific pieces of information. In other words, previous actions or recently experienced events create a mindset that guides decision making.

At the topmost level, if the operator has no knowledge of a prior task, then that task has no carryover cognitive effect for the operator. If the operator has knowledge of the prior task, then we must consider what that knowledge could affect. For example, a higher level of stress will result if the prior task has failed. This may negatively influence subsequent task performance. For available time, the important factor is whether excessive time required to take one action leaves less time for the next and influences the failure rate.

Dependency comes about from factors that create a context where the human has no stimulus to look at the situation with a fresh perspective. The actions are being taken from a developed context and the human is executing a script. So, how often is the human involved in operations where little new stimulus occurs? In fact, procedures have different steps and different tasks, requiring different experience, with humans that possess different memorial antecedents. In many cases there are long periods of time between actions being taken in different locations. All of these factors may disrupt the continuity of task flow. It is for this reason we believe that dependency should not be required often.

In order to determine if the dependency tables should be applied, one critical question must be considered: *Does a compelling reason exist to suggest that tasks are dependent?* If there is no compelling reason to believe that actions are dependent, then dependency should not be applied. Simply because one task follows another does not make the two tasks dependent. This is particularly true for actions on different systems, or actions that are being performed for different reasons (e.g., involving different mindsets).

3 Guidance for Application of Dependency

We have identified eight insights from examination of the literature and applications of SPAR-H, including development of HRAs across 70 generic nuclear power plant models and special applications in aerospace [5] and offshore oil [6]. Further, these insights reflect guidance [7] developed in response to questions from analysts at the U. S. Nuclear Regulatory Commission and individuals engaged in SPAR-H at nuclear utilities. These insights are:

1. *Direct dependency should not be applied when there is indirect dependency.* The PSFs that describe the context of a given situation in methods such as SPAR-H also allow the analyst to account for some of the impact of factors associated with dependency in the error rate. If two tasks feature similar PSFs, this may imply indirect dependency. In that case, the dependency error mechanism is already accounted for in the PSFs. Additionally applying a correction for direct dependency might risk double counting negative influences on performance. Applying dependency requires a thoughtful examination of all the variables, including those that indirectly affect dependency.
2. *Dependency should be applied to all second and subsequent checkers.* Experience has shown that analysts sometimes take undue credit in HRAs for recovery by second (and subsequent) checkers. This results in unrealistically low error probabilities. Not only should analysts be cautious in crediting recovery and checkers, it may actually be more appropriate to consider the adverse effects caused by dependency. Although operators may enter a recovery path, success is mitigated by the factors that contributed to the initial human error. Recovery does not necessarily disrupt the mindset that led to human error, and analysts should be cautious not to assume that the tendencies of previously faulty behaviors are fully overcome by a slight course correction. The effects of recovery should therefore be adjusted for the appropriate level of dependency. Of course, it may be the case that no dependency should be assessed for recovery actions due to the entry of other factors such as additional salient cues or shift changes that truly disrupt task linkages.
3. *Almost never use recovery for normal operations.* For example, a second checker on the same crew in normal operations executing repetitive tasks that are highly routine should not be credited for recovery, because there is a high likelihood that there is complete dependency between the two operators. There would be no compelling additional cues to highlight the need for a thorough second check. Instead, the second check becomes perfunctory. That is not to say that second checkers should not be included in operating environments. Rigorous conduct of operations and high reliability organizations are capable of improving the effectiveness of this strategy.
4. *Almost always use dependence for subtasks of a recovery sequence.* As noted in Insight No. 2, once a recovery sequence has begun, it is not a guaranteed success path. As such, it is necessary to account for the relative impact of earlier failures on subsequent tasks. A recovery sequence is typically performed by the same crew who committed the initial error, which means dependency continues to manifest

throughout the sequence of tasks. This guidance applies broadly to all manners of recovery, not just to second checkers as is the focus of Insight No. 2.

5. *Actions with different functions or purposes that are close in time are likely not dependent.* Although the timing of tasks in fact affects dependence, if these tasks are unrelated in terms of function or purpose, it is unlikely that they are connected in any meaningful way in the operators' mind and therefore must be considered independent. For example, two tasks in sequence such as completing a valve lineup followed by adjusting rod positions are not closely associated, involving different systems and functionality. These differences inherently cause a shift in thinking by the operators.
6. *Actions with the same functions or purposes that are not close in time are likely dependent.* In this case, the fact that the actions are functionally related and attempting to achieve a common goal will cause the operator to more strongly link those actions across time. Time will not erase the basis for the decision to start the task sequence. Time by itself is not an active cue for changing a course of action.
7. *Different crews (i.e., fresh eyes) are likely to break dependency.* Different crews are different from second checkers in that they have not been involved in actions or events up to the point of their assumption of duty. They are briefed by the prior crew in terms of the plant status and conditions but have not been biased by any incorrect observations or ongoing operational mindset. This allows them to approach the task in an independent fashion.
8. *Dependency may apply over long periods of time.* Current HRA methods, including SPAR-H, do not adequately deal with dependency over very long time frames. Dependency may persist due to fundamental misunderstandings of the root cause of an event. Crews may treat symptoms rather than an underlying cause. Without corrective action, the underlying human error may continue uncorrected. Good examples of this may be found in monitoring tasks where indirect observations or measurements are used to infer component or system status. Workarounds performed without proper knowledge of the underlying fault may continue for weeks to months, potentially resulting in further degradation to the component or system without operator awareness. Severe accidents may be another case of extended time windows for which dependency may play a role.

4 Conclusions

The application of dependency has generally been conservative, resulting in elevated HEPs. It is the case that any assessment of dependency in the moderate to high range will, in fact, dominate the error rate. We encourage analysts to consider the recommendations given in this paper to provide a more accurate account of human performance. Dependency can drive the human error probability in ways that do not accurately reflect human behavior. Decision making and cognition are affected in complex ways through the factors we have discussed. Because of this, analysts need to be attentive to the application of dependency. Simply following a dependency table may not yield accurate results.

References

1. Swain, A.D., Guttmann, H.E.: Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, NUREG/CR-1278. U.S. Nuclear Regulatory Commission, Washington, DC (1983)
2. Blackman, H.S., Byers, J.C.: ASP Human Reliability Methodology Development, INEL-95/0139. Idaho National Engineering Laboratory, Idaho Falls (1995)
3. Gertman, D., Blackman, H., Marble, J., Byers, J., Smith, C., O'Reilly, P.: The SPAR-H Human Reliability Analysis Method, NUREG/CR-6883. U.S. Nuclear Regulatory Commission, Washington, DC (2005)
4. Whaley, A.M., Kelly, D.L., Boring, R.L.: Guidance on dependence assessment in SPAR-H. In: Joint Probabilistic Safety Assessment and Management and European Safety and Reliability Conference, Helsinki (2012)
5. Chandler, F.T., Chang, Y.H.J., Mosleh, A., Marble, J.L., Boring, R.L., Gertman, D.I.: Human Reliability Analysis Methods: Selection Guidance for NASA. NASA Office of Safety and Mission Assurance, Washington, DC (2006)
6. Bye, A., Laumann, K., Taylor, C., Rasmussen, M., Øie, S., van de Merwe, K., Øien, K., Boring, R., Paltrinieri, N., Wærø, I., Massaiu, S., Gould, K.: The Petro-HRA Guideline. Institute for Energy Technology, Halden, Norway (2017)
7. Whaley, A.M., Kelly, D.L., Boring, R.L., Galyean, W.J.: SPAR-H step-by-step guidance. In: Joint Probabilistic Safety Assessment and Management and European Safety and Reliability Conference, Helsinki (2012)

Simulation-Based Optimization of Resilient Communication Protocol for Nuclear Power Plant Outages

Cheng Zhang[1], Zhe Sun[1], Pingbo Tang[1](✉), Shawn W. St. Germain[2], and Ronald Boring[2]

[1] Del E. Webb School of Construction, Arizona State University,
660 South College Avenue, Tempe, AZ 85281, USA
{cheng.zhang.7, zsun43, tangpingbo}@asu.edu
[2] Idaho National Lab, 1955 N. Fremont Avenue, Idaho Falls, ID 83415, USA
{shawn.stgermain, ronald.boring}@inl.gov

Abstract. During Nuclear power plant (NPP) outages, communications between outage participants (e.g., workers, managers, engineers) can be tedious and error-prone due to complicated organization of outage processes and crews, the extremely busy schedule with a 10-minute level of detail, and many people working in the field. Therefore, precisely predicting and controlling the time wasted during communications and remedying miscommunications can improve the NPP outage productivity. A communication protocol is a set of rules defining the organization structures, timing, channel, and content of communication according to the information transition needs of a workflow. In order to reduce the time wasted and ineffective information transition due to human errors in communications, the authors propose a communication protocol optimization according to the as-planned workflow. This optimization study evaluates how different communication protocols in an outage workflow will influence the time wasted under the influence of human errors and task duration uncertainties. The simulation results indicate that the proposed method can provide a reliable reference of improving the communication protocol in NPP outage workflows.

Keywords: Nuclear power plant outages · Communication protocol · Human error · Agent-based modeling

1 Introduction

Nuclear power plant (NPP) outages are among the most challenging projects involve large number of maintenance and repair activities with a busy schedule and zero-tolerance for accidents. During an outage, more than 2,000 workers will be working around the NPP and finishing the maintenance work including more than 2,000 tasks within about 20 days, while the planning stage of a typical NPP outage is more than 4 months. Moreover, an one-day delay in a NPP outage could cost $1.5 million. These features of NPP outages call for a real-time, robust, effective workflow control to

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_3

ensure the construction safety and productivity while reducing wastes and resource consumption.

During NPP outages, the communication activities are time-consuming, which is caused by the complicated organization of outage participants and processes [1–3]. The approval of each task involves multiple stakeholders in order to ensure safety. For example, an outage task should be confirmed by the following organizational units before the execution [4]: (1) the outage control center (OCC), which determining the necessity of a task based on the progress of the entire project and plant conditions; (2) schedulers, who configure the task schedules; (3) maintenance shops, who manage and assign workforces for tasks; (4) the main control room staff, who configures the NPP according to the requirements of certain tasks (e.g. mechanics can process turbine maintenance tasks only when its temperature is cooled down); (5) the work execution center, which inspects the site preparation for safe execution of a given task. Complicated communications between all these organizational units are necessary for safety, but will cause possible time wastes [5]. For example, the OCC need to have 30-minute meetings up to every three hours to know the as-is status of the outage progress and performance. In addition, the communication activities in NPP outages are error-prone, and communication errors could introduce additional communications and delays. Hobbins et al. [6] analyzed the errors, incidents, and near misses documented in the Licensee Event Reports (LERs). They found that for all the human-related incidents, about 50% of them are related to communication errors. Furthermore, the extremely busy schedules with a 10-minutes level of detail delays or mistakes due to communication could propagate to more tasks, which could compromise the productivity and safety of the entire workflow and even the whole outage. Therefore, precisely predicting and controlling the time wasted and information loss caused by human errors during communications can improve the productivity of NPP outages [7].

To reduce the time wasted and the information losses due to human errors in communication, the nuclear industry needs a system to help people to transit information effectively and efficiently in collaborative workflows. Such system is called a communication protocol, which is a set of rules defining the organization structures, timing, channel, and content of communication according to the information transition needs of a workflow. All these elements in a communication protocol can influence the performance of the collaborative workflow. For example, Sorense and Stanton [8] found that the different type of organizational structures between participants will influence the task success rate. However, current approaches about communication optimization only focuses on the specifically designed lab experiments and the result cannot be applied to NPP outage workflow. In addition, the theoretical framework of communication protocol optimization is missing. A gap exists between the needs to reduce the time wasted and miscommunications in NPP outage workflows and current understanding of communication of the academy.

To bridge this gap, this research will evaluate how different communication protocols in an outage will influence the time wasted under human errors and task duration uncertainties. This methodology will first model the spatial relationship, time relationship, and the human-resource relationship between tasks, which is termed as the outage workflow process of the workflow. Then this method identifies the uncertainties in the workflow including the indeterminate task durations and the possible human errors. The next of this methodology is modeling the communication protocol to mitigate the impact of the uncertainties. Finally, the simulation model enables the optimization of the parameters in the communication protocol. This research has two objective functions to evaluate the communication protocol in a collaborative workflow. The first objective function of communication protocol optimization is minimizing the duration of the entire workflow. Shorter workflow duration means less time wasted due to communication or human errors. The second objective function is minimizing the probability of critical path change in the schedule. Lower chance of critical path change means the management team can only focus on the progress of a few critical tasks to control the progress of the overall workflow, so that the labor and resource can be saved. To validate this methodology, the authors developed an agent-based modeling platform to test the performance of the NPP outage workflows using different communication protocols. Simulation result shows that the proposed method can provide a reliable reference of improving the communication protocol in NPP outage workflows.

2 A Method for Optimizing the Communication Protocols Used in NPP Outages

Optimizing the communication protocols for NPP outage workflows consist of four steps, which is shown in Fig. 1. Section 2.1 through Sect. 2.4 will introduce these steps in detail. The first step is to define the outage workflow process, which identifies the spatial, temporal and human relationship between tasks. The second step is to identify uncertainties, such as the as-is task duration that could deviate from as-planned durations and random communication errors in the workflow. Based on the identified uncertainties, the third step is to adjust the communication protocol based on the identified uncertainties in the workflow. Finally, the last step is to use the simulation result to optimize the parameters in the communication protocol.

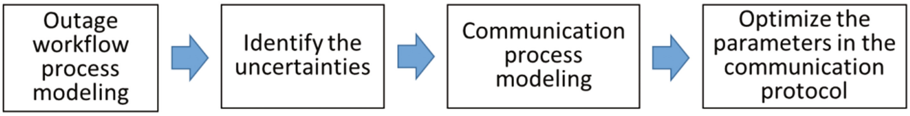


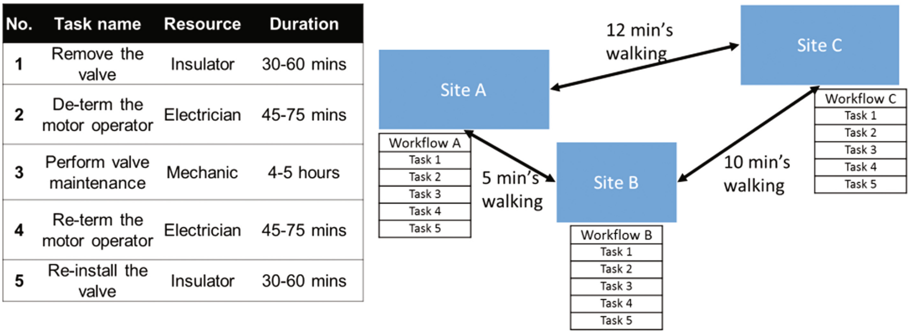
Fig. 1. Framework of optimizing the communication protocols for NPP outage workflows

2.1 Outage Workflow Process Modeling

The first step of optimizing the communication protocol in a teamwork oriented workflow is to define the outage workflow process, which consist of three types of relationship between tasks:

(1) *Spatial relationships*, which represent distances between sites of tasks. The combination of the spatial relationship between all tasks forms the workspace layout;
(2) *Time relationship*, which is the re-requisite relationship between tasks and as-planned task duration. The combination of the pre-requisite relationship between all tasks forms the as-planned schedule.
(3) *Human-resource relationship*, which defines which worker is in responsible for which tasks.

Figure 2 visualizes the three types of relationship between tasks in the NPP outage workflow used in this research. The tasks simulated in this research are valve maintenance activities during an NPP outage project spanning over three different sites (they are on Site A, B, and C, respectively) which is shown in Fig. 2. The workers need to complete five tasks on each site: remove the valve, de-term the motor operator, perform the valve maintenance, re-term the motor operator, and re-install the valve. Each task is assigned to a specific worker. In this workflow, the workers can only work on one site at a time, which makes them the shared resources.



No.	Task name	Resource	Duration
1	Remove the valve	Insulator	30-60 mins
2	De-term the motor operator	Electrician	45-75 mins
3	Perform valve maintenance	Mechanic	4-5 hours
4	Re-term the motor operator	Electrician	45-75 mins
5	Re-install the valve	Insulator	30-60 mins

Fig. 2. The spatial and temporal relationship between tasks in the outage workflow

Three workers collaborate with each other to complete the workflow. The as-planned critical path of the simulation model is A1-A2-A3-B3-C3-C4-C5 (shown in Fig. 3), which is the longest possible task chain between the start and the end of the workflow. The durations of the tasks on the critical path will decide the duration of the entire workflow. This critical path is calculated using the mean value of the task duration. In addition, the dispatching rule of the workflow is "first come, first serve". This means that if the worker has two available tasks, he or she will work on the task assigned to the worker earlier. Without losing the generality, the authors assume that the workers will follow the priority queue of Site A, B and C if the tasks at different sites are assigned to him or her at the same time. Therefore, the tasks sequence could be influenced by the uncertainty of the duration of each task.

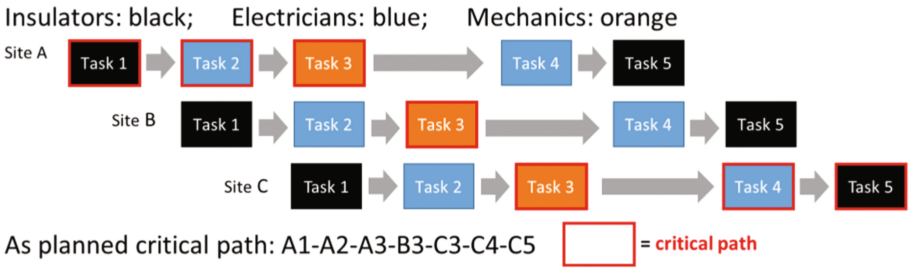


Fig. 3. Visualization of the as-planned schedule

2.2 Identifying and Modeling Uncertainties in the Test Case Workflow

If no uncertainty exists in any tasks, the workers do not need to communicate but just carry out the tasks following the as planned schedule. In this workflow model the duration of each task follows uniform distribution. For example, the electrician need to know that the insulator has finished the first task in Site A (noted as A1) before he/she can determine the start time of A2. For example, the uncertainty of task duration may cause the sequence of tasks to change. For example, after finishing task B2 the electrician could work on C2 or A4 because either of them may be ready for the electrician first. Such possibility of task sequence change shows the necessity of effective communication in complex NPP outage workflows.

On the other hand, humans are no perfect. Therefore, randomly occurred human errors should be considered in the communication protocol optimization. This research considers one type of human error that workers or the supervisor may forget to call when they need to should be considered in the communication protocol. If a worker forgot to call the supervisor when a task is finished, the whole team will never know that and the workflow cannot continue. Similarly, the workflow will also fail if the supervisor forgot to call. Such two types of uncertainty call for the effective and reliable information transition between workers, which is the need of communication protocol optimization.

2.3 Communication Process Modeling

The communication protocol needs to guarantee the success and efficiency of the NPP outage workflow under the influence of task duration uncertainty and randomly occurred human errors. To start from a simple case, the communication in this workflow is centralized, which means a supervisor will organize the communication of the entire team. Three workers (i.e. the insulator, the mechanics, and the electrician) can only talk with the supervisor but they are not allowed to allowed to talk with each other. Figure 4a visualizes the part of “Report finish task” in the communication protocol between the workers and the supervisor. Without losing generality, the insulator should call the supervisor when he/she finished A1 and report. After the talking on the phone with the insulator, the supervisor should call the electrician who is responsible for task A2 which is the successor of A1. After this phone call, the electrician will know that task A2 is

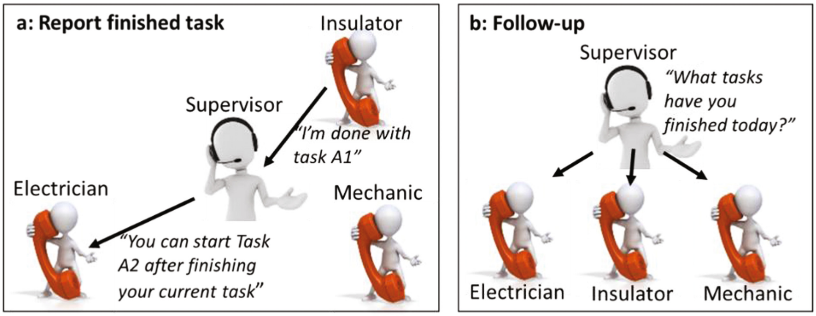


Fig. 4. Communication protocol of the team

ready for him/her to work on. To remedying miscommunications caused by human errors, which is forgetting to make phone calls, Fig. 4b visualizes the follow-up process in the communication protocol. At certain amount of time interval, the supervisor will call all the workers asking about what tasks have been finished in all. In this way, all the information about finished task can be recovered even if workers or the supervisor forget to communicate.

2.4 Optimize the Parameters in the Communication Protocol

This research has two indices to evaluate the communication protocol in a collaborative workflow. The first index of communication protocol optimization is minimizing the duration of the entire workflow. Shorter workflow duration means less time wasted due to communication or human errors. The second index is minimizing the probability of critical path change in the schedule. In outage schedules, the critical path could change due to newly discovered tasks, human errors, or uncertain resource availabilities. Moreover, tight schedules and packed workspaces cause high probability of error and uncertainty propagation throughout tasks that share the limited time, space, and resources on job sites, further increasing the probability of critical path shifting. The time wasted during communication will also interact with the critical path of a workflow. Therefore, a good communication protocol should not increase the chance of the changing of critical path.

The simulation model of the workflow can help optimize the interval of follow-up call according to the error rate of the team. In the simulation model, the researchers assume that: (1) the chance of the worker or supervisor forgetting to call ranges from 1%, 2% and 5% (2) the supervisor will not forget the follow-up call; (3) each phone call takes two minutes. If the forget-to-report rate is low and the follow-up call is frequent, the team will waste too much time on phone-calls. On the other hand, if the forget-to-report rate is high and the follow-up call is not frequent enough, the team will waste too much time in waiting. Therefore, the goal of the research is to optimize the interval of follow-up calls according to the probability of people forgetting to call.

3 Simulation-Based Communication Protocol Optimization Result

This section uses the simulation results to show how different communication protocol will influence the duration of the entire workflow and probability of critical path change. Shorter workflow duration means less time wasted due to communication or human errors. Lower chance of critical path change means the management team can only focus on the progress of a few critical tasks to control the progress of the overall workflow, so that the labor and resource can be saved. Sections 3.1 and 3.2 will compare these two optimization functions under different error rate of human (i.e. the probability of the worker or supervisor forgetting to call when they need to) and different follow-up call intervals.

3.1 Impact of Different Follow-Up Intervals on the Workflow Duration

In this section, the simulation results quantified the trade-off between making follow-up calls in a higher frequency, causing more time wasted during communication, or make follow-up calls in a lower frequency, causing more time wasted due to delayed report of task finishing information. First a baseline has been setup to be used for further comparison. For the baseline model with no communication errors, simulation results show that the average total duration for the work flow is 495 min. Then the authors have tested several scenarios considering different error rates with different time intervals between status checking. Results shown as the table below (Table 1).

Table 1. Delay of workflow duration (minutes) under the influence of different error rates and follow-up call intervals

Error rate	0.5 h	1 h	1.5 h	2 h	2.5 h	3 h	3.5 h
0%	52.8	28.2	19.9	17.8	15.0	14.9	11.0
1%	54.4	30.7	23.4	20.0	19.0	20.0	17.8
2%	56.2	32.2	27.5	24.6	24.7	27.4	25.1
5%	57.0	37.7	33.9	34.2	38.3	41.2	43.3

According to the results, when the error rate is one percent, which means out of one hundred times the worker will forget to report his status to the supervisor, the 3.5 h checking would be the best option. In other words, when the error rate is small enough, hourly-checking might not be necessary since delay can also be caused by unnecessary phone calls. However, when the error rate increased to 5%, the optimal option would be checking every 1.5 h. Considering 5% error rate would be large enough to fail the entire workflow, a 1.5 h checking strategy would be necessary to ensure the effectiveness of the entire workflow.

In addition, compare to the baseline result, even though the optimal option has prevented failure of the workflow in a relative shorter time period, the total duration is still delayed. Thus, further simulation is still need to find the optimized option to shorten the delay compare to the baseline.

3.2 Impacts of Different Follow-up Intervals on the Probability of Critical Path Change

Critical path method (CPM) is the major method used to identify important tasks to monitor in the workflows [9]. Complicated scheduling and resource allocation in NPP outages could enlarge the probability of shifting of the critical path because uncertainties can propagate to tasks sharing resources with the changed tasks [10, 11]. Therefore, quantifying the probability of critical path changes and the probabilities of tasks falling on critical paths can help outage management teams adjust schedules according to field changes.

In the baseline simulation model without human errors or follow-up calls, 72% runs hold the same critical path as planned. However, in other 28% cases, the critical path is different. Table 2 shows the statistics of the probability distribution of possible changes of critical path. This result shows that if the outage personnel pay much attention to the tasks on the as-planned tasks to make these tasks are on schedule, the overall workflow may still be delayed some non-critical tasks may become critical because of the uncertainty of task durations.

Table 3 shows the simulation results of critical path changing in the workflows with different probability of human errors and follow-up call intervals. All the values in the first row of Table 3 is larger than 72%, which means the activity of make follow-up call will reduce the chance of critical path changing. This is because during the follow-up call, the task on the critical path is definitely delayed, while the non-critical path task might still be waiting. Therefore, the tasks that have a higher chance on the critical path will also have a higher chance being delayed by the follow-up call. The tasks on the as-planned critical path are more likely to remain on the critical path with the frequent follow-up calls. The simulation result shows that the chances to maintain as-planned critical path is decreasing with the increase of human error rate and with the increase on the time interval between follow-up calls.

Table 2. All the possible critical path of the baseline workflow (No human error, no follow-up calls)

Critical path type	%
A1-A2-A3-B3-C3-C4-C5	72
A1-A2-B2-C2-A4-A5-B5-C5	9.2
A1-A2-B2-C2-A4-B4-B5(C4)-C5	7.7
A1-A2-B2-C2-C3-C4-C5	1.8
Others	9.3

4 Discussion

The simulation-based communication protocol optimization provides us a method to optimize the communication protocol considering the interaction between human error rate, delay of the workflow duration, and the critical path change. The simulation result (shown in Table 4) shows that, frequent status checking can help reduce the chance of

Table 3. Comparison of the probability of critical path change under different follow-up call interval and different human error rate

Error rate	0.5 h	1 h	1.5 h	2 h	2.5 h	3 h	3.5 h
0%	0.7%	7.1%	11.4%	19.4%	15.2%	14.2%	21.4%
1%	1.1%	5.6%	10.9%	20.1%	17.8%	20.6%	25.9%
2%	1.1%	8.4%	12.1%	22.5%	22.2%	20.7%	31.3%
5%	0.5%	8.2%	19.3%	27.6%	31.8%	29.7%	40.7%

critical path changing and mitigate the delay caused by human errors, but the communication time caused by frequent follow-up call will delay the entire workflow also. In order to balance the critical path change and delay of workflow duration considering different human error rate, the management team can set a threshold of “acceptable rate critical path change” and then choose the communication protocol that can minimize the workflow duration. For example, we can set the acceptable rate critical path change at 28% because it is the probability of critical path change in the baseline workflow without any human error or follow-up calls. Then we can choose the commutation protocol that satisfy this threshold and minimize the workflow duration. Table 4 tells that the optimized follow-up call interval is 3.5 h, 2 h, and 1.5 h (which are highlighted in bold) when the human error rate is 1%, 2%, and 5%, respectively.

Table 4. Comparison of workflow duration delay (WDD) and the probability of critical path change (CPC) under different follow-up call interval and different human error rate

Error rate	Function	0.5 h	1 h	1.5 h	2 h	2.5 h	3 h	3.5 h
1%	WDD	54.4	30.7	23.4	20.0	19.0	20.0	**17.8**
	CPC	1.1%	5.6%	10.9%	20.1%	17.8%	20.6%	**25.9%**
2%	WDD	56.2	32.2	27.5	**24.6**	24.7	27.4	25.1
	CPC	1.1%	8.4%	12.1%	**22.5%**	22.2%	20.7%	31.3%
5%	WDD	57.0	37.7	**33.9**	34.2	38.3	41.2	43.3
	CPC	0.5%	8.2%	**19.3%**	27.6%	31.8%	29.7%	40.7%

5 Conclusion

This research proposed a communication protocol optimization according to the as-planned workflow. According to different error rate of human, this method can optimize the duration of the entire workflow and the probability of critical path chance by modifying the parameters in the communication protocol. Simulation result shows that the proposed method can provide a reliable reference of improving the communication protocol in NPP outage workflows.

Acknowledgements. This material is based upon work supported by the U.S. Department of Energy (DOE), Nuclear Engineering University Program (NEUP) under Award No. DE-NE0008403. DOE’s support is acknowledged. Any opinions and findings presented are those of authors and do not necessarily reflect the views of DOE.

References

1. Petronas, Z.G.G., Shamim, A., Teknologi, U.: Managing People in Plant Turnaround Maintenance: the Case of Three Malaysian Petrochemical Plants (2016)
2. Tang, P., Zhang, C., Yilmaz, A., Cooke, N., Boring, R.L., Chasey, A., Vaughn, T., Jones, S., Gupta, A., Buchanan, V.: Automatic imagery data analysis for diagnosing human factors in the outage of a nuclear plant. In: Duffy, V.G.G. (ed.) DHM 2016. LNCS, vol. 9745, pp. 604–615. Springer, Cham (2016)
3. Zhang, C., Tang, P., Yilmaz, A., Cooke, N., Chasey, A., Jones, S.: Automatic crane-related workflow control for nuclear plant outages through computer vision and simulation. In: International Symposium on Automation and Robotics in Construction (ISARC), Auburn, AL (2016)
4. St. Germain, S.W., Farris, R.K., Whaley, A.M., Medema, H.D., Gertman, D.I.: Guidelines for implementation of an advanced outage control center to improve outage coordination, problem resolution, and outage risk management (2014). doi:10.2172/1168631
5. Gorman, J.C., Cooke, N.J., Winner, J.L.: Measuring team situation awareness in decentralized command and control environments. Ergonomics **49**, 1312–1325 (2006). doi:10.1080/00140130600612788
6. Hobbins, S., Cooke, N.J., Tang, P.: Analyzing licensee event reports for improved teamwork and nuclear power plant safety during outages. In: 22nd Occupational Health Week "Knowledge-Based Networks, Worldwide Connections for the Global Workplace (2016)
7. Kim, M.C., Park, J., Jung, W., Kim, H., Joong Kim, Y.: Development of a standard communication protocol for an emergency situation management in nuclear power plants. Ann. Nucl. Energy **37**, 888–893 (2010). doi:10.1016/j.anucene.2010.01.003
8. Sorensen, L.J., Stanton, N.A.: Y is best: how distributed situational awareness is mediated by organisational structure and correlated with task success. Saf. Sci. **56**, 72–79 (2013). doi:10.1016/j.ssci.2012.05.026
9. Siu, M.F.F., Lu, M., Abourizk, S.: Bi-level project simulation methodology to integrate superintendent and project manager in decision making: Shutdown/turnaround applications. In: Proceedings of - Winter Simulation Conference, pp. 3353–3364, January 2015. doi:10.1109/WSC.2014.7020169
10. Blackstone, J.H., Phillips, D.T., Hogg, G.L., Chaudhry, I.A., Khan, A.A.: A research survey: review of flexible job shop scheduling techniques. Int. Trans. Oper. Res. **23**, 551–591 (2016). doi:10.1111/itor.12199
11. Kapanoglu, M., Alikalfa, M.: Learning IF-THEN priority rules for dynamic job shops using genetic algorithms. Robot Comput. Integr. Manuf. **27**, 47–55 (2011). doi:10.1016/j.rcim.2010.06.001

Task and Procedure Level Primitives for Modeling Human Error

Ronald Boring[1(✉)], Martin Rasmussen[2], Thomas Ulrich[1],
Sarah Ewing[1], and Diego Mandelli[1]

[1] Idaho National Laboratory, Idaho Falls, ID, USA
{ronald.boring, thomas.ulrich, sarah.ewing,
diego.mandelli}@inl.gov
[2] NTNU Social Research, Trondheim, Norway
martin.rasmussen@svt.ntnu.no

Abstract. In this paper, we describe the development of behavioral primitives for use in human reliability analysis (HRA). Previously, in the GOMS-HRA method, we described the development of task level primitives, which model basic human cognition and actions. Like generic task types found in some HRA methods, the task level primitives provide a generic or nominal human error probability. These generic task types are often modeled at the task level—grouped according to a high-level goal that includes many activities. In contrast, task level primitives represent a finer level of task decomposition, corresponding not to a group of actions that comprise an overall task but rather individual steps toward that task. In this paper, we further elaborate on the task level primitives by grouping task level primitives into procedure level primitives. This terminology reflects standard groupings of activities that are performed by reactor operators when following operating procedures. For the purposes of HRA, it is desirable to model operator actions according to these prescribed procedure categories. We present mappings of the procedure level to the task level primitives found in the GOMS-HRA method. We provide examples and conclude that procedure level primitives are a useful tool to streamline HRA modeling and quantification, especially for dynamic HRA applications.

Keywords: Human reliability analysis · GOMS-HRA · Subtask modeling · Procedures · Dynamic human reliability · Analysis

1 The Importance of Subtasks in Human Reliability Analysis

In practice, human reliability analysis (HRA) looks at causes of human error at the *task* level. A task consists of a series of activities related to a common goal. This goal is typically centered on the function of a particular hardware component or system. For example, the goal to put feedwater into service at a nuclear power plant may entail multiple steps by a reactor operator, including preparatory actions like checking the availability of systems, the culmination of which is a functioning feedwater system at the plant. This division of a task into subtasks is mirrored in many operating procedure hierarchies, in which procedure steps feature substeps and in which groups of procedure steps have collective headings to indicate groupings of related steps.

© Springer International Publishing AG (outside the USA) 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_4

The common unit of analysis in HRA is the human failure event (HFE), which is defined as the failure of a function, component, or system due to human action or inaction [1]. This definition may be seen as top-down—from the hardware to the human—in the sense that only significant hardware failures that affect the safety of the facility are considered for modeling. The human presents one of many sources of such hardware failures. As noted in [2], most human factors methods associated with task analysis tend to build a bottom-up structure of human actions—from the human to the hardware with which he or she interfaces. The bottom-up approach may capture human errors that are not readily modeled in the top-down approach, but they may not all be risk significant. Logically, therefore, it only makes sense to include human errors that actually have an impact on facility safety. However, it is possible that the top-down approach overlooks some opportunities for significant errors, such as those caused by errors of commission—actions taken by the operator that aren't required and may change the facility from its expected configuration. Moreover, the top-down approach may omit a consistent or complete modeling level of human actions, focusing instead on the most salient actions of the operator.

The risk of the HFE as a unit of analysis is that it is very high level, potentially encompassing dozens to hundreds of human subtasks related to hardware in a top-down fashion. This level of task composition is very rough and may highlight certain salient actions while overlooking seemingly less significant actions. Importantly, the HFE level is difficult to replicate consistently between analysts, as the question of what is omitted from the task list is left to analyst discretion and expertise. For this reason, inconsistent task modeling within the HFE was implicated as a significant reason for variability between human reliability analysts in one HRA benchmark [3].

The Technique for Human Error Rate Prediction (THERP) [4], arguably the original HRA method, did not explicitly model HFEs. Instead, it considered groups of human actions within an HRA event tree, a unique representation for linking subtasks. This structure was particularly important for THERP, because it specified how the human error probability (HEP) should be calculated. Each node in a THERP HRA event tree is a human subtask; the tree models how HEPs, recoveries, and dependence between subtasks occur. In most cases, the HRA event tree as a whole may be seen as an HFE, and the total HEP is associated with the interconnections between subtasks. The subtasks are quantified individually through lookup tables, which create a type of scenario-matching approach. Novel subtasks are mapped to similar subtasks in THERP tables, for which HEPs are provided. Dependence in the propensity of error to beget error, resulting in increased error rates for the second and subsequent subtasks in a series. In contrast, recovery breaks the error chain and puts the operator back on a success path. Whereas dependence increases the HEP, recovery decreases it.

It can be extremely labor intensive to complete an HRA in THERP, in large part due to the necessity to model many subtasks. One of the chief simplifications of later HRA methods was the focus on the HFE instead of the subtask level of analysis. While no HRA method to our knowledge specifically advocates omission of a thorough task analysis, there is nonetheless an erosion of such efforts in common practice, because most probabilistic risk models only require input of a single HEP for the entire HFE. Even THERP and methods directly derived from it (e.g., [5]) currently tend to be used primarily for quantification of overall HFEs, not subtask quantification.

While there appears to be a trend away from subtask modeling in HRA as practiced, subtask modeling remains especially important for dynamic HRA, in which human activities are placed into a simulation. For example, in a recent dynamic HRA effort called the Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER) [6], a virtual human operator is modeled alongside thermohydraulic code. HUNTER models each required human intervention in the plant with corresponding plant response. HUNTER actions are therefore being modeled for each step in the operating procedures, autocalculating the HEP [7] and required completion time for each step. While dynamic HRA may be considered creating a virtual operator, it is also possible to conceive of this type of modeling as creating a virtual analyst whose job it is to calculate the HEP automatically for each HFE based on available information in the model [8]. Regardless of whether it is a virtual operator or analyst, dynamic HRA requires the subtask level of granularity when considering human actions. In this paper, we focus on subtask modeling to accommodate current efforts at modeling operator actions in the HUNTER dynamic HRA framework.

2 GOMS-HRA for Subtask Modeling

The Goals-Operators-Methods-Selection rules (GOMS) approach is a method developed to factor human information processing (i.e., cognition) into empirical observations [9]. GOMS is an important method for decomposing human activities into their constituent subparts. It may be considered a task analysis approach focused on the subtask level of mental operations. Adaptations of GOMS, like the Keystroke Level Model (KLM) [10] have been used to assign action specific timing data to human subtasks. Combining these subtask durations allows analysts to determine the relative efficiencies of different system designs, for example.

GOMS was recently adapted to HRA to create GOMS-HRA [11] to encompass the subtask level of human activities for dynamic HRA. Because GOMS is a framework more than a specific method for considering human activities at the subtask level, GOMS-HRA pays homage to GOMS, but it should not be considered an adaptation of any of the previous instantiations of the method.

GOMS-HRA is linked to a taxonomy of human subtask primitives, corresponding to basic human activities likely to be performed in conjunction with operating a nuclear power plant. The Systematic Human Error Reduction and Prediction Approach (SHERPA) [12] serves as the basic taxonomy, with adaptations for cognitive decision making [13] and periods of relative inactivity such as waiting and monitoring. The modified SHERPA taxonomy for use in GOMS-HRA is provided in Table 1. In our nomenclature, we call this list *task level primitives* (TLPs).

Note that the action (A), checking (C), retrieving (R), and selecting (S) TLPs make a distinction between control boards (i.e., main control room) and field (i.e., balance of plant) operations, as denoted by a subscripted $_C$ or $_F$, respectively. The instruction (I) TLP distinguishes between producing ($_P$) and receiving ($_R$), respectively. The decision making (D) TLP considers decisions guided by procedures ($_P$) or without procedures ($_W$).

Table 1. GOMS Operators used to define Task Level Primitives.

Primitive	Description
A_C	Performing required physical actions on the control boards
A_F	Performing required physical actions in the field
C_C	Looking for required information on the control boards
C_F	Looking for required information in the field
R_C	Obtaining required information on the control boards
R_F	Obtaining required information in the field
I_P	Producing verbal or written instructions
I_R	Receiving verbal or written instructions
S_S	Selecting or setting a value on the control boards
S_F	Selecting or setting a value in the field
D_P	Making a decision based on procedures
D_W	Making a decision without available procedures
W	Waiting

Note also that the difference between checking (C) and retrieving (R) has to do with the level of information being sought. Checking is the simple act of confirming information like a status light on the control boards. In contrast, retrieving requires greater cognitive engagement such as reading the exact value on a gauge or trending a value over time. Generally speaking, the more cognitive effort that is required, the more the categorization falls to retrieval.

This taxonomy serves not only to decompose human activities into elemental subtasks; the taxonomy also affords the ability to anticipate common error types for each subtask. The adapted SHERPA taxonomy from Table 1 yields the following types of errors at a high level:

- *Action Errors*—Performing the required action incorrectly or failing to perform the action
- *Checking Errors*—Looking for required information in wrong place or failing to look for that information
- *Retrieval Errors*—Obtaining wrong information such as from control room indicators or failing to obtain required information
- *Information Communication Errors*—Communicating incorrectly or misunderstanding communications
- *Selection Errors*—Selecting the wrong value or failing to select a value
- *Decision Errors*—Making wrong decision or failing to make decision.

Waiting is not a TLP in the sense of modeling failed actions and HEPs; instead, it acts as a placeholder for tasks such as monitoring that involve extended periods of time. Therefore, waiting is not modeled as a potential error type, although we acknowledge that there is opportunity for human errors to occur while waiting. Each of the errors associated with TLPs can, in turn, be decomposed into further error types similar to what is found in [14]. GOMS-HRA stops short of providing a catalog of possible failure mechanisms for each TLP, although such a catalog may be the topic of future research efforts.

As noted, HRA methods like THERP use a scenario-matching approach for quantification. The task or subtask at hand is compared against a lookup table of prequantified nominal HEPs and subsequently fine-tuned through further analysis. Similarly, the TLPs can serve as a series of generic task types with associated nominal HEPs. Table 2 includes nominal HEPs for each of the TLPs, as aligned to THERP subtasks [4] using expert judgement. Unlike THERP, which includes fairly specific scenario matches, the GOMS-HRA TLPs are characterized by the type of human activity rather than a specific scenario. As such, we believe the TLPs are more generalizable than the scenarios found in THERP. The TLPs allow maximum flexibility for modeling human activities in a dynamic simulation.

Table 2. HEPs associated with each talk level primitive.

Operator	Nominal HEP	THERP source[a]	Notes
A_C	0.001	20-12 (3)	Assume well-delineated controls
A_F	0.008	20-13 (4)	Assume series of controls
C_C	0.001	20-9 (3)	Assume well-delineated indicators
C_F	0.01	20-14 (4)	Assume unclear indication
R_C	0.001	20-9 (3)	Assume well-delineated indicators
R_F	0.01	20-14 (4)	Assume unclear indication
I_P	0.003	20-5 (1)	Assume omit a step
I_R	0.001	20-8 (1)	Assume recall one item
S_C	0.001	20-12 (9)	Assume rotary style control
S_F	0.008	20-13 (4)	Assume series of controls
D_P	0.001	20-3 (4)	Assume 30-minute rule
D_W	0.01	20-1 (4)	Assume 30-minute rule
W	n/a	n/a	n/a

[a] Corresponds to THERP [4] Table values from Chap. 20.

3 Introducing Procedure Level Primitives

The TLPs are at a more basic level than are the actions commonly prescribed to reactor operators. Reactor operators follow operating procedures ranging from standard operating procedures, annunciator response procedures, abnormal operating procedures, emergency operating procedures, to severe accident management guidelines (SAMGs). SAMGs tend to be different from the rest of the procedures in that they provide problem solving strategies rather than step-by-step processes. The remaining procedures articulate step-by-step actions the operators should follow to maintain production and safety at the plant. In fact, there are often license penalties for deviating from the procedures, making them legally binding process manuals.

Because the procedure steps serve as the script for operating the plant, the logical level of task decomposition for HRA is the procedure step. Procedure steps and substeps explicate exactly what the reactor operators need to be doing at the plant, including any interfacing with components, and in which sequence. Procedures also point to specific decision branch points. As such, it is possible to create a simplified model of the reactor operator without necessarily creating a full-blown artificial intelligence system [15]. The level of procedural detail and the high procedural compliance create a perfect context for using procedures for simplified dynamic modeling frameworks like HUNTER.

To link the TLPs to the procedures, we developed *procedure level primitives* (PLPs) that map common procedure steps to TLPs. In many cases, a single procedure step may actually entail a series of TLPs. Consider, for example, the common procedure step to `check` something such as an indicator (see Fig. 1). This procedure step corresponds to the TLPs of check (C_C), making a decision based on procedures (D_P), verbalizing the value to the shift supervisor (I_P), selecting or setting a value (S_C or A_C) if necessary, and potentially waiting (W) while monitoring the value.

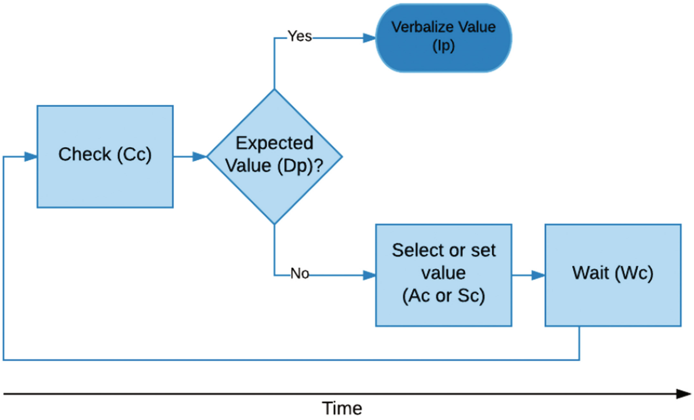


Fig. 1. Procedure level primitive decomposition into task level primitive example.

To simplify the process of modeling TLPs, we have mapped a number of common procedure steps to TLPs (see Table 3). These mappings, which constitute the PLPs, may be reused across analyses and may make it possible to extract TLPs in an automated fashion from operating procedures.

To arrive at a standard list of PLPs, we referenced the Procedure Professionals Association (PPA) *Procedure Writer's Manual* [16]. The PPA manual provides an extensive list of action verbs and their definitions to guide procedure development at nuclear power plants and other facilities. An example definition for `check` is:

```
CHECK: Observe an expected condition exists
          (no actions to correct)
```

Table 3. Common procedure level primitives mapped to task level primitives.

Procedure level primitive	Task level primitive	Mapping Notes
Determine*	C_C or R_C	Information type dependent
Ensure*	C_C or R_C and/or A_C and/or S_C	Information and control action type dependent
Initiate	A_C	–
Isolate	A_C	–
Minimize	S_C	–
Open	A_C	–
Verify*	C_C or R_C	Information type dependent
Check*	D_P, A_C, S_C, W_C, and/or I_P	Information type dependent

* These procedure level primitives, or action verbs, can be decomposed into multiple task level primitives. Figure 1 depicts the check procedure primitive decomposed into D_P, A_C, S_C, W_C, and I_P task level primitives and the relationship between these task level primitives.

The list of procedure steps (or action verbs) provided by PPA is extensive, but it is not necessarily exhaustive of all possible procedure steps at plants, nor does it narrow the list of procedure steps according to frequency. Moreover, the PPA manual is only a guideline, meaning individual plant's use of particular preferred procedure terminology or adherence to the suggested definitions will vary. Indeed, the consistency between procedures within individual plants varies, depending on the procedure writers and operators involved in generating the procedures as well as the system being proceduralized. There can be, for example, considerable differences in preferred nomenclature between procedures for the main control room vs. field operations. Mapping all procedure steps in the PPA list as PLPs would prove difficult and be fraught with necessary subjective interpretations to meet plant specific details. Instead of a complete mapping, we have mapped PLPs on a case-by-case basis as we have encountered new steps in procedures we are modeling as scenarios in HUNTER. This process, over time, is building a library of PLPs. Common PLPs that recur across many procedures are found in Table 3.

4 Discussion

4.1 Complex Mappings

As demonstrated with the `check` procedure step, PLPs can consist of multiple TLPs. The challenge with the reuse of PLPs across analyses is that complex mappings may not always be consistent. In one case, a `check` step may be synonymous with the checking (C_C) TLP, but, in another case, it may require multiple TLPs. Where complex mappings occur, model building will still require expertise by the human analyst to ensure the mappings are appropriate. There is still value in reusing PLPs, since the suggested mappings can serve as a template for reuse, but complex PLPs will likely require manual fine-tuning to suit their use context.

Using text mining, we have explored the possibility to extract procedure steps to derive PLPs automatically from operating procedures [17]. This process is still in the early stages of exploration. One of the main challenges of text mining PLPs is with complex PLPs, whereby the exact mappings to a series of TLPs requires more contextual information than can readily be extracted automatically. Other challenges remain. The formatting of procedures (e.g., the common double column format for procedures for Westinghouse pressurized water reactors) presents puzzling logic and parsing for current text mining algorithms. Certain placeholder words like `if` and `not` are at least as meaningful as standard action verbs, yet these words are semantically excluded from text mining. Finally, there are many challenges in differentiating what actions a reactor operator is doing versus what the plant is doing. For example, consider the following illustrative procedure step:

```
check that valve is closed and pressure is deceasing
```

The operator action is `check`. However, there are potential points of confusion over the related word stems of `close` and `decrease`. There remains considerable development work to refine the automatic extraction of PLPs from operating procedures. Currently, we are manually extracting the procedure steps and logic to arrive at accurate models of operator actions.

4.2 The Problem with Procedures

Of course, one of the main limitations of the PLP approach is that it relies on operating procedures. A few specific limitations are detailed below:

- *Variability in procedures*. Terminology varies considerably between plants, but there may also be intra-plant variability between procedures depending on the system addressed. Such variability is not a limitation of the procedures or the plant, but it makes the process of creating a global set of PLPs implausible. To address this reality, we have crafted PLPs on a procedure-by-procedure basis, tailoring the underlying TLPs as needed. In our anecdotal experience, the PLPs have proved robust in mapping TLPs. Initial experience suggests strong suitability to reuse of the PLPs across procedures, although we are carefully vetting individual instantiations of them to account for variability across procedures.
- *Procedures as scripts*. Procedures do not represent the totality of operator actions. There are routine tasks such as monitoring and communicating that are so frequent as not to warrant mention in specific procedures. While it is tempting to use the procedures as a script for modeling operator actions in HRA, doing so could result in an unrealistic account of what operators do. It would omit key routine activities, but it would also suggest a linearity of actions that is not representative of operator actions. Operators sometimes need to deviate from procedural paths or select alternate procedures in response to emerging plant conditions. While procedures may be a good starting point for modeling operator tasks, they actually underspecify

everything the reactor operators are doing. At best, procedures can be used to create simplified models of operator actions.[1] One key advantage of PLPs is that they can incorporate some of the underspecified actions—the between-the-lines activities of the operators—as part of the PLP. If, for example, each operator action is assumed to include some form of threeway communication between crew members, the PLPs for each procedure step should include the instructions (*I*) TLP. PLPs can ensure that the right TLPs are included in the analysis.

- *Task roles*. Procedures generally provide good detail about what tasking needs to be completed, but they do not specify who is completing that tasking. This is not an omission in the procedures; rather, it is by design, because the assignment of specific tasks or procedure steps is the job of the control room supervisor (CRS). The CRS maintains situation awareness of plant conditions and processes, follows and often reads aloud procedures, and delegates tasking between available crew members. The net effect whether the operator at the controls or the balance-of-plant operator performs a particular procedural step is negligible in most HRA models. However, for dynamic HRA, where the goal is to create a virtual operator, crew roles do matter, especially for determining relative workload of specific operators. As such, procedures cannot be used blindly, but rather must be augmented to specify which operator is performing the tasks.
- *Complex logic and compound steps*. As discussed briefly in the context of text mining, procedures are rarely as simple as one step equals one action. Instead, procedures often feature complex AND/OR logic and compound steps. Complex logic can be accounted for with TLPs related to decision making, and compound steps are simply chains of TLPs. The PLPs can likewise track the TLPs for added complexity and steps, but the generalizability of such PLPs may prove minimal for complex steps.

It is telling that many HRA methods include two considerations of procedures as a performance shaping factor. The first aspect is the *procedural quality for the task*. Procedure writers strive to cover plant conditions as completely as possible, but it is never possible to anticipate every process permutation at the plant. Compound faults, for example, may force operators to prioritize their response and choose between competing procedures. The second aspect of procedures considered in many HRA methods is *procedural adherence*. Reactor operators are trained both to follow operating procedures and to recognize when procedures may not adequately cover the breadth of possible responses. The reactor operators must exercise their expertise, which may on occasion take them outside of the procedural script or branch them to a new procedure that is a better fit to the plant conditions. There are multiple success paths to recover from a plant upset, for example, and crews may respond differently throughout the evolution of the transient.

Since the procedural quality and procedural adherence are known to vary, these will certainly limit the ability of the PLPs to become one size fits all across analyses. Of

---

[1] Note that a simplified model based on PLPs will likely be incrementally more detailed than a simplified HFE-level model. One premise of GOMS-HRA is that the more detail that is available in the HRA model, the higher the fidelity and scrutability of the quantification.

course, the purpose of PLPs is not to be generic constructs that are interchangeable across all contexts. The PLPs are simply a way of bundling common groups of TLPs that are associated with procedure steps. If the procedure steps do not lend themselves to PLPs, a more detailed case-by-case TLP analysis is warranted.

4.3 Advantages of PLPs

While the preceding discussion has highlighted some challenges and shortcomings of using PLPs, we believe there is merit in the approach. Where appropriate, PLPs give a way of linking procedure steps to TLPs in GOMS-HRA. This approach can greatly benefit efforts at dynamic HRA modeling in frameworks like HUNTER by providing the basis for error quantification and even task timing [18]. Additionally, the TLPs provide a basis for anticipating certain types of errors that might occur in the context of procedure following. In many cases, PLPs can be reused, thereby reducing the laborious efforts associated with model building. Ultimately, the PLP approach provides a consistent way to decompose procedure steps into meaningful subtasks in HRA. This approach is especially useful for dynamic HRA for heavily proceduralized nuclear power plant activities, but PLPs hold equal promise for any HRA that requires subtask modeling.

Acknowledgements. This work of authorship was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights. Idaho National Laboratory is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517.

Hoomaikai Hawaiʻi, no ka leʻaleʻa ka hihia oiai kākau i kēia pepa. The authors express their thanks to Dr. Curtis Smith at Idaho National Laboratory for championing the research of HUNTER and GOMS-HRA and funding various parts under the U.S. Department of Energy's Light Water Reactor Sustainability program and Idaho National Laboratory's Laboratory Directed Research and Development program.

References

1. American Society of Mechanical Engineers: Addenda to ASME/ANS RA-S-2008 Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications, ASME/ANS RA-Sb-2013, New York (2013)
2. Boring, R.L.: Top-down and bottom-up definitions of human failure events in human reliability analysis. In: Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting, pp. 563–567 (2014)
3. Poucet, A.: Human Factors Reliability Benchmark Exercise, Synthesis Report, EUR 12222 EN. Office for Official Publications of the European Communities, Luxembourg (1989)

4. Swain, A.D., Guttmann, H.E.: Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, NUREG/CR-1278. U.S. Nuclear Regulatory Commission, Washington, DC (1983)
5. Gertman, D., Blackman, H., Marble, J., Byers, J., Smith, C., O'Reilly, P.: The SPAR-H Human Reliability Analysis Method, NUREG/CR-6883. U.S. Nuclear Regulatory Commission, Washington, DC (2005)
6. Boring, R., Mandelli, D., Rasmussen, M., Herberger, S., Ulrich, T., Groth, K., Smith, C.: Human unimodel for nuclear technology to enhance reliability (HUNTER): a framework for computational-based human reliability analysis. In: 13th International Conference on Probabilistic Safety Assessment and Management (PSAM 2013), paper A-531 (2016)
7. Rasmussen, M., Boring, R.L.: Implementation of complexity in computation-based human reliability analysis. In: Proceedings of the European Safety and Reliability Conference on Risk, Reliability and Safety: Innovating Theory and Practice, pp. 972–977 (2016)
8. Rasmussen, M., Boring, R., Urich, T., Ewing, S.: The virtual human reliability analyst. In: Proceedings of the 1st International Conference on Human Error, Reliability, Resilience, and Performance, Los Angeles (2017, in press)
9. Card, S.K., Newell, A., Moran, T.P.: The Psychology of Human-Computer Interaction. Lawrence Erlbaum Associates, Hillsdale (1983)
10. Card, S.K., Moran, T.P., Newell, A.: The keystroke-level model for user performance time with interactive systems. Commun. ACM **23**(7), 396–410 (1980)
11. Boring, R.L., Rasmussen, M.: GOMS-HRA: a method for treating subtasks in dynamic human reliability analysis. risk, reliability and safety: innovating theory and practice. In: Proceedings of the European Safety and Reliability Conference, pp. 956–963 (2016)
12. Stanton, N.A., Salmon, P.M., Rafferty, L.A., Walker, G.H., Baber, C.: Human Factors Methods: A Practical Guide for Engineering and Design, 2nd edn. Ashgate Publishing Co., Aldershot (2013)
13. Boring, R.L.: Defining human failure events for petroleum applications of human reliability analysis. Procedia Manufact. **3**, 1335–1342 (2015)
14. Whaley, A.M., Xing, J., Boring, R.L., Hendrickson, S.M.L., Joe, J.C., Le Blanc, K.L., Morrow, S.L.: Cognitive Basis for Human Reliability Analysis, NUREG-2114. U.S. Nuclear Regulatory Commission, Washington, DC (2016)
15. Coyne, K.A.: A Predictive Model of Nuclear Power Plant Crew Decision-Making and Performance in a Dynamic Simulation Environment. PhD Dissertation, University of Maryland, College Park (2009)
16. Procedure Professionals Association: PPA AP-907-005, Procedure Writer's Manual, Revision 2 (2016)
17. Ewing, S.M., Boring, R.L., Rasmussen, M., Ulrich, T.: Text mining for procedure-level primitives in human reliability analysis. In: Proceedings of the 1st International Conference on Human Error, Reliability, Resilience, and Performance, Los Angeles (2017, in press)
18. Ulrich, T., Boring, R., Ewing, S., Rasmussen, M., Mandelli, D.: Operator timing of task level primitives for use in computation-based human reliability analysis. In: Proceedings of the 1st International Conference on Human Error, Reliability, Resilience, and Performance, Los Angeles (2017, in press)

Operator Timing of Task Level Primitives for Use in Computation-Based Human Reliability Analysis

Thomas Ulrich[1(✉)], Ronald Boring[2], Sarah Ewing[2],
and Martin Rasmussen[3]

[1] University of Idaho, 709 S Deakin St, Moscow, ID 83844, USA
thomas.ulrich@inl.gov
[2] Idaho National Laboratory, PO Box 1625, Idaho Falls, ID 83415, USA
{ronald.boring, sarah.ewing}@inl.gov
[3] NTNU Social Research, Studio Apertura, Dragvoll Allé 38 B,
7491 Trondheim, Norway
Martin.rasmussen@ntnu.no

Abstract. Computation-based human reliability analysis (CoBHRA) provides the opportunity for dynamic modeling of human actions and their impacts on the state of the nuclear power plant. Central to this dynamic HRA approach is a representation of the human operator comprised of actions and the time course over which those actions are performed. The success or failure of tasks is time dependent, and therefore modeling different times at which the operator completes actions helps predict how timing differences affect the human error potential for a given task. To model the operators' timing variability, Goals, Operators, Methods, and Selection rules (GOMS) task level primitives were developed based on simulator logs of operators completing multiple scenarios. The logs have sufficient detail to determine the timing information for procedure steps and to map the procedure steps into the task level primitives. The task level primitives can then be applied to other procedures that were not evaluated, since they represent generic task level actions applicable to all procedure steps. With these generic task level primitives, untested scenarios can be dynamically modeled in CoBHRA, which expands the usefulness of the approach considerably. An example is provided of a station blackout scenario, which demonstrates how the operator timing of task level primitives can enhance our understanding of human error in nuclear process control.

Keywords: Human reliability analysis · Computation-based human reliability analysis · Human error · GOMS-HRA · Station blackout

1 Introduction

The work presented here represents one component of a larger effort to dynamically model human actions and their consequences on the plant. Specifically, this work consists of developing the timing components for a virtual operator completing procedure steps in a computation-based human reliability analysis (CoBHRA) approach

© Springer International Publishing AG (outside the USA) 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_5

developed for nuclear process control. This CoBHRA framework is called Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER) [1]. HUNTER relies on a virtual operator that interfaces with a realistic plant model capable of accurately simulating plant thermohydraulic physics behaviors [1, 2]. Ultimately, the virtual operator will consist of comprehensive cognitive models comprised of artificial intelligence, though at this time a much more simplified operator model is used to capture the diagnosis and action behaviors of a typical operator. HUNTER is a merger between an area where HRA has previously been represented—probabilistic risk models—and an area where it has not—realistically simulated plant models through mechanistic thermohydraulic multi-physics codes. Through this approach, it is possible to evaluate a much broader spectrum of scenarios, both those based on previous experience and those that are unexampled, i.e., have not been assessed with static human reliability analysis (HRA).

A significant influence on plant behavior and performance comes from the human operators who use that plant. The computational engine of the virtual plant model therefore needs to interface with a virtual operator that models operator performance at the plant. In current nuclear power plants (NPPs), most plant actions are manually controlled from the control room by reactor operators (ROs) or locally at the physical plant systems by field operators. Consequently, in order to have a non-idealized model of plant performance to support HUNTER, it is necessary to account for those human actions that ultimately control the plant. A high fidelity representation of an NPP absolutely requires an accurate model of its human operators in order to faithfully represent real world operation.

While it is tempting simply to script human actions at the NPP according to operating procedures, there remains considerable variability in operator performance despite the most formalized and invariant procedures to guide activities [3]. Human decision making and behavior are influenced by a myriad of factors at and beyond the plant. Internal to the plant, the operators may be working to prioritize responses to concurrent demands, to maximize safety, and/or to minimize operational disruptions. While it is a safe assumption that the operators will act first to maintain safety and then electricity generation, the way they accomplish those goals may not always flow strictly from procedural guidance. Operator expertise and experience may govern actions beyond rote recitation of procedures. As a result, human operators may not always make decisions and perform actions in a seemingly rational manner. Modeling human performance without considering the influences on the operators will only result in uncertain outcomes. To create the procedure step and timing based virtual operator model, the procedure steps were broken into subcomponents using an approach based on the task analysis technique called Goals, Operators, Methods, and Selection rules (GOMS). The approach, when adapted to HRA, is called GOMS-HRA [4].

2 GOMS-HRA Based Virtual Operator Model Development

Traditional HRA methods quantify tasks at an overall task level, which is adequate for static HRA purposes. However, computation-based HRA requires subtask level quantifications in order to model the virtual operator's actions in order to more accurately represent the scenario as it unfolds. Computation-based HRA affords a higher resolution of analysis and therefore it requires a finer grained quantification of operator tasks. To address this finer grained analysis, GOMS-HRA was created by categorizing subtasks as GOMS primitives and linking them to HEPs associated with each subtask [4]. These GOMS primitives were then assigned completion times based on empirical simulator studies and the subsequence timings for each subtask could then be mapped onto procedures steps to create a virtual operator model with variable completion times for each procedure step.

2.1 Goms-Hra

The GOMS model is a human cognitive model that provides analysts with a formal description of user behaviors as the user works with a human-machine interface (HMI) [5]. The GOMS acronym stands for Goals, Operators, Methods, and Selection Rules. *Goals* refer to the high-level objective users are attempting to accomplish, *Operators* are the individual actions that the user can make, *Methods* are collections of operators and consist of the steps and/or subgoals the human completes to move towards achieving the Goals, and *Selection* rules represent decisions the human makes to select particular methods and operators. GOMS is particularly useful for characterizing a user's procedural knowledge as they complete procedural tasks as evidenced by its extensive use in human factors. GOMS is categorized as a form of task analysis since the methods and operators form a hierarchical structure that can be used to complete subgoals and ultimately achieve the overall task goal. By categorizing particular types of actions, GOMS affords predictive capabilities that can be used to evaluate existing systems and model user interactions with human-computer interfaces. System designers have used GOM's predictive abilities to supplant user studies, though this use has been criticized for being time-consuming and overly laborious [6]. With the advent of discount usability methods centered on streamlined and cost-efficient data collection for user studies [7], the popularity of GOMS modeling as an alternative to such studies has subsided.

The Keystroke-Level Model (KLM) [8] is the simplest instantiation of the GOMS model. The KLM provides timing data for individual actions at the most fundamental level for interacting with a computer system. As such, this resolution of subtask decomposition into individual actions makes it possible to map human actions and predict how long certain activities, or collections of actions, will take to complete. This approach proved useful for routine and repetitive tasks like call center operations, where each scripted action was decomposed into its basic elements and then translated into its overall duration. Thus, it was possible to determine processes or even software use sequences that were inefficient. Through using the KLM, human factors researchers have been able to optimize human-computer interfaces. Indeed, such optimizations became the poster child of human factors, because it was easy to map the repetitive tasks to cost and thereby achieve cost savings with more efficient processes and interfaces. Usability engineering still lives under the shadow of the easy cost savings realized through KLM, whereas it can be difficult to cost-justify other human factors methods in comparison.

2.2 Defining GOMS-HRA Task Level Primitives

NPP operators activities are procedural in nature and therefore, the GOMS-HRA is well suited to serve as the foundation for the virtual operator model used in HUNTER. To populate the model with suitable operators, we examined HRA error taxonomies and selected the Systematic Human Error Reduction and Prediction Approach (SHERPA) [9] to derive Operators. In HRA, SHERPA is often used in conjunction with hierarchical task analysis to cluster subtasks into meaningful tasks suitable for defining human failure events [10].

The GOMS-HRA Operators generally distinguish between control room actions and field actions, the latter of which may be performed by ROs working as balance-of-plant operators or by technicians and field workers. Table 1 below depicts the task level primitives derived from the SHERPA error taxonomy.

Table 1. GOMS Operators used to define Task Level Primitives

Primitive	Description
Ac	Performing required physical actions on the control boards
Af	Performing required physical actions in the field
Cc	Looking for required information on the control boards
Cf	Looking for required information in the field
Rc	Obtaining required information on the control boards
Rf	Obtaining required information in the field
Ip	Producing verbal or written instructions
Ir	Receiving verbal or written instructions
Sc	Selecting or setting a value on the control boards
Sf	Selecting or setting a value in the field
Dp	Making a decision based on procedures
Dw	Making a decision without available procedures
W	Waiting

These Operators, derived from the SHERPA error taxonomy, can serve as the task level primitives which can be mapped onto the procedure steps followed by the operators to providing timing information relevant to Human Error Probabilities (HEPs).

2.3 Assigning Timing Values to the GOMS-HRA Task Level Primitives

Task primitive completion times were quantified based on empirical data collected during a series of operator-in-the-loop studies conducted as part of a separate control room modernization project [11]. The empirical data consists of simulator logs recorded by an observer shadowing a crew of operators during a series of turbine control scenario simulations. The simulator logs provided a detailed account of each procedure step, relevant actions, completion times for those actions, and crew communications. The simulator logs contained a total of 283 observations spanning five separate scenarios, each of which lasted approximately half an hour. Though the scenarios were specific to turbine control, the task primitive timing data extracted from the simulator logs represent universal actions that are applicable throughout the entirety of the main control room interfaces.

3 Operator Timing Station Blackout Example

A station blackout (SBO) represents a time sensitive worst case scenario in which offsite power and the diesel generators fail. This forces the plant to rely on battery backup power until these other two more reliable power sources can be restored to maintain the necessary cooling to prevent core damage. The station blackout scenario used to illustrate the GOMS-HRA method contains procedure steps containing specific verb terminology. Procedure writing guidelines suggest following the convention of consistently using a single verb to denote a particular action. Operators are trained to interpret the verb during training so that each procedure step is clearly defined and intuitive for the operator to complete. We followed the standard conventions to define each verb used in each procedure step of the Post Trip Actions (PTA) and Station Blackout procedures [12, 13] used in this example. Defining the verbs with standardized definitions enables the HRA task primitives to map onto each specific procedure step and provide timing data. Each verb represents a single primitive or a series of combined primitives required to complete the procedure step. At each step in the procedure, the realistic plant model is provided with the appropriate timing and Human Error Probability (HEP) data to more accurately reflect the true nature of the event in contrast to more traditional state HRA approaches.

The *procedure level primitive* (PLP) used within each procedure step represents a cluster of actions that must occur in the proper sequence in order for the operator to successfully complete the step [14]. These procedure level primitives can be decomposed into sequences of *task level primitives* (TLPs) for complex mappings. To check a value in a procedure step, a series of activities is performed. After reading and interpreting the procedure step, the operator walks to the board and looks for the required information. If the expected value or state is observed, the operator verbally conveys the value or state to the RO and the sequence of primitives concludes. If the expected value or state is not observed, the operator then must take corrective actions by setting a state or specific value and waiting for those actions to take effect. The sequence of task level primitives repeats until the desired value or state is achieved and the step is concluded. The task level primitives were mapped following this method for each procedure step in order to support the estimation completion times for each step.

Each action verb contained within a procedure step is decomposed into one or more task level primitives. Some procedure level primitives are comprised of multiple task level primitives, while others represent a single action. The procedure level primitives, i.e. action verbs, are generically defined in order to support mapping onto any procedure steps used in the main control room. The action verbs or procedure level primitives can be further broken down into task level primitives depending upon the context of the action verb for a given procedure step.

3.1 Defining Nominal Timing Data for HEPs

In order to analyze a specific scenario, such as the station blackout event, and calculate the nominal HEP and task timing values, the procedure must be evaluated at the procedure level and then at the task level. The procedures included in this simulation are based on the post trip action and station blackout procedures from a nuclear utility. To protect the proprietary procedures, the procedure text cannot be publicly disseminated. Since the procedure steps cannot be shared, an example procedure step in Table 2 serves to provide an overview of how a step is mapped to the procedure level and task level primitive. For example, procedure step 2 of the post trip action (PTA) procedure contains two procedure level primitives, which are *determine* and *verify*. Determine is an abstract procedure level primitive that can be decomposed into three verify substeps. These substeps of procedure step 2 are mapped onto the task level primitive of verify, which corresponds to the task level primitive, C_C, or looking for required information on the control boards.

Table 2. Example mapping of procedure step to procedure and task level primitives for a post-trip action (PTA) procedure.

PTA	2	–	Determine maintenance of vital auxiliaries acceptance criteria are met	Determine	–
PTA	2	a	Verify the main turbine is tripped	Verify	C_C
PTA	2	b	Verify the main generator output breakers are open	Verify	C_C

To reiterate the process, two mappings are involved:

- The plant procedures are classified in terms of procedure level primitives
- These procedure level primitives are comprised of task level primitives from GOMS-HRA.

Because there is a high degree of nuclear industry consensus on terminology in operating procedures, the procedure level primitives represent commonly and consistently deployed types of activities. It is therefore possible to create a universal mapping of GOMS-HRA task level primitives to the procedure level primitives. This universal mapping affords the opportunity for reuse of the building blocks in HUNTER across different analyses.

The procedures used in this station blackout example are an approximation of the actual series of events that would unfold during the scenario. Though this reduces some of the realism captured in the simulation, it was necessary due to the procedures' proprietary nature. Furthermore, this is the first attempt at performing an integrative HRA model with dynamic HEPs and corresponding thermohydraulic computations, which was made possible by restricting the scope of the simulation to these the post trip actions (PTA) and station blackout (SBO) procedures. To illustrate this analysis further, SBO procedure step 5a stating "Ensure letdown is isolated" will be described at each stage of the analysis process (see Tables 3 and 4). The procedure level primitive in this step is defined as the verb, *Ensure*. Ensure could be decomposed into different task level primitives, so the context of the procedure step, in this case letdown isolation, must be evaluated to determine which of the task level primitives are applicable. In this instance, the valve positions are a status indicator with a simple state control as opposed to a continuous numerical value setting. As a result, this procedure level primitive translates to the task level primitives of C_C (look for required information on the control board) and A_C (perform physical actions on the control board).

Table 3. SBO Step 5 showing mapping of *Ensure* procedure level primitive.

SBO	5	–	Minimize reactor coolant system leakage	Minimize	–
SBO	5	a	Ensure letdown is isolated	Ensure	C_C
SBO					A_C
SBO	5	b	Ensure reactor coolant pump controlled bleedoff is isolated	Ensure	C_C
SBO	5	c	Ensure reactor coolant system sampling is isolated	Ensure	C_C

The procedure steps for the PTA and SBO procedures were mapped to procedure and task level primitives as shown in Table 3. Following the analysis of the procedures to map procedure level and task level primitives, timing data were estimated for each procedure step as derived from GOMS-HRA (see Table 4). Additionally, the procedure steps were aligned with the two primary events in which the loss of offsite power occurs and the loss of diesel generators, and loss of battery during the station blackout event. By aligning the procedure steps to the primary events of the scenario, the simulation can be repeated with slight variations on the timing for the operator to expeditiously complete procedure steps. This provides insight into the effects of the operators taking longer or shorter to complete task and can aid the identification of key steps that either lead to success if completed swiftly or place the plant in danger of core damage if completed too slowly. The different timings modelled in each repetition of the same scenarios simulation are drawn from the empirically calculated values for each task primitive as depicted in Table 4. For example, the amount of time to complete a physical action on the control board, Ac, took on average 18.75 s, but could be

Table 4. Average, 5th percentile and 95th percentile time (seconds) for completing each GOMS task level primitive as calculated by mapping the task level primitives to the empirical simulator log data for completing procedure steps.

Task level primitive	5th percentile	Time (seconds)	95th percentile
Ac	1.32	18.75	65.26
Cc	2.44	11.41	29.88
Dp	2.62	51	152.78
Ip	3.35	15.56	40.66
Ir	1.47	10.59	31.84
Rc	3.08	9.81	21.90
Sc	3.01	34.48	115.57
W	1.79	14.28	113.61

vary in the completion time used for a given simulation run between 1.32 and 65.26 s in order to capture the variability in the amount of time operators empirically require to complete this task level primitive.

4 Conclusions

The development of the virtual operator with a variable timing component for completing procedure steps adds a level of fidelity to HUNTER that was nonexistent before. The effort to map the procedure steps with procedure level and task level primitives was substantial. Mapping the procedure steps with procedure and task level primitives is time consuming. Future steps involve automating this process with statistical techniques to extract the verbs used in a plants entire catalogue of procedures and create a larger database of procedure primitives and their associated task level primitives in order to apply this method to other scenarios. This process also requires more empirical data on actual operator timing to further enhance the accuracy of the timing data used to assign timing values for each task level primitive. Though expanding the database for automated procedure mapping of procedure and task level primitives represents a substantial research endeavor, it is quite advantageous due to the enormous benefit of being able to model unexampled events and more accurately conduct HRA.

Acknowledgments. The authors would like to thank Cutis Smith and Diego Mandelli of the Idaho National Laboratory for their exceptional contributions to this research. This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U. S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

References

1. Boring, R., Mandelli, D., Joe, J., Smith, C., Groth, K.: A Research Roadmap for Computation-Based Human Reliability Analysis, INL/EXT-15-36051. Idaho National Laboratory, Idaho Falls (2015)
2. Joe, J.C., Shirley, R.B., Mandelli, D., Boring, R.L., Smith, C.L.: The development of dynamic human reliability analysis simulations for inclusion in risk informed safety margin characterization frameworks. Procedia Manufact. **3**, 1305–1311 (2015)
3. Forester, J., Dang, V.N., Bye, A., Lois, E., Massaiu, S., Broberg, H., Braarud, P.Ø., Boring, R., Männistö, I., Liao, H., Julius, J., Parry, G., Nelson, P.: The International HRA Empirical Study. Lessons Learned from Comparing HRA Methods Predictions to HAMMLAB Simulator Data, NUREG-2127. U.S. Nuclear Regulatory Commission, Washington, DC (2014)

4. Boring, R.L., Rasmussen, M.: GOMS-HRA: a method for treating subtasks in dynamic human reliability analysis. In: Proceedings of the 2016 European Safety and Reliability Conference (2016)
5. Card, S.K., Newell, A., Moran, T.P.: The psychology of human-computer interaction (1983)
6. Rogers, Y., Sharp, H., Preece, J.: Interaction Design. Wiley, New York (2002)
7. Nielsen, J.: Usability engineering at a discount. In: Salvendy, G., Smith, M.J. (eds.) Designing and Using Human-Computer Interfaces and Knowledge Based Systems, pp. 394–401. Elsevier Science Publishers, Amsterdam (1989)
8. Card, S.K., Moran, T.P., Newell, A.: The keystroke-level model for user performance time with interactive systems. Commun. ACM **23**(7), 396–410 (1980)
9. Stanton, N.A., Salmon, P.M., Rafferty, L.A., Walker, G.H., Baber, C.: Human Factors Methods: A Practical Guide for Engineering and Design, 2nd edn. Ashgate Publishing Co., Aldershot (2013)
10. Boring, R.L.: Defining human failure events for petroleum applications of human reliability analysis. Procedia Manufact. **3**, 1335–1342 (2015)
11. Boring, R., Lew, R., Ulrich, T., Joe, J.: Light Water Reactor Sustainability Program Operator Performance Metrics for Control Room Modernization: A Practical Guide for Early Design Evaluation, INL/EXT-14-31511. Idaho National Laboratory, Idaho Falls (2014)
12. Procedure Professionals Association: Procedure Writers' Manual (PPA AP-907-005, Rev. 2) (2016)
13. Jang, T.I., Lee, Y.H., Park, J.C., Jeong, Y.S., Chu, G.I.: Standardization of the Action Verbs used in Emergency Operating Procedures of NPPs (2010)

A Systematic Method to Build a Knowledge Base to be Used in a Human Reliability Analysis Model

Yunfei Zhao(✉), Keith Altman, Kreteeka Chaudhury,
Muhammad Anandika, and Carol Smidts

Department of Mechanical and Aerospace Engineering,
The Ohio State University, Columbus, OH 43210, USA
`{zhao.2263,smidts.1}@osu.edu`

Abstract. A human's knowledge base is a key component for the development of a mechanistic model of human response to be used for human reliability analysis. This paper proposes a new method for constructing this knowledge base. The proposed method is comprised of three steps: (1) systematic literature review, which is used to collect data pertinent to the subject under study; (2) summarization, the goal of which is to extract key points that are expressed in the literature; (3) qualitative coding, a process in which codes closely related to the topic are derived and the relationships between these codes are expressed. As a case study, the proposed method is being applied to construct an operator's knowledge base concerning severe accident phenomenology in a nuclear power plant. Part of this application is explored in this paper. With the proposed method and the resulting knowledge base, it is expected that an individual's response when presented with a specific context can be modeled in more detail.

Keywords: Human reliability analysis · Mechanistic model · Knowledge base · Knowledge representation · Severe nuclear accident

1 Introduction

We are currently witnessing an increase in the complexity and scale of technological systems, examples of which include nuclear power plants and air traffic control systems. These trends herald greater economic gains due to accompanying economies of scale and enhanced technical features. However, these complex systems may also be subject to an escalating number of incidents, mishaps, or accidents, as unexpected failures are more likely to occur since these systems will be more difficult to operate. The Three Mile Island accident and the Bhopal disaster are just a few examples of such accidents [1]. As a result, both governmental entities and utilities themselves are putting more emphasis on safety and reliability during design and operation. This emphasis also translates into the creation of effective mitigation strategies that would minimize the consequences of an incident if one were to occur.

Human response is a major component involved with the implementation of these mitigation measures, which poses a problem when considering the difficulty that modern

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_6

day technological systems pose to their human counterparts. Therefore, the call for credible methods of human reliability analysis (HRA) has become prevalent. Nuclear power plants represent large and complex technological systems, for which a number of HRA methods have been proposed, including THERP [2], SPAR-H [3], etc. These methods have been widely used in nuclear power plants probabilistic risk assessments due to their ease of use. Most of these methods use a set of performance shaping factors (PSFs), which represent the specific context of an accident, to adjust the probability of human error. However, these methods have their limitations. For example, it is not possible to represent a context with a large number of distinct dimensions via a few PSFs. Human response when presented with a specific context is also more complex to represent than what can be achieved by the simplified models. Therefore, mechanistic models, which are based upon state-of-the-art research on human cognition and behavior, have emerged in the field due to their ability to describe human response in higher fidelity. These include models such as IDA [4] and IDAC [5].

A human's knowledge base, which stores information related to various aspects of a nuclear power plant, is a key element in such type of mechanistic models. It provides the basis for an individual's diagnosis of the system state and for the planning of corresponding actions, which are two important components of human response [3]. Therefore, the knowledge base should contain the cause-effect relationships between the system state and the corresponding phenomena or parameters. This knowledge provides the basis for a diagnosis. It should also contain alternative actions in face of a specific system state and the features of the actions such as the cost, the availability, and the potential impact. This knowledge provides the basis for planning human actions.

However, there has been little research on how to construct the knowledge base. In IDA, an operator's knowledge base is derived from Emergency Operating Procedures and represented using *if-then* rules. In the latest version of IDAC [6], expert questionnaires are used to collect related information. The operator's knowledge base is derived from the questionnaires. It consists of the cause-effect relationships between systems and monitoring signals and the weights assigned to each link in the network. Learning algorithms used to construct an operator's knowledge base from historical operational data have also been proposed in the literature, but these algorithms have not been used in the context of nuclear power plants and the main goal of these studies has differed greatly from that of this paper [7].

In this paper, a systematic method to construct a knowledge base is proposed. This method consists of three steps. First, papers and reports related to the topic of interest are collected. They serve as the original data set of the knowledge base. Second, each of the collected items is summarized to extract the key points of each paper in preparation for the next step. Third, codes closely related to the topic under study and their relationships are derived from the summaries. The proposed method is applied to construct an operator's knowledge base for severe accident management in nuclear power plants.

The paper is arranged as follows. In Sect. 2, the proposed method is explained in detail. In Sect. 3, said method is applied to the topic of severe accident management in nuclear power plants. The method and the resulting knowledge base are discussed in Sect. 4. The paper is concluded in Sect. 5 with a summary of the paper and suggestions on future research.

2 Method

One of the issues during the construction of a knowledge base is the great variation and lack of repeatability between different modelers. To address this issue, at least to some extent, a new method is proposed in this section, which can be implemented systematically, reducing the variation caused by the inconsistencies amongst different researchers' level of understanding. As stated in Sect. 1, the method is comprised of three steps, as shown in Fig. 1. Each step will be described in further detail in the following sub-sections.

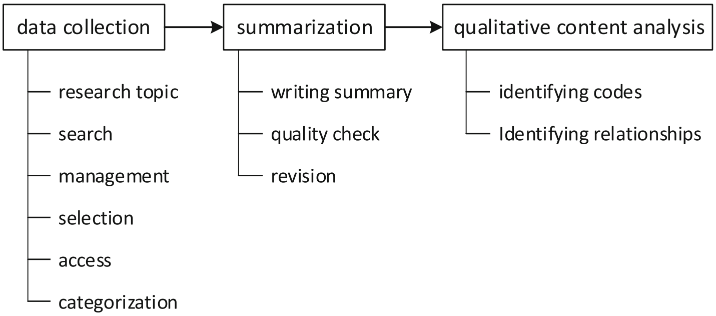


Fig. 1. General process followed to construct the knowledge base.

2.1 Data Collection

Different from previously used data sources such as operating procedures, expert questionnaires, and historical data, the data used to construct the knowledge base is extracted from the published literature, which includes papers and reports related to the topic under study. To collect the literature in a trustworthy, rigorous, and auditable way, a systematic literature review is adopted. This technique has been widely applied in software engineering [8], and serves as a means of evaluating and interpreting all available research relevant to a particular research question, topic area, or phenomenon of interest [9]. The method designed initially for software engineering has been adapted slightly in this paper to achieve the objective of data collection. It consists of three steps which are depicted in detail below.

In the first step, the research that is relevant to the knowledge base to be constructed, is identified. This includes (1) the identification of the research topic, i.e. the topic of the knowledge base; (2) searching primary studies relevant to the identified topic from databases; (3) documentation of the search result. For example, the research topic can be a specific accident initiator such as loss of coolant accident (LOCA), a specific phenomenon during an accident like fission product behavior, or any other topic that the modeler wants to include in the knowledge base. As for the literature search, to keep the process transparent and replicable, the information pertaining to the database such as the name, and the search strategy (keywords used, etc.), need to be recorded. Then the search result can be documented in a specific format, which consists of information such as the title of the publication, publication year, etc.

In the second step, the primary studies are selected and access to each selected item is obtained. Before incorporation of the paper in the final data set, the selection criteria need to be defined. These could be whether the content in the literature is actually relevant to the research topic, the type of the paper/report, and its date, etc. As to the question of access, most of the journal and conference papers can be obtained through University Library Databases. For other types of literature, such as technical reports, access may be obtained through the publishing organization.

In the last step, all the obtained items are categorized based on specific content. For example, for a knowledge base focused on fission product behavior during an accident, the related papers and reports can be categorized according to specific fission products. In addition, the items in each main category can be further divided into lower level categories, such as the different forms the fission product may take. This step aims to facilitate the following analyses, and to enable the modeler to have a better knowledge of the research areas covered by the selected papers and reports.

2.2 Summarization

The search conducted in Sect. 2.1 will generate a large volume of references. To reduce the workload of the analysis, each obtained item is summarized before further analysis. Information extraction in this step serves as the bridge between the original data and a more detailed analysis in the next step. The purpose of this step is to extract the key points presented in the paper or report. These key points are usually found in the abstract and conclusion sections of the document. However, the analyst should also pay attention to the points raised in the intermediate sections of the document for a more detailed understanding of the material.

To ensure the quality of the summaries, quality checks are necessary once the summaries have been completed. With respect to quality, this paper focuses on how consistently a summary can be created between two different authors from the same original document. The procedure for checking the quality of the summaries is described as follows. A number of documents, the lower limit of which is usually 10% of all the documents [10], are sampled from each category and used as the samples for quality check. Then for each document, the summary is rewritten by a different author, without referring to the existing summary drawn by the first author. The two summaries for document i are denoted as S_{io}, the original summary, and S_{ir}, the rewritten summary. The key points in S_{io} and S_{ir} are compared by a third author, and the consistency level for document i is assessed through the following formula:

$$C_i = \left(1 - \frac{\frac{m_{io}}{N_{io}} + \frac{m_{ir}}{N_{ir}}}{2}\right) \times 100\%. \tag{1}$$

In Eq. (1), C_i is the consistency of document i, $N_{i\mathrm{o}}$ and $N_{i\mathrm{r}}$ are the total number of key points extracted from S_{io} and S_{ir} respectively, and $m_{i\mathrm{o}}$ and $m_{i\mathrm{r}}$ are the number of key points that are different in S_{io} compared with S_{ir}, and in S_{ir} compared with S_{io} respectively.

In the end, the average consistency level of the documents in each category is obtained. This quality check needs to be compared with a predefined criterion, for example a level of consistency of 80%. If the consistency of the documents in one category is below this criterion, the summaries need to be revised and reassessed until the consistency criterion is met.

2.3 Qualitative Content Analysis

The summaries created in Sect. 2.2 are analyzed in further detail in this step. As stated in [6], a human's knowledge base can be represented as a semantic web. In this paper, qualitative content analysis is adopted and adapted to extract the relevant codes or concepts and their relationships which together constitute this semantic web.

Qualitative content analysis is a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns [11]. Its goal is to provide knowledge and understanding of the phenomenon under study. Qualitative content analysis has been widely used in health studies. Based on whether prior knowledge is available, different qualitative content analysis approaches may be used. When prior knowledge of the topic under study is limited, inductive [12] or conventional [11] analysis is usually used, where the codes are derived from the text. However, when prior knowledge exists, deductive [12] or directed [11] analysis is usually adopted, which allows the analyst to identify the concepts in the text based on existing ones. The main processes in qualitative content analysis include reading the text thoroughly and understanding the content, identifying the codes based on the analyst's understanding or the existing codes, and categorization of the codes.

It needs to be noted that different from the application in health studies, where the main goal is to identify codes and then categorize them, the application of qualitative content analysis in this paper is more complex. More specifically, the relationships between the codes, especially the logical relationships, constitute another important part of a human's knowledge base. Therefore, traditional qualitative content analysis is adapted in this paper and goes beyond merely coding and categorization in order to examine the text more carefully for the purpose of identifying the relationships between the codes. The relationships can be roughly divided into two types. The first type is an affiliation relationship and is generally simpler, for example one component *is part of* a system. The second type is a logical relationship and is more complex, for example *if* a LOCA occurs *then* the primary system pressure decreases.

Each summary can be analyzed separately, and the codes and relationships in individual summaries are then combined to form a full knowledge base, which contains all of the extracted codes and their relations.

3 Application to Nuclear Severe Accidents

For illustration, the method proposed is applied to the construction of a knowledge base of severe accidents in nuclear power plants. The analyses and partial results in each step are described as follows.

3.1 Data Collection

Severe accident and *nuclear* were used as keywords to search relevant literature through Google Scholar. Four hundred and eighty-six relevant items were retrieved in total, and their basic information stored in MS Excel. The information stored includes: title, publication venue, publication year, authors of the publication, affiliation of the authors, keywords of the publication, date of search, number of citations, abstract, URL, and contact information of the authors. Among the 486 items selected, 281 items were accessible and could be downloaded. All these items were kept as the original data set for the knowledge base. The 281 items were divided into 11 main categories based on a preliminary analysis of their contents. The items in each main category were further categorized into two lower level tiers. The number of publications in each main category and the secondary categories of *accident management* are shown in Fig. 2, and the number of items published in each year is shown in Fig. 3. The variety of data collected ensures that diverse research on the topic of nuclear severe accident can be covered in the constructed knowledge base.

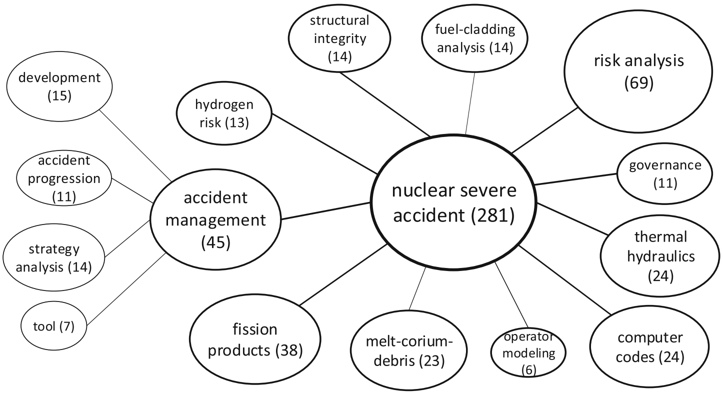


Fig. 2. The number of publications in each topical category.

3.2 Summarization

The 281 items were assigned to three undergraduate students to extract information from and write a short summary for each item. Before writing the summaries, they were trained by an experienced postdoctoral researcher to make sure that they understood the objective of this step. The postdoctoral researcher also taught them how to extract

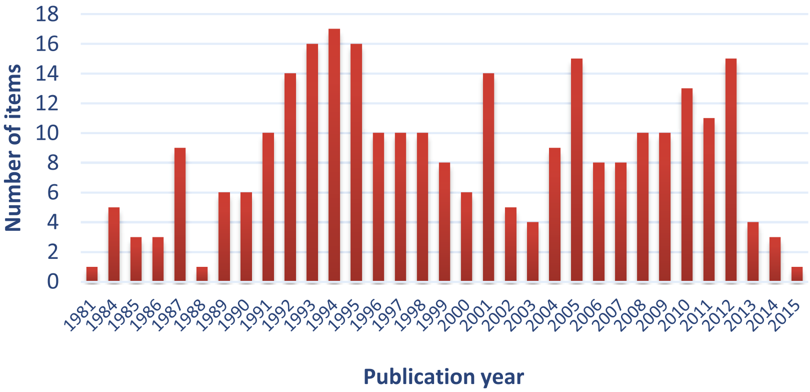


Fig. 3. The number of items published in each year.

information and write an effective summary. The undergraduate students' summary writing capabilities were then tested using several examples and evaluated by the postdoctoral researcher until the summaries captured all the key points in the original publications. After the training and testing phase, the undergraduate students were believed to have sufficient expertise to write effective summaries. Then each student was assigned a specific number of topical categories.

After all the summaries were completed, a quality check was conducted. At least 10% of the items in each category were randomly chosen, which generated 33 items in total. Then the randomly chosen items that were assigned to one student initially were now assigned to a different student. Lastly, the rewritten summaries were compared with the existing summaries to assess their consistency.

The blue columns in Fig. 4 show the consistency level of each category from the first round summaries. The last blue column shows the overall average consistency level. Although the overall average consistency is 82%, the consistency level of some categories is low. For example, the consistency level of the governance category is below 40%.

Based on the evaluation of the degree of consistency, the items in four categories, namely governance, hydrogen risk, fuel cladding, and melt-corium-debris were revised by the students, after an in-depth discussion of the problems encountered while drafting the first round of summaries. A quality check was conducted again for the revised summaries, in which process nine items in total were sampled randomly. The green columns in Fig. 4 show the consistency levels of the four revised categories. The last green column shows the overall average consistency level after the revision. It can be seen that the consistency levels for the four categories improve significantly and the average consistency level is 91%, which indicates the high quality of the summaries.

3.3 Qualitative Content Analysis

Qualitative content analysis is being conducted for each summary in this step. First, the summary is read thoroughly and the relevant codes or concepts in the text are

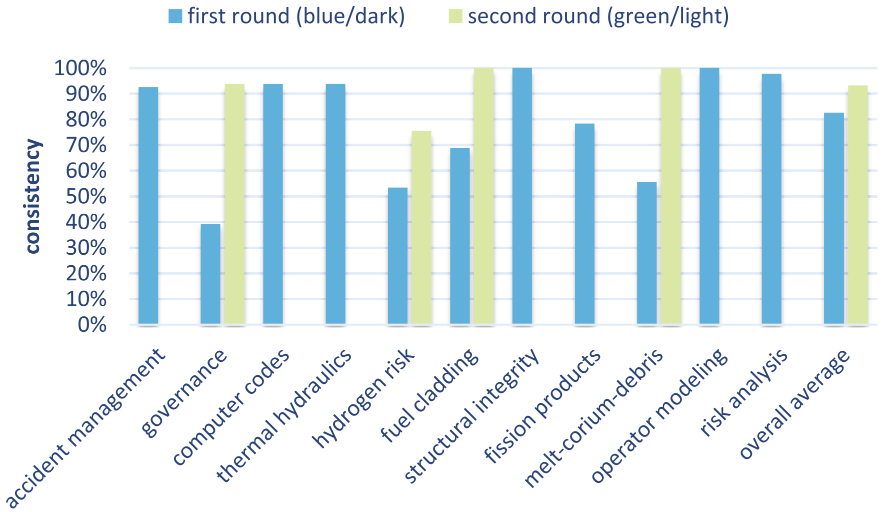
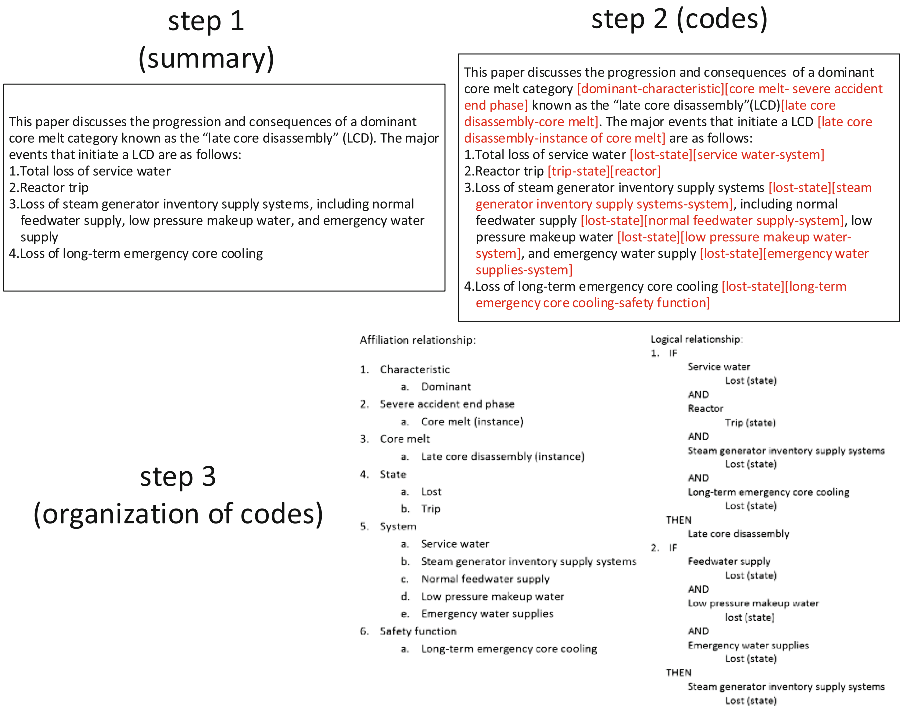


Fig. 4. Result of quality check of summaries.



step 1
(summary)

This paper discusses the progression and consequences of a dominant core melt category known as the "late core disassembly" (LCD). The major events that initiate a LCD are as follows:
1.Total loss of service water
2.Reactor trip
3.Loss of steam generator inventory supply systems, including normal feedwater supply, low pressure makeup water, and emergency water supply
4.Loss of long-term emergency core cooling

step 2 (codes)

This paper discusses the progression and consequences of a dominant core melt category [dominant-characteristic][core melt- severe accident end phase] known as the "late core disassembly"(LCD)[late core disassembly-core melt]. The major events that initiate a LCD [late core disassembly-instance of core melt] are as follows:
1.Total loss of service water [lost-state][service water-system]
2.Reactor trip [trip-state][reactor]
3.Loss of steam generator inventory supply systems [lost-state][steam generator inventory supply systems-system], including normal feedwater supply [lost-state][normal feedwater supply-system], low pressure makeup water [lost-state][low pressure makeup water-system], and emergency water supply [lost-state][emergency water supplies-system]
4.Loss of long-term emergency core cooling [lost-state][long-term emergency core cooling-safety function]

step 3
(organization of codes)

Affiliation relationship:

1. Characteristic
    a. Dominant
2. Severe accident end phase
    a. Core melt (instance)
3. Core melt
    a. Late core disassembly (instance)
4. State
    a. Lost
    b. Trip
5. System
    a. Service water
    b. Steam generator inventory supply systems
    c. Normal feedwater supply
    d. Low pressure makeup water
    e. Emergency water supplies
6. Safety function
    a. Long-term emergency core cooling

Logical relationship:
1. IF
    Service water
        Lost (state)
    AND
    Reactor
        Trip (state)
    AND
    Steam generator inventory supply systems
        Lost (state)
    AND
    Long-term emergency core cooling
        Lost (state)
  THEN
    Late core disassembly
2. IF
    Feedwater supply
        Lost (state)
    AND
    Low pressure makeup water
        lost (state)
    AND
    Emergency water supplies
        Lost (state)
  THEN
    Steam generator inventory supply systems
        Lost (state)

Fig. 5. Illustration of the qualitative content analysis- process.

identified. One can iterate through this process for several rounds until no important code is missing. Then the relationships between the codes are identified based on the understanding of the content of the text. The work on this section is still ongoing, but some results have been obtained. The qualitative content analysis process is illustrated with one publication [13] in Fig. 5, and part of the results are shown as a network in Fig. 6.

The codes in the example network (Fig. 6) consist of the main systems or components in the nuclear power plant, the concepts related to accident like core melt, and the different states of the systems. The two types of relations between the codes were also identified and included in the example network. For example, *state* and *instance* belong to the first type, affiliation relationship, and *if-then* belongs to the second type, logical relationship.

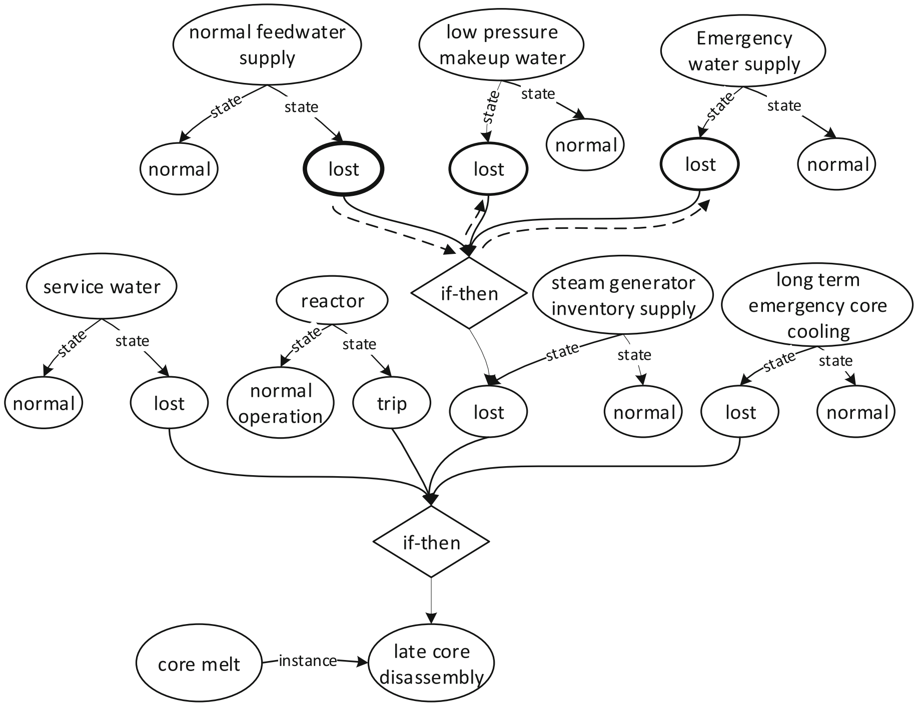


Fig. 6. Illustration of the qualitative content analysis- result.

4 Discussion

Compared to other methods that rely too much on expert opinion, the systematic method proposed in this paper has two main advantages. First, it decomposes the process of extracting relevant information from large volumes of datasets into three steps: data collection, summarization, and qualitative content analysis. The process in each step can be recorded in detail to maintain transparency and to improve repeatability, and the result in each step can be assessed so that the quality of the analysis in each step is assured. This advantage ensures that the analysis can be conducted in a systematic way. The second advantage is that the proposed method reduces the adverse effect of an expert's subjectivity to a great extent. By applying a systematic literature review for data collection, a variety of areas related to the topic under study are selected rather than the areas which the expert is familiar with. Qualitative content analysis ensures that the knowledge base is grounded in a rigorous analysis of the content, although it is partly based on the modeler's understanding as well. Through the proposed method, the consistency between knowledge bases constructed by different modelers can be improved.

In addition to the proposed method itself, the resulting knowledge base exhibits some advantages in modeling human reasoning. The final knowledge base constructed through the proposed method is a form of semantic network, which is comprised of concepts and their relationships. In addition to being used for human reasoning through the various relationships included, the knowledge base can also be used to propagate activation [14] easily, which is illustrated through an example in Fig. 5. In this example, the reasoning process is initiated first by the loss of normal feedwater supply, which has the highest activation level and is marked by the circle with thickest border. Then its activation spreads first to the *if-then* logic gate, shown as the dashed line in Fig. 5, and then through the logic gate to the state of the other two systems: low pressure makeup water and emergency water supply. After this process, the activation levels of the state of the two related systems are increased, which are marked by circles with lighter borders than the initiating node, but thicker than all the other nodes in the figure. Activation propagation in this example directs the operator to check the state of relevant systems.

5 Conclusion and Future Research

This paper proposes a systematic method to construct an operator's knowledge base, which is a key component of a mechanistic model enabling a more detailed human reliability analysis. The procedures and quality assurance in each step of the method improve the repeatability of the process and reduce the adverse effect of expert subjectivity. The proposed method is illustrated with an example application of the construction of an operator's knowledge base for nuclear severe accident regime, with part of the results in each step shown. Based on the results obtained, future research is expected in the following areas to make the method more effective. First, artificial intelligence techniques may be introduced in the process to reduce the workload. For example, natural language processing methods may be used to perform a preliminary analysis of the text. Second, research on human reasoning may serve as the prior knowledge in qualitative content analysis. For example, logical relationships that are essential in human reasoning may be included before initiating qualitative content analysis to guide extraction of similar logical relationships in the text.

Acknowledgments. This research is funded by U.S. Nuclear Regulatory Commission under Grant NRC-HQ-60-15-G-0002.

References

1. Meshkati, N.: Human factors in large-scale technological systems' accidents: three mile Island, Bhopal. Chernobyl Ind. Crisis Q. **5**(2), 133–154 (1991)
2. Swain, A.D., Guttmann, H.E.: Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report, Sandia National Labs., Albuquerque, NM, USA (1983)
3. Gertman, D., Blackman, H., Marble, J., Byers, J., Smith, C.: The SPAR-H Human Reliability Analysis Method. US Nuclear Regulatory Commission (2005)
4. Smidts, C., Shen, S.H., Mosleh, A.: The IDA cognitive model for the analysis of nuclear power plant operator response under accident conditions. part I: problem solving and decision making model. Reliab. Eng. Syst. Saf. **55**(1), 51–71 (1997)
5. Chang, Y.H., Mosleh, A.: Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents: part 1: overview of the IDAC model. Reliab. Eng. Syst. Saf. **92**(8), 997–1013 (2007)
6. Li, Y.: Modeling and simulation of operator knowledge-based behavior. Ph.D. dissertation, University of Maryland, College Park (2013)
7. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Syst. Man Cybern. SMC **15**(1), 116–132 (1985)
8. Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering–a systematic literature review. Inf. Softw. Technol. **51**(1), 7–15 (2009)
9. Kitchenham, B.: Procedures for performing systematic reviews. Technical report, Keele University (2004)
10. Huang, F., Smidts, C.: Causal mechanism graph – a new notation for capturing cause-effect knowledge in software dependability. Reliab. Eng. Syst. Saf. **158**, 196–212 (2017)
11. Hsieh, H.F., Shannon, S.E.: Three approaches to qualitative content analysis. Qual. Health Res. **15**(9), 1277–1288 (2005)
12. Elo, S., Kyngäs, H.: The qualitative content analysis process. J. Adv. Nurs. **62**(1), 107–115 (2008)
13. Dick, J.E., Nath, V.I., Kohn, E., Min, T.K., Prawirosoehardjo, S.: Event sequence of a severe accident in a single-unit CANDU reactor. Nucl. Technol. **90**(2), 155–167 (1990)
14. Anderson, J.R.: A spreading activation theory of memory. J. Verbal Learn. Verbal Behav. **22**(3), 261–295 (1983)

Bridging Human Factors and Human Error

A Literature Study to Explore Empirically: What Is the Scientific Discipline of Human Factors and What Makes It Distinct from Other Related Fields

Karin Laumann[1(✉)], Martin Rasmussen[1], and Ronald L. Boring[2]

[1] Department for Psychology, Norwegian University of Science
and Technology (NTNU), Dragvoll, 7491 Trondheim, Norway
`karin.laumann@ntnu.no`
[2] Human Factors, Controls, and Statistics, Idaho National Laboratory,
P.O. Box 1625, Idaho Falls, ID 83415, USA

Abstract. The aim of this paper is to investigate which topics are studied within human factors, what are the "levels" studied (individuals, work group, organizations, societies), and which methods are used. The questions were answered by investigating 183 papers published in the *Human Factors* journal for 2015 and 2016. The results showed that more than five papers included the topics; car driving, physical workload, human-automation interaction, design and usability, human machine interface (displays, controls and alarms), mental workload, cognition, team work, training/simulations, and anthropometry. The topics that seem to be unique for human factors are all the topics that are about human-computer/technology interactions and the topic of design and usability. Experiments are the main method used in human factors and almost all of the studies are at the individual level.

Keywords: Human factors · Literature study · Human computer interaction · Human-Technology interaction · Design and usability

1 Introduction

The most often cited definition of human factors is The International Ergonomic Association's definition (IEA) [1]: "Ergonomics (or human factors) is the scientific discipline concerned with the understanding of the interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance. Ergonomists contribute to the design and evaluation of tasks, jobs, products, environments and systems in order to make them compatible with the needs, abilities and limitations of people".

From our experience, when this definition is presented to students, researchers from other fields and others they do not do seem to understand the definition. It seems especially difficult to understand what makes the scientific field of human factors different from other fields.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_7

There might be different reasons why this definition is difficult to understand. The word “system” is used twice in this definition. A system could be defined as something that consists of interacting or interdependent components [2, 3]. With this definition, a system could thus be everything from a cell, to a telephone, to a society. In addition, when the use of the words “element of a system” are added to the definition, the definition says that human factors study the interaction among humans and everything else that might exist.

Also Hollnagel [3] describes the difficulties with using the word system in the definition of human factors. Hollnagel [3] describes that a system could also be defined from a system theoretic or cybernetic perspective as “a set of mutually dependent functions, hence characterized by what it does rather than what it is.” This definition does not increase the clarity of the word “system” since it is equally difficult to understand which functions we refer to within human factors.

In addition, the human factors definition says that the purpose of human factors is to optimize human well-being and overall system performance. It is difficult to find any study in the entire field of psychology that is not concerned with either “well-being” or “performance” or both. However, this seems to apply for all sciences where humans are the main object of study, which applies for most of the social sciences (anthropology, sociology, and pedagogics), medicine, physiotherapy, and business management science.

The Human Factors and Ergonomic Society have collected several definitions of human factors. What characterizes all of these definitions is that they are so broad that they include the entire social science field with words like “systems” and “performance.”

There has been some concern that human factors is not used as much as it should when for example new products are being developed for example [4]. Part of this problem could be that the field of human factors is not defined or described in a way that is easy to communicate and understand for people from other fields like engineering or psychology. The purpose of this paper is to: (a) investigate the questions presented below and to propose a more understandable description of what human factors is based on human factors literature, and (b) describe how human factors is different from other fields.

The questions investigated in this paper are:

a. Which topics are studied within human factors?
b. Which methods are used in human factors research?
c. What are the “levels” studied (individuals, work group, organizations, societies)?

To answer these questions, a literature review was performed.

2 Method

In this study we chose to investigate all papers from 2015 and 2016 in the journal, ‘*Human Factors: The Journal of the Human Factors and Ergonomics Society*.’ This journal was selected, because of its name and because it has a broad definition of human factors that resembles the definition presented in the introduction of this paper. In addition, this is a highly-ranked journal. We argue that the last two years of papers in this journal should give us an answer to which topics are included within contemporary human factors, which methods are used, and which levels (individuals, work group, organization, society) are investigated.

In all 183 papers were selected for this study. One of the papers from 2015/2016 was left out because it was a short introduction to a special issue.

To investigate which topics where included in the papers, a thematic analysis inspired by Brown and Clark method [5] was performed. The paper titles and abstracts, which for almost all papers consist of; objective, background, method, results, applications and keywords were copied to a table in a word document. The papers' themes or main topics were coded/interpreted, based on this information, for each paper. For some papers, further information from the papers had to be read to define their topic. Thereafter, all the codes from all the papers were compared with each other and color coding (one color for each topic/theme) was used to group the papers in broader topics or themes.

The papers were first coded by the first author. Subsequently the second author looked at the information from the papers and the codes and made his own judgements about the codes. There was a general consensus between the authors. Some small differences in judgements of themes/topics were discussed and based on these discussions some minor changes in the topics were performed. Finally all the themes/topics and the papers sorted into them were described by the first author.

The first author coded the methods and levels used in the papers. This information needed less interpretation and was therefore analyzed by one author. Since this paper is a conference paper with a limited length, there is not enough space to refer to all the papers included into the analysis.

3 Results

As can be seen in Table 1, we found twenty main themes or topics within the papers in the journal *Human factors*. The themes were: car driving, physical workload, human-automation interaction, design and usability, human machine interface (displays, controls and alarms), mental workload, cognition, team work, training/simulations, anthropometry, safety, virtual reality, human reliability, procedure, human factor method, cyber security, physical environment, stress, dynamic systems modeling and psychomotor test. Totally 164 papers, or approximately 90 percent of the papers, were about the first ten themes. The remaining ten themes included less than five papers each. Below we provide a short description of each of the topics.

3.1 Car Driving

Car driving was the topic most found and a total of 38 papers were sorted under this topic. There was a special issue, (issue 8, in 2015) about assessing cognitive distraction in driving. However, without that particular issue, car driving would still be the topic with the most papers included. All the papers within the car driving topic were related to safety. Within the category car driving, the topic most studied is distractions, which

Table 1. Topic/themes into which the papers in the journal Human Factor for 2015 and 2016 are sorted, the numbers of papers that are included in each theme for 2015 and 2016, and the total of these two years.

Topics/themes	2015	2016	Total
Car driving	24	14	38
Physical workload	12	14	26
Human-automation interaction	12	8	20
Design and usability	7	9	16
HMI (Displays, controls and alarms)	9	5	14
Mental workload	6	7	13
Cognition	6	6	12
Team work	4	6	10
Training/simulations	2	5	7
Anthropometry	4	2	6
Virtual reality	4	0	4
Safety	1	3	4
Human reliability	2	0	2
Procedure	1	1	2
Human factor method	1	1	2
Cyber security	1	1	2
Physical work environment	1	1	2
Stress	0	1	1
Dynamic systems modeling	0	1	1
Psychomotor test	1	0	1
Total	98	85	183

was covered in seventeen papers. There were two papers on attention and driving, which are related to the topic distraction since this topic is concerned with distraction from internal sources. There are four papers in this topic, which are related to autonomous car driving. Four studies investigate the usability of different support systems within the car, and three study the field of view. There are two papers that investigate range anxiety in electrical vehicles There is also one paper on crossing behavior and decision making, driving styles and road safety, and pedal application and safety. One study tested if drivers take into account the action boundaries of their car when overtaking. One paper on car speeding in simulated hazard scenes and differences between young and novice drivers. There were two papers included within this category that are in the boundary of the topic, since their main concern is pedestrians. However, these studies were included in this category, since this is clearly related to safety and car driving.

3.2 Physical Workload

All the papers sorted into this topic are related to how different forms of physical workload such as standing, prolonged sitting, lifting, and hand force have a negative effect on the human body (such as muscles and back pain).

Twenty-six papers were included in this topic. In Issue 5 for 2016 there was a special section on the impact of Thomas Water on the field of ergonomics. This issue included nine papers that were included in this topic. However, without taking into account the special issue, a high number of papers were also included into this category. Four of the papers in this topic have also investigated physical workload in different occupations.

Some of the studies are related to another topic that we found namely design and usability in that they investigate how different interventions reduce physical workload. If papers seemed to focus more on the physical workload than the design and usability they were included into this category.

3.3 Human-Automation Interaction

All the papers in this topic are related to how humans interact with automation. Most of these papers seem to focus on how humans perceive autonomous systems and how their perceptions affect performance. In these studies, trust in automation seems to be a rather large topic. Some studies look at individual differences in perceiving, experiencing and performing, with automation One study investigated how a specific system affected performance, workload, and situation awareness and one study investigated how different types of automation failure affect performance.

One study looked at human automation interaction used in different fields and one study looked at different human sub-systems that are affected by human-automation interaction. There is one study investigating tactile language for human computer interaction, and one study that investigate human aware motion in human-robot collaboration. There is also one study that investigates the cooperation behavior of an agent.

3.4 Design and Usability

Sixteen papers were grouped into this topic. Most of the studies within this topic are related to usability tests (performance and/or preference) of the design of a product for a particular situation, purpose, a group of people, or a combination of these three, or the testing and development of guidelines that have an aim to increase usability. There was also one study that developed and tested a user experience satisfaction scale. One study is concerned with culture in design. In this topic, there were two studies that evaluated a design with physiological measurements. These two studies are in this topic and not in the physical workload topic, because the main purpose seems to be the test of a design rather than the physiological workload in itself.

3.5 Human Machine Interface (Display, Controls and Alarms)

Fourteen papers were included into this topic. This topic is very much related to the last topic of design and usability. However, the studies that are included in this category are usability studies of the human-machine interface or, more specifically of displays, controls and alarms. Eight of these studies are concerned with displays, three studies on alarms, and one study on control responses. There was also one paper that studied the overall HMI. The studies within this topic could have been included in design and usability, however, since all of these papers studied human machine interface, they were collected under one topic.

3.6 Mental Workload

Thirteen papers were included under this topic. The topics in the papers were related to different types of mental workload such as sustained attention, interruptions, sleep disruptions, watch schedule, transitions in task demand, night shifts, break length, boredom in the workplace, shift length, and trajectory uncertainty. There are some papers that are concerned with measurement of mental workload.

3.7 Cognition

Twelve papers were included under this topic. The main purpose of the papers that were sorted into this topic is to study the different cognitive processes and their effects on task performance. The main topics in these papers were: multitask resource allocation, uncertain contact location in simulated submarine track management, factors that influence the predictions of uncertain spatial trajectories, task switching when concurrence is impossible, goals and strategic mental models, effects of standing or walking on mental functions, sustained attention and loss of inhibitory control, reducing disruptive effects of interruption, operational decision making, situation awareness in offshore drillers, individual differences in verbal-spatial conflicts, and situation awareness in submarine track management.

3.8 Team Work

Eleven papers were included in this topic. In three of the papers, the main topic was team training and measuring effects of team training. These papers could also be in the training/simulation topic. The other papers in this category studied the effect of coacting observers, coordination strategies, haptic communication, increase of task relevant information, strategies for pre-handover preparation, and the effectiveness of brainstorming.

3.9 Training/Simulations

Seven papers were included in this topic. In two of these papers simulator or simulation training were the main topic. Two of the papers evaluated the realism and transferability of training from the simulator. For two of the studies, the topics studied were related to training in general (and not in a simulator). All the studies that were included in this topic were related to the context of training and technology.

3.10 Anthropometry

Six papers were included in this topic. All the papers that were included in this topic studied the measurement of the size and proportions of the human body that is useful for design.

3.11 Virtual Reality

Four papers were included in this topic. Each paper studied one of these topics: hand gesture, exertion of force, haptic perception, and localization of spatial differentiated virtual audio signals in virtual reality.

3.12 Safety

Four papers were included within this topic. The papers sorted under this topic are very different; however, all of them investigate some form of safety as the main topic. One of the papers investigated intervention to reduce slips, trips and falls in hospitals. In one study a maintenance questionnaire is developed. One investigated a system to analyze trading incident. Finally, one study explored interdependencies of human and organizational subsystems of multiple complex, safety-sensitive technological systems.

3.13 Human Reliability

Two papers were included under this topic. One study investigated visual inspection reliability, and one study investigated a new human reliability analysis method.

3.14 Procedures

Two papers were included in this topic. In one study, the effects of hand-off protocols were investigated, and in another study, an intervention to study procedural errors was investigated.

3.15 Human Factor Method

Two papers were included in this topic. One of these papers studied if thinking aloud influenced perceived time, and the other studied the use of link analysis.

3.16 Cyber Security

Two papers were included under this topic. One of the papers studied vulnerability to phishing attacks and the other paper studied the role of human factors in security in cyberspace.

3.17 Physical Environment

Two paper where included in this topic. One of the papers investigated motion sickness and the second paper investigated perceived spaciousness.

3.18 Stress

One paper was included in this topic, which investigate multidimensional assessment of task stress.

3.19 Dynamic System Modeling

One paper was included in this topic, which describes the modeling of factors influencing long-term viability of a food supply network.

3.20 Psychomotor Test

One paper was included in this topic, which estimated finger-tapping rates and load capacities.

4 Method Used and Level Studied in the Papers

Table 2 shows in how many papers used the different methods. The experiment was the most used method. Literature review, questionnaire, other qualitative methods (than experiment and questionnaire) and discussion, were used in some studies. Both qualitative and quantitative, quantitative meta-analysis and qualitative method was used in a few studies.

The levels (individuals, workgroup, organization or society) that the papers describe are shown in Table 3. The table shows that an individual level was investigated in almost all (160) papers, a work group level was investigated together in

Table 2. Methods used in the papers in Human Factors journal for 2015 and 2016

Methods	Number
Experiment	125
Literature review	16
Questionnaire	12
Other quantitative methods[a]	11
Discussion paper	9
Both qualitative and quantitative	5
Quantitative meta-analysis	3
Qualitative	2

[a] Than experiment and questionnaire.

Table 3. Levels investigated in the Human Factor journals paper for 2015 and 2016

Levels	Number
Individual	160
Work group	10
Organization	1
Society	2
Individual/team/organization	4
Not possible to interpret level	6

fourteen papers and an organizational level in together five papers and a society level in two studies. In six papers, it was not possible to interpret a level.

5 Discussion

The results show that the main topics that are included by more than five papers in the journal ‘Human factors’ are; car driving, physical workload, human-automation interaction, design and usability, human machine interface (displays, controls and alarms), mental workload, cognition, team work, training/simulations, anthropometry. Some smaller topics, where less than five papers were included are: virtual reality, safety, human reliability, procedure, human factor method, cyber security, physical work environment, stress, dynamic systems modeling and psychomotor test.

The topics that seem to be unique for human factors are all the topics that are about (a) human-computer/technology interactions (human-automation interaction, human machine interface (displays, controls and alarms), virtual reality, cyber security), and (b) the topic design and usability. Both anthropometry and physical and mental workload are related to and important for design and usability. Safety or human reliability could also be important for human computer/technology interactions, HMI, and design and usability.

However, from the list and the description of the papers within the topics human factors seems to consist of topics that are unrelated and which are connected to different academic fields. For example, physical workload and anthropometry seems to be connected to physiology while topics like mental workload, team work, cognition and training/simulations are connected to psychology. Car driving or traffic, at least in Norway is an academic field in itself.

Even if human errors are often used as a dependent variable in the human factors experiments, safety and human reliability are small topics within the Human Factors journal, with altogether six papers that are sorted into these topics.

In the definition of human factors and in the general literature one might get the impression that human factors is a field that tries to include very much from other fields. It is a question if that is the best way for the academic field of human factors to proceed. If human factors overlap with several academic fields (e.g. cognitive psychology, work and organizational psychology, organizational science, safety science, traffic, physiology, anthropometry, and occupational therapy), it might be difficult to describe what a human factors expert is. One person would usually not have training in more than a few of these academic fields, and if he/she does have training in one or two of the fields, is that then sufficient to be a human factors expert? It is also difficult to know when you look for a human factors expert, what types of knowledge the person possesses.

We here argue that human factors should limit itself and not include several large scientific fields. The papers in the Human Factors journal do not reflect that human factors is covering the entire human field. Additionally, it is a bit ironic that in a field, where one of the main topics is usability, has been so vague on what human factors include or excludes.

The method used in most of the studies in the journal is experiments. Few paper used other methods.

Almost all of the papers described an individual level. Hence human factors is studying individuals, some studies on the work group level, and studies at the organizational and society level do almost not exist. From this, it could be argued, that human factors are more related to work psychology than to organizational psychology or to organizational science.

This paper has investigated the topics, methods and levels studied in the papers in the Human Factors journal. The papers that are included into a journal might again depend on several factors. The first one is the journal selection process by the authors. Usually authors would look at the former papers in the journal to see if their papers fit there. Furthermore, choosing a journal might depend on other journal options where the paper might fit better as well as the ranking of the journals. Thereafter, the editor(s) and the reviewers also make a decision whether a paper fit within the journal or not. It might be that not all research that is contemporary representing human factors research would be included in this journal. However, we think that the collective process between authors, reviewers and editors should give a broad representation of the research within human factors.

6 Conclusion

Our conclusion from investigating the papers in the Human Factor journal is that human factors should be limited to the study of humans interaction with computers/technology and usability of different types of design. It could be a good idea to split human factors and ergonomics where the ergonomic deals with topics like physical workload and anthropometry. The study of cognitions in itself, should continue to be included in cognitive psychology. However, cognition is also relevant for human interaction with computers/technology and for the usability of different types of design and in this context, it is relevant for human factors.

Teams, mental workload and stress have been studied in work and organizational psychology and can continue to belong there. It seems like organizational safety is included within safety science and organizational science, which is very different from human factors. The main method used in human factors is the experiment, and the data are analyzed at an individual level, which shows that human factors is not an organizational science. Reliability and human reliability also seem to be an academicals field in itself, however human reliability, which is often analyzed at an individual or work group level, seems closer to human factors than to organizational safety.

References

1. IEA: IEA Triennial Report, 2000–2003. IEA Press, Santa Monica (2003)
2. Dul, J., Bruder, R., Buckle, P., Carayon, P., Falzon, P., Marras, W.S., Wilson, J.R., van der Doelen, B.: A strategy for human factors/ergonomics: developing the discipline and profession. Ergonomics **55**, 377–395 (2012)
3. Hollnagel, E.: Human factors/ergonomics as a systems discipline? "The human use of human beings" revisited. Appl. Ergon. **45**, 40–44 (2014)
4. Sætren, G.B., Hogenboom, S., Laumann, K.: A study of a technological development process: human factors—the forgotten factors? Cogn. Technol. Work **18**, 595–611 (2016)
5. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**, 77–101 (2006)

Use of Risk Information to Support NRC Human Factors License Amendment Reviews, Inspections, and Research

Brian Green(✉) and Jing Xing

U.S. Nuclear Regulatory Commission, 11555 Rockville Pike, Rockville,
MD 20852, USA
{Brian.Green, Jing.Xing}@NRC.gov

Abstract. The mission of the U.S. Nuclear Regulatory Commission (NRC) is to protect the health and safety of the public and the environment with regards to the commercial use of radiological material. This paper describes various ways in which the NRC uses risk information to guide various human factors processes (such as nuclear power plant control room design and modification technical reviews, control room construction inspection efforts, and human factors research). Included are descriptions of observations and lessons learned utilizing risk insights and challenges associated with incorporating risk into NRC human factors processes and recommendations for improving this process within the organization.

Keywords: Human factors · Risk assessment · Human reliability analysis · Risk-informed regulation

1 Introduction

The U.S. Nuclear Regulatory Commission (NRC) licenses and regulates the nation's civilian use of radioactive materials to protect public health and safety, promote the common defense and security, and protect the environment. Since the Three Mile Island accident of 1979, the NRC has considered the important role that organizational and human factors design principles can play on the safe operation of nuclear facilities.

Human factors at the NRC has a very specific scope: to promote safe operation of nuclear plants. Other common human factors goals such as usability evaluations and efficiency are typically not considered. This safety focus is not unique by any means, but it is different from the norm. Despite this intentional limiting of scope, the number of safety considerations can still be quite large.

For instance, an important part of the human factors work at the NRC reviews the design of and modifications to the main control room (MCR) of nuclear power plants. This is the part of the plant where licensed operators monitor the status and control the operations of the plant. The MCR contains many human-system interfaces (HSI) (i.e. controls, displays, and alarms). NRC human factors staff review proposed changes to the MCR design, procedures, training, etc. to ensure that the changes ultimately support the safe operation of the plant.

© Springer International Publishing AG (outside the USA) 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_8

In addition to the MCR, NRC human factors staff also consider other systems such as the design of local control stations (HSIs located outside of the MCR), remote shutdown workstations (to be used to shutdown the plant in the case the MCR is uninhabitable or no longer functioning properly), and emergency response facilities located either on or off-site.

Human factors reviewers also consider changes to important human actions (IHAs) which are credited in the plant safety analyses or determined to be of special importance during a probabilistic risk assessment (PRA).

Given the broad scope of activities described above, it is challenging for NRC human factors staff to apply these considerations to the roughly 100 operating nuclear units, 4 units under construction, and several plants in the design certification process in the U.S. today. Therefore, it is necessary to use a rational and systematic method to ensure that the human factors staff identify potential risks and adequately addresses the areas that can influence safety.

In 1995, the NRC published the Probabilistic Risk Assessment (PRA) policy statement [1], which created a framework for NRC staff to utilize risk information to conduct a risk-informed regulatory process. Since then, NRC staff have used risk information in a variety of ways to ensure that staff resources focus on structures, systems, components that are most likely to affect safety. In recent years, there has been considerable incentive to increase the use of risk-informed methods to systematically address safety, gain organizational efficiency, and reduce unnecessary regulatory burden on applicants and licensees. The PRA policy statement provides a rational means to achieve all of these goals.

Human factors is just one of many disciplines that has benefited from risk insights gleaned from PRA and human reliability analysis (HRA) (an essential part of PRA). NRC human factors staff are striving to develop and improve these methods to perform technical reviews, inspections, and research methods to support these goals and to ensure that they focus their efforts on areas that have the great influence on safety. The remainder of this paper will highlight some of these human factors methods and describe how risk is used to improve the human factors processes.

2 Overview of the Human Factors Review Model and Risk Assessment Framework at the NRC

2.1 Human Factors Review Model

10 CFR 50.34(f)(iii) requires licensees of US commercial nuclear power plants to use a “state-of-the-art” human factors program prior to constructing or modifying the MCR. The NRC uses its “Human Factors Engineering Review Model,” known as NUREG-0711 [2], to conduct technical reviews to determine if 10 CFR 50.34(f)(iii) and other regulations have been met. NUREG-0711 is NRC staff guidance that contains review criteria used to assess the quality of a proposed human factors program. Licensees who develop and maintain a human factors program that is consistent with NUREG-0711 are determined to be compliant with 10 CFR 50.34(f)(iii).

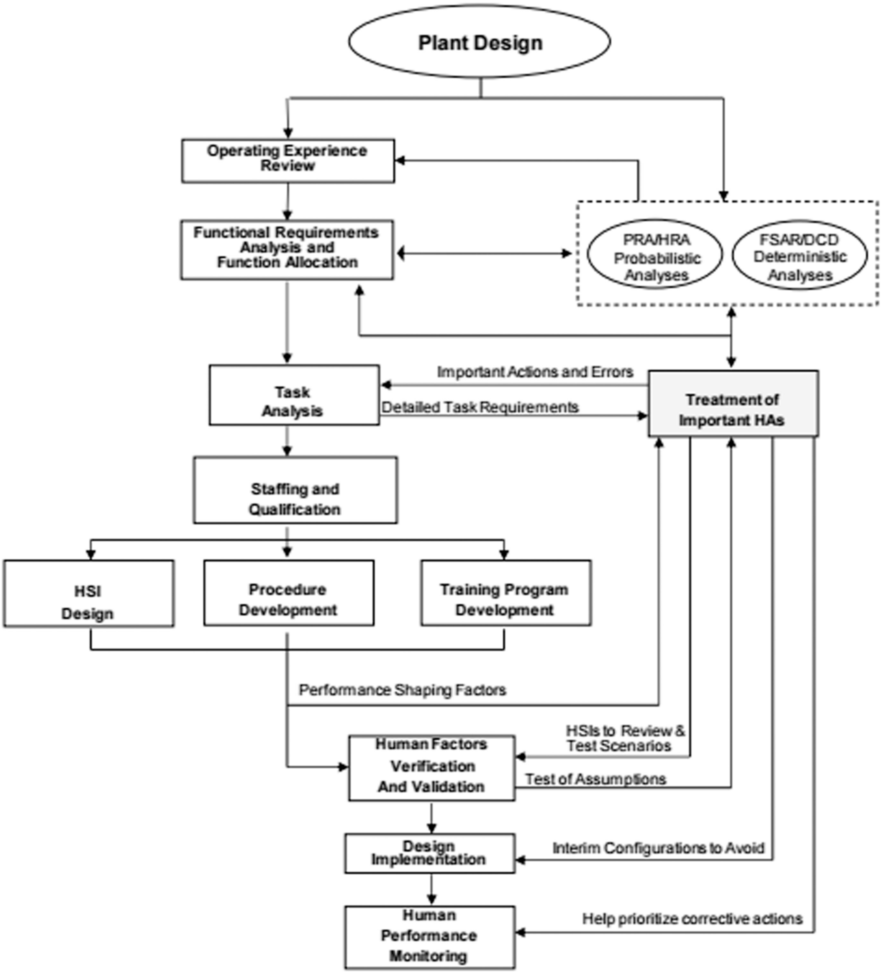


Fig. 1. This figure illustrates eleven of the twelve elements of a NUREG-0711 human factors program in the solid rectangles (the twelfth is Human Factors Engineering Program Management, which is a plan for ensuring these eleven elements are conducted in an integrated manner). (Reproduced from [2]).

Figure 1 illustrates the NUREG-0711 framework. It includes 12 review elements that span the entirety of the design and operation of the plant. For new designs, a review is conducted using criteria from all 12 elements. For plant modifications, depending on the scale of the proposed modification, only a subset of the elements may be applicable.

2.2 Human Reliability Analysis Process and Resulting Risk Information

The NRC uses PRA for its risk-informed regulatory and licensing activities. HRA is an essential part of the PRA; it provides a systematic understanding of human

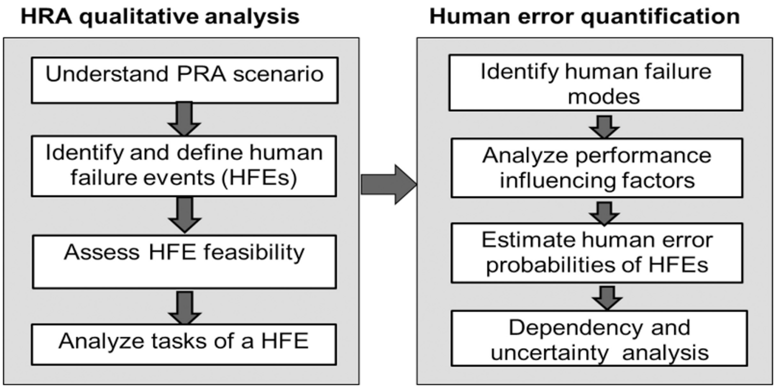


Fig. 2. The framework of human reliability process [3].

performance of risk-important actions, referred to as human failure events (HFEs). Figure 2 shows the generic steps in a typical HRA. While the outcome of one HRA step is the input to the steps following, the outcome by itself provides the understanding of human performance from the perspective of the step. Below is a brief summary of the steps and the outcomes:

- Analyze PRA scenario – A PRA scenario describes a sequence of events happening, including the initiating events, the involved system responses, and required human actions. The output of the scenario analysis includes the operational narrative that describes what may happen (especially under unexpected, non-typical conditions) and context information describing the conditions that may negatively influence human performance.
- Identify human failure events (HFEs) – The output is a set of safety-important human factions, failure of which would affect system safety.
- Assess HFE feasibility – The output is the assessment of whether an HFE can be performed by operators for the given context of the scenario.
- Analyze tasks – The output is the sequences of tasks that operators perform in order to be successful for the human action defined in the HFE. Operators may have multiple alternative sequences to achieve the same human action.
- Identify failure modes – The output is a set of failure modes describing the potential ways that operators may fail to perform the tasks or actions.
- Analyze performance-influencing factors – The output is a set of factors that can lead to the failure modes. Examples of such factors are time available for performance, task complexity, HSI design, training, and procedures.
- Estimate human error probabilities – The output is the likelihood of the HFEs.
- Analyze dependency and uncertainties – The output is the dependency of the HFEs and its effect on human error probabilities. HRA also requires documenting uncertainties in its process and results.

Given the above process, some risk information that can be gained from HRA and inform human factors includes the following:

- Identified imperfect, unexpected, and non-typical conditions that challenge human performance
- Identification of human actions that may lead to unsafe plant status
- Potential ways that crews may fail required actions
- Performance influencing factors that impact crew performance
- Likelihood of personnel successfully performing the actions

Such risk information helps to ensure that the NRC staff focuses on safety-important areas based on the potential consequence and likelihoods of operators failing to perform required safety-important actions or performing unsafe actions because of proposed designs or modifications. The next section provides an in-depth discussion on utilization of risk information to inform human factors processes.

3 Use of Risk Information in NRC Human Factors Processes

Below you will find descriptions of how NRC staff use risk to inform human factors technical reviews of nuclear power plants and the research supporting these reviews.

3.1 Treatment of Important Human Actions in NUREG-0711

Chapter 7 of NUREG-0711 addresses the treatment of important human actions (IHA). This review element describes a method for determining the actions that have the greatest impact on safety. These actions fall into two basic categories.

The first category of actions, are those deterministic operator actions, which if not successfully completed, will have significant negative impact on the plant during certain prescribed postulated scenarios (such as a loss of coolant accident). These can be thought of as the "worst-case" scenarios, therefore it is important to ensure that operators can prevent and mitigate the consequences in those rare instances when they occur.

The second set of IHA are called risk-important actions. Risk-important actions are determined based on the plant probabilistic risk assessment (PRA) and the human reliability analysis (HRA). These analyses often identify a set of actions that are important to plant safety that are different from those identified by the deterministic analysis. These actions may have lower consequences to safety (than the deterministically identified actions described above) but may occur more frequently, or may be related to a particularly error prone operator action. Many human actions modeled in PRA are those that operators need to perform to recover the plant systems to a safe status when there are failures in the plant systems, structures, or components.

Together, these deterministic and probabilistic actions are considered as IHAs and they take special importance in the NUREG-0711 process. The results of these analyses are fed into the iterative human factors design process and used the during verification and validation activities. This paper focuses on the use of IHAs during the human factors validation process, but the reader should be aware that IHAs are considered throughout the human factors process. Figure 1 illustrates how IHAs are used throughout the human factors process and shows how the analyses in the Final Safety

Analysis Report (FSAR) and Design Control Document (DCD)[1] are used as source material.

The human factors integrated system validation process (ISV) is the main opportunity for the designers of the plant to demonstrate that operators can safely operate the plant under normal and emergency conditions using actual procedures, training, and realistic staffing considerations. Validation is conducted in a high-fidelity MCR simulator that closely resembles the design of the final plant in design and functionality.

Chapter 11 of NUREG-0711 describes an acceptable method for conducting ISV. This chapter relies heavily on the results of the treatment of important human action elements (Chap. 7) to focus the review of ISV implementation plans and inspections of the ISV process. For instance, Sect. 11.4.3.2 "Test Objectives," indicates that applicants should develop detailed test objectives to determine if the actual operator performance can be used to validate the assumptions about performance of important human actions. Several criteria like this use risk as a consideration during the review of applicant implementation plans for the ISV process.

When the design is complete, an inspection of the ISV is conducted to ensure that the applicant followed the approved human factors engineering program and that the results support the conclusion that the plant can be safely operated as designed. The number of actions conducted by operators is too numerous to closely review them all, so the highest level of scrutiny is typically focused on those IHAs identified in NUREG-0711 Chap. 7. Although it may not be apparent from the formal inspection reports, the ISV inspection typically focuses primarily on IHAs and the design of the HSIs used to conduct IHAs.

3.2 Scaling Operator Manual Action Reviews Based on Risk

Operator manual action (OMA) reviews are conducted when a licensee proposes a change to an operator action that is credited in licensing documents. This may occur with or without a change in the physical plant, (i.e. a licensee wants to credit an operator with opening/closing a valve to prevent flooding in certain buildings after a revision to the postulated flood data). Staff use NUREG-1764 [4] for the review of changes to operator manual actions. NUREG-1764 is much like NUREG-0711 in that it covers the same human factors processes, but in a greatly abbreviated manner (it has the same 12 review elements, but with fewer review criteria in each). These review elements are scoped into the review based upon the risk associated with the OMA under review. In other words, those OMAs with large potential to affect plant safety get a relatively more thorough review compared to those OMAs with little potential to impact safety.

Risk is assessed using one of four methods described in NUREG-1764 (generic human action method, estimate importance method, use of available risk information, and calculate importance method). It is outside the scope of this paper to describe these methods in detail here. However, some of these methods require the assistance of

---

[1] The DCD and FSAR are two important documents in the NRC licensing process. Information from these documents is used to support the various technical reviews including the human factors review.

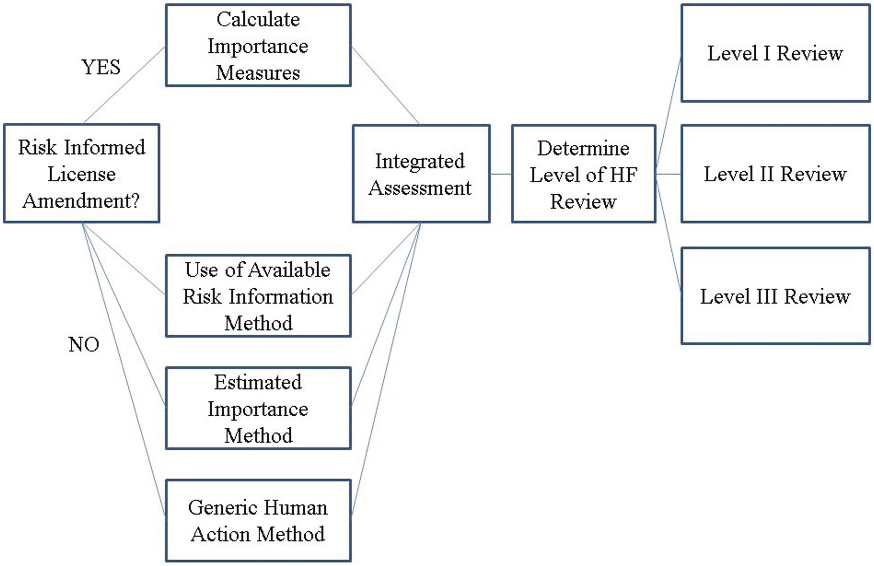


Fig. 3. This figure provides a simplified overview of the NUREG-1764 risk-screening process. Level I reviews receive a high amount of scrutiny. Level II and III reviews receive less scrutiny because the risk associated with the amendment is progressively less.

qualified risk analysts to modify PRA/HRA models to quantify the change to safety for a proposed amendment. Other methods are qualitative and can be conducted by a human factors expert without the assistance of risk analysts (Fig. 3).

Regardless of the risk assessed during scoping, human factors reviewers can decide to add review elements that are deemed necessary (or subtract elements that add little to the review) in the review process. This is an important part of the process, which provides an extra layer of protection against erroneously screening a high risk OMA as a moderate or low risk OMA.

3.3 Informal Use of Risk During HSI Design Reviews

NRC staff use risk to assess HSI design characteristics as well. Control rooms have hundreds, if not thousands, of controls, displays, and alarms. It would be very costly to examine the properties of each, therefore staff use risk information to help narrow the scope.

Applicants and licensees create style guides, usually based on NUREG-0700 [5], the NRC approved HSI design guidelines. HSIs that are designed according to this guidance are deemed appropriate under 10 CFR 50.34(f)(iii). NUREG-0700 has several hundred review criteria. Therefore, the nuclear industry and NRC staff must use discretion when applying this guidance to ensure that the final design ultimately supports safe operation of the plant.

The staff typically prioritize the HSIs that support IHAs and review the criteria that are most likely to affect safe operation. IHAs are not the only actions considered for this review but they are an important set of the sample. The staff also consider factors such as the novelty of the system (i.e. a new type of control will get more scrutiny than more familiar controls), industry operating experience (identifying systems that have been problematic for others) and other considerations.

3.4 Risk Considerations in Human Factors Research

Risk information can also be used to shape the course of human factors research. A recent study conducted jointly by the NRC and the Halden Reactor Project examined the use of computer-based procedure (CBP) systems for evidence of trust-induced operator complacency errors [6]. The CBP system used in this study mimics paper-based industry procedures but with added HSI features. The CBP system reads system parameters from various sensors and provides indications for each procedure step whether the relevant parameter is in or out of specification. The study inserted intentional failures of the CBP into certain specific steps, and then observed to see if operators successfully identified the failures of the CBP. The initial results suggest that operators may suffer from complacency-induced errors similar to users of other automated systems [7].

Additional research is being planned to follow up on these findings. Several different study designs are currently being proposed. Among those being proposed, is a tabletop analysis that uses existing PRA models and HRA methods to assess the consequences of complacency errors for each step of the procedure. This assessment can then be used to gauge the severity of a complacency-induced error for a particular procedure step. Consider the following example:

Procedure step 37 instructs operators to ensure that the vibration of a particular pump is less than X mils. The CBP provides an indication when a parameter is within specification and a different indication when it is not within specification. In this case, the operator sees an indication that the parameter is within specification. If he trusts the CBP, he may be more likely to go on to step 38 without verifying that vibration levels are actually less than X mils using independent indications. If the CBP provides erroneous indication to a complacent operator when vibration levels exceed X mils, then the pump will likely be damaged. The resulting damage to the pump may cause an increase in the risk estimated in the PRA models that can then be quantified.

The process illustrated in the example above can be repeated for each step in a procedure to estimate the corresponding impacts of failing to perform the steps from the PRA perspective. It is expected that operator failure of some procedure steps will have more influence on the overall plant risk than others. These risk estimates can provide researchers with important insights about identifying the most important steps in procedures and can provide insights about whether these types of complacency errors are ultimately important to plant safety.

4 Observations and Discussion

The examples in Sect. 3 demonstrate a variety of ways in which NRC human factors staff use risk-information. This section discusses some of the benefits and complications that arise when using risk information during human factors processes.

4.1 Increased Efficiency: Prioritizing Review Areas and Reducing Requests for Additional Information

The uses of risk information described in Sects. 3.1 and 3.2 help improve the efficiency of the technical review and inspection processes in which they are used. For instance, consider how the NUREG-1764 (Sect. 2.2) reduces the number of review elements for moderate and low risk OMAs. Elimination of unnecessary review elements improves the efficiency of the process and allows the licensee and the NRC to focus their resources on safety-important areas.

Reducing the scope of the review also typically reduces the need for issuing requests for additional information (RAI). RAIs are a formal method for NRC staff to gain more information when a licensee submitted document is incomplete or unclear. RAIs are issued and responded to on a publically available docket; therefore, the amount of management involvement in this process can be substantial, increasing the cost. Issuing and tracking RAIs expends NRC resources. Licensees must also expend resources responding to them; therefore, it is significant beneficial to both parties to minimize the issuance of RAIs.

4.2 Schedule Slippage When Risk Assessment Is Complex

Section 3.2 describes the formalized OMA review process that includes four structured methods of risk assessment, some of which are much more complex compared to others. In all of these methods, the formal risk assessment must occur *prior* to beginning the human factors review. Under some circumstances, this can cause complications in the review process.

NRC and licensee project managers negotiate a schedule for completion of the technical review process when a license amendment is received. Although it is possible to adjust review schedules, the NRC generally strives to stick to the agreed upon schedule unless there are evident safety concerns. Therefore, the amount of time to conduct the technical review can be considered fixed. The amount of time allotted to conduct the technical review includes the time needed for the risk assessment. Therefore, any time taken to conduct the risk assessment is time that is no longer available for the technical review.

One potential positive outcome of the risk assessment is a significant reduction in the time needed for the human factors review. The increased efficiency is noticed most in cases where a quick risk assessment justifies a reduced level of human factors review. In this case, staff can complete the process and move on to other projects.

However, in cases where the risk assessment is complex, or takes a long time to complete, and the results support a high-level of human factors review, then there actually is a loss experienced by the human factors staff. The time that is used to conduct the risk assessment is no longer available to the human factors staff for the review. In this case, the same high-level review must ultimately be done with *less* time available. However, this situation is uncommon.

Successfully managing project schedules, as described above, involves a decision to use the more complex risk-analyses when accuracy of the PRA is advantageous (such as for first of a kind or potentially contentious reviews), or to use the more qualitative approach when accuracy can be sacrificed (such as for routine reviews). An integrated risk assessment and human factors process is a promising way for conducting safety reviews more efficiently and effectively, but it relies heavily upon the ability of staff to make good decisions about risk.

4.3 Concerns for Using Risk Information

In HRA, human error probabilities are often based on expert judgment with limited support from empirical operator performance data. Collecting data to support these estimations is challenging because of the low probability of occurrence (estimates vary significantly but it is not usual to see estimates in the 1 E^{-6} range). As a result, some staff have concerns about relying on these estimates without supporting data. Therefore, the staff typically uses the results of the risk assessment in combination with other approaches. The staff can use the estimates as a starting point, or a guide, and then rely on more traditional methods to build confidence. For instance, if operators are estimated to make error A at a rate of 6.7 E^{-4} and error B at a rate of 3.9 E^{-5} then one could rationally choose to spend more time looking at preventing error A than error B.[2] In this case, there is no official threshold that makes an error acceptable. It simply tells us that error A is an expected order of magnitude worse than error B. There is no implication that error B is acceptable or should not be prevented or mitigated. Another way to address the concern is to understand the uncertainties documented along with the error probability estimates in risk assessment. The uncertainties provide the confidence level of the estimates along with the factors contributing to the uncertainties. The uncertainty information helps the staff make a decision about the scope and focus of the subsequent human factors review. This practice is consistent with the NRC PRA policy statement, which does not remove the need for deterministic processes. Rather, it uses risk to *supplement* existing deterministic methods.

4.4 Development of Human Factors and Risk Assessment Skills at the NRC

In using risk information to enhance human factors reviews, it is essential that the staff understand the appropriate uses and limitations of risk information and feel comfortable

---

[2] The staff may similarly look at core damage frequency (CDF) estimates, a result of the PRA that considers both frequency of occurrence and consequence, or other risk figures to make similar decisions.

using it. Since some staff may be reluctant to embrace these methods, and some review processes do not mandate their use, the staff may be less likely to rely on risk information, and are therefore less likely to acquire the skills necessary to use risk effectively and safely. Therefore, it is imperative that the NRC should continually develop staff skills in both human factors and risk assessment.

The following are some examples of the skill development efforts at the NRC:

- The NRC makes several training courses in PRA/HRA available to staff across the agency. In addition, there is a growing body of staff enrolled in a program dedicated to becoming experts in risk assessment and advocates for its use.
- Two of the three human factors organizations at the NRC are currently embedded in larger organizations dedicated to risk assessment providing staff access to applicable resources.
- In addition, senior human factors staff are working on developing new ways of familiarizing the human factors staff about the benefits and appropriate use of risk.

Other organizations that use risk like the NRC does, or are contemplating using it, should consider developing similar programs that help to educate their staff about the appropriate and inappropriate uses of risk and the potential rewards and costs.

4.5 Risk-Informed Principle for NRC Regulatory Activities

One important principle that the NRC uses is that our work should be risk-informed, rather than risk-based. For instance, achieving a desirable core damage frequency (CDF) in a PRA does not absolve a nuclear plant from doing work to prevent an accident, rather it may mean that the NRC will perform less intense reviews or fewer inspections of that area to ensure compliance with regulations. This way, NRC inspectors and reviewers can spend more time assessing systems and processes that are more likely to affect safety. While this concept is deeply engrained into the NRC staff, it is not uncommon to hear risk-based suggestions at meetings. Risk assessors, and others who are well versed in the principles outlined in the PRA policy statement are typically quick to coach the staff when this occurs, but it is a symptom indicating that additional education may be necessary.

It is important that staff become flexible in their ability and willingness to use risk information during informal processes. In some circumstances, it may be inappropriate to use risk, and staff should lean on other selection criteria. While in others circumstances, it may be possible to increase the use of risk to gain additional efficiency. Of course, this cannot occur unless the staff are competent and confident.

5 Conclusions

The NRC has been using risk information to increase the efficiency and effectiveness of NRC human factors technical reviews and inspections for many years. Risk information helps staff to prioritize their time, focusing their effort on areas that are most likely to have a safety effect. The use of risk information within the human factors processes has generally been positive. In some rare circumstances, using risk information can add time to reviews, but this cost is usually minimal and is far outweighed by efficiencies gained during the majority of reviews. It is important that human factors staff understand the risk assessment process and limitations of the results when applying risk information in their work. It is desirable for the organization to further integrate risk assessment and human factors review processes.

Acknowledgments. The positions described in this paper are those of the authors and do not necessarily represent the views of the U.S. Nuclear Regulatory Commission or the U.S. Government. Thank you to Aida Rivera-Varona, Sean Peters, Sunil Weerakkody and Nathan Siu for your input, assistance, and support.

References

1. U.S. Nuclear Regulatory Commission: Use of Probabilistic Risk Assessment Methods in Nuclear Regulatory Activities; Final Policy Statement. 60 FR 42622. Published in the Federal Register, 16 August 1995
2. O'Hara, J.M., Higgins, J.C., Fleger, S.A., Pieringer, P.A.: Human Factors Engineering Program Review Model (NUREG-0711 Revision 3). U.S. Nuclear Regulatory Commission (2012)
3. Xing, J., Parry, G., Presley, M., Forester, J., Hendrickson, S., Dang, V.: An Integrated Human Event Analysis System (IDHEAS) for Nuclear Power Plant Internal Events At-Power Application (NUREG-2199), vol. 1. U.S. Nuclear Regulatory Commission (2017)
4. Higgins, J.C., O'Hara, J.M., Lewis, P.M., Persensky, J.J., Bongarra, J.P., Cooper, S.E., Parry, G.W.: Guidance for the Review of Changes to Human Actions (NUREG-1764 Revision 1). U.S. Nuclear Regulatory Commission (2007)
5. O'Hara, J.M., Brown, W.S., Lewis, P.M., Persensky, J.J.: Human-System Interface Design Review Guidelines (NUREG-0700 Revision 2). U.S. Nuclear Regulatory Commission (2002)
6. Taylor, C., Hildebrandt, M., McDonald, R., Hughes, N.: Operator Response to Failures of a Computerised Procedure System: Results from a Training Simulator Study (Halden Work Report No. 1198). Institute for Energy Technology, Halden, Norway (2017)
7. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**(1), 50–80 (2004)

Task Analysis as a Cornerstone Technique for Human Reliability Analysis

Claire Taylor(✉)

OECD Halden Reactor Project, Institute for Energy Technology,
P.O. Box 173, Halden 1751, Norway
`claire.taylor@ife.no`

Abstract. Qualitative analysis is essential for obtaining credible and useful HRA results, yet most HRA method descriptions fail to adequately describe how to perform qualitative analysis. Research by the Halden Reactor Project identified task analysis as one of the key qualitative techniques for HRA, and also one of the most challenging for less experienced analysts. This paper makes an argument for why task analysis should be considered a cornerstone technique for qualitative HRA, and also describes current Halden research activities to investigate the role of task analysis in HRA, and to develop support tools for analysts to address these challenges.

Keywords: Human reliability analysis · Human factors · Task analysis

1 Introduction and Background

The importance of qualitative analysis for human reliability analysis (HRA) is becoming more apparent, as evidenced by recent HRA research [1–6], HRA guidance [7] and recently developed HRA methods such as IDHEAS [8] and Petro-HRA [9]. It is acknowledged that there is a need for more detailed qualitative understanding of human operator tasks and constraints in order to lend greater credibility to the quantitative results of the HRA.

There is also a growing acknowledgment of the role of HRA in developing improvement strategies to reduce the risk from human error, i.e. beyond the scope of probabilistic safety/risk analysis (PSA/PRA) [10]. For example, the Petro-HRA method guideline includes a section on how to utilize qualitative HRA findings as input to human error reduction efforts. Such utilization of HRA findings beyond quantitative analysis emphasizes need for a more thorough qualitative analysis to better understand the human performance drivers that sit behind the human failure events (HFE) and human error probabilities (HEP).

Recent research carried out by the Halden Reactor Project on how to improve HRA practices identified an issue with many current HRA methods, which is that the method description often fails to adequately describe how to perform qualitative analysis [11]. This can result in less experienced analysts lacking confidence in their approach and findings, increased variability between analysts, incomplete substantiation of HRA findings, and a lack of transparency and traceability of results.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_9

1.1 The Problem with HRA

The International HRA Empirical Study [1] and the US HRA Empirical Study [2] both identified that a detailed qualitative analysis is required in order for the analyst to develop the necessary understanding of the context of the analysis scenario and the drivers affecting human performance in that scenario. The necessary level of analysis is often beyond what is described in the documentation accompanying most HRA methods. In fact, the International HRA Empirical Study noted that study participants with higher levels of HRA experience tended to perform a much more detailed qualitative analysis than what was strictly required by the HRA quantification method that they used.

This highlights a problem with HRA method guidance, in that it does not adequately describe what is generally considered as necessary for a comprehensive qualitative analysis to inform and substantiate the HRA results. This is particularly of concern for less experienced analysts who are more reliant on following the instruction provided by the available HRA method guidance, and who may not be aware of the necessity to perform additional qualitative analyses beyond this scope.

A study was established at the Halden Reactor Project to further investigate this mismatch between HRA method guidance and applied HRA practices, with the goal of developing practical guidance on how to perform qualitative analysis for HRA. Interviews were conducted with analysts ranging in experience, and a short questionnaire was distributed to collect data on how analysts *actually* perform HRA and the difficulties they experience. From this, we established that many experienced analysts follow a similar approach to HRA in terms of the techniques they apply and the order in which they apply them, regardless of the quantification method that they use. This approach is broadly similar to that presented in Fig. 1 (adapted from [12]).

In particular, many analysts confirmed that they would *always* perform some kind of qualitative task analysis as part of the HRA, and usually a combination of task analysis techniques. Yet task analysis also appears to present some of the biggest challenges for some analysts. When describing what make HRA difficult the analysts (and, in particular, the less experienced analysts) mostly described difficulties such as:

- How far should I decompose the tasks?
- How do I identify critical tasks or actions?
- How do I identify human errors in the tasks?
- How do I model human errors?
- How can I link the human error model and quantification back to the task analysis (i.e. to substantiate the findings of the HRA)?

As noted earlier, most HRA method descriptions do not provide guidance on how to address these challenges, instead focusing solely on how to perform quantification calculations to generate HEPs. Indeed, many HRA methods do not describe any form of qualitative analysis, and simply assume that the analyst will have performed some without specifying what that should be. The good news is that this trend appears to be changing with recent publications as the US Nuclear Regulatory Commission's (US NRC) Fire PRA guidance [7], the IDHEAS method [8] and the Petro-HRA method [9] incorporating more detailed description of qualitative analysis requirements for HRA.

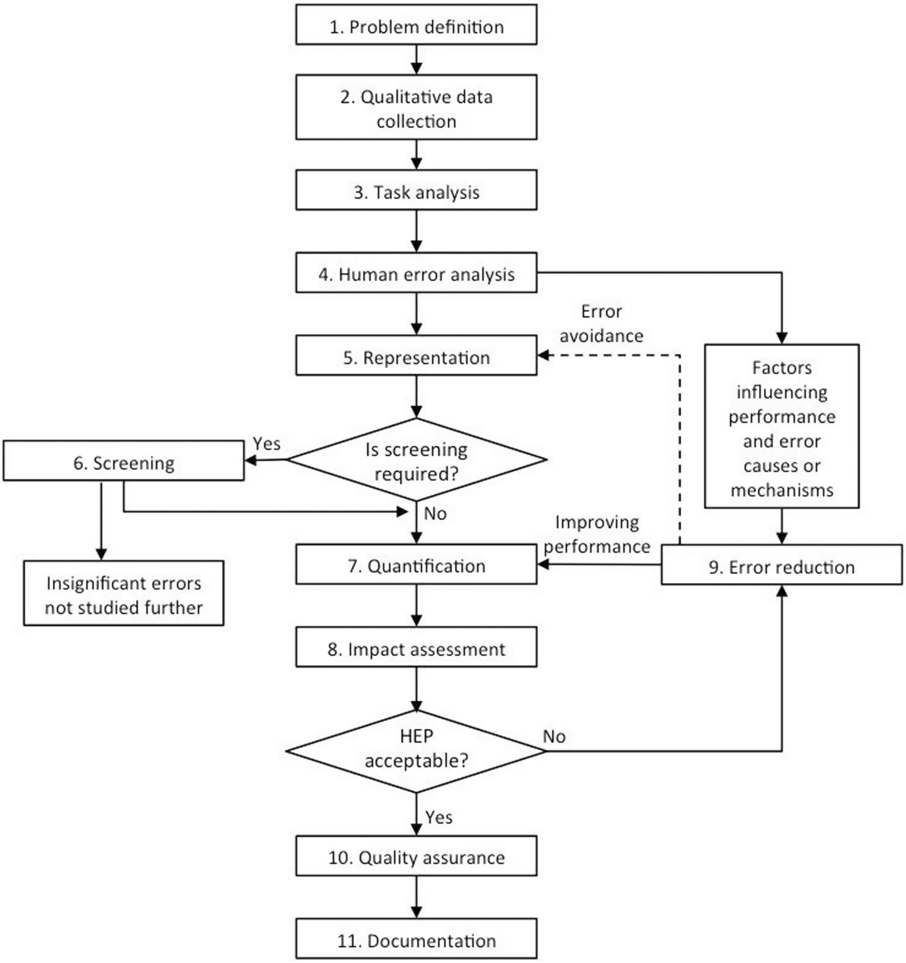


Fig. 1. A generic HRA process (adapted from [12]).

2 A Quick Overview of Task Analysis

"Task analysis covers a range of techniques used by ergonomists, designers, operators and assessors to describe, and in some cases evaluate, the human-machine and human-human interactions in systems. Task analysis can be defined as the study of what an operator (or team of operators) is required to do, in terms of actions and/or cognitive processes, to achieve a system goal. Task analysis methods can also document the information and control facilities used to carry out the task" [13, p. 1].

Table 1 lists the ten task analysis techniques recommended by Kirwan [12] for use during HRA.

It is not expected that analysts would use all ten techniques during the course of a single HRA. Rather, the analyst would typically select the most appropriate technique(s),

Table 1. Task analysis techniques for HRA

Technique name	Technique description
1. Hierarchical Task Analysis (HTA)	One of the most commonly used task analysis techniques; the HTA is usually depicted graphically and it describes the task in terms of its top-level goals and the individual operations or actions below these
2. Observation	This method involves watching the operator(s) perform tasks in the field or in a simulated environment
3. Interviews & documentation review	One of the primary sources of information for task analysis. Documents are typically reviewed to collect information about system operation, risk analyses and proceduralized tasks. Interviews with experienced operational personnel are carried out to clarify information, confirm assumptions and answer questions
4. Link analysis	A method to record and graphically represent the nature, frequency and importance of links in a system, e.g. how often an operator walks over to, consults or operates a particular control or display. Generally used for analysis of workspace layout
5. Verbal protocols	An oral commentary of what is being done by the operator while he/she carried out the task. Useful for collecting information about cognitive activities that cannot be observed
6. Decision Action Diagrams (DADs)	A method to describe decision-making tasks where the operator(s) must chose an action path. Useful for describing decisions which would otherwise be difficult to display in HTA format
7. Critical Incident Technique (CIT)	A simple data collection method to identify, via interviews with experience personnel, incidents or near misses which have occurred.
8. Tabular Task Analysis (TTA)	A task description method that records information in a columnar tabular format. The column titles will vary depending on the purpose and main focus of the TA. It is usually used in conjunction with the HTA to further investigate and provide more detail about the tasks described in the HTA
9. Walk-Through/Talk-Through (WT/TT)	A method that uses experts to commentate their way through a scenario either in the field, pointing to the controls and displays that would be used during the scenario (walk-through) or verbally in a different location usually referring photographs, diagrams and procedures to describe tasks and actions (talk-through)
10. Timeline analysis	A task description method to determine the overall time taken to perform the tasks in the scenario, the sequence and intervals at which particular task steps or actions occur, and the duration of particular task steps or actions

based on the needs of the analysis. Some techniques may be used every time, regardless of the context of the scenario being analyzed – for example, interviews and documentation review. Other techniques may be selected because they provide specific information needed for that analysis – for example, if there is some concern that operators might not complete the necessary tasks in time, then a timeline analysis should certainly be performed. Alternatively, if the focus of the analysis is on the usability of a particular display screen, then a link analysis might be performed to document how often and for which task steps the operators use that screen.

2.1 Task Analysis as a Cornerstone Technique for HRA

When one reviews the list in Table 1 of recommended task analysis techniques, it is evident that a combination of techniques can provide a comprehensive insight into how operators perform the task(s) of interest, how they interact with other people and systems around them, and the factors that could influence the likelihood of them making an error during the performance of these task(s).

For example, the following techniques are often used in combination in an HRA to describe, investigate and evaluate operator tasks:

1. First, a HTA maybe developed based on some initial information from the PSA/PRA or subject matter expert to identify the main task goals and sub-goals of the scenario of interest.
2. Then a task walk-through and/or talk-through may be performed with operational personnel, in conjunction with interviews, to collect more information about the task.
3. Next, a tabular task analysis may be developed to collate the information collected to date, to identify knowledge gaps and to identify areas of focus. The tabular task analysis may be updated several times throughout the HRA as new information and clarification is received.
4. A timeline analysis might then be performed, if the timing of the scenario is identified as important to a successful outcome.
5. Throughout this whole process, documents will be reviewed and interviews will be held with various subject matter experts to inform and focus the analysis.

The information collected from the combination of task analysis techniques forms the basis of both the qualitative analysis and the quantitative HEP calculation. With this information, the analyst can develop a clear understanding of the basic task(s) or action(s) that make up the HFE, and can make an informed judgment about the performance shaping factors (PSFs) that can influence the success or failure of that HFE. In addition, the task analysis information can be used to substantiate the HEP calculation, and can provide the necessary transparency for future reviewers of the analysis who may wish to trace the origin of the inputs to the HEP calculation.

The tabular task analysis (TTA) format in particular can be an invaluable tool for screening tasks and human errors, and for documenting the link between the different qualitative analysis techniques and between the qualitative and quantitative analyses. The TTA has the capability to contain all of the information needed to model and quantify human errors, and to document the links between the evidence to substantiate quantification calculations.

Furthermore, task analysis can be used as a powerful tool for risk-informed decision making to input to continuous improvement efforts at the plant by identifying areas of weakness in the plant systems, and by defining the factors that can positively or negatively influence the performance of those systems. The qualitative results from the task analysis can be used in conjunction with the calculated HEPs to develop a risk-prioritized list of improvement recommendations to reduce the likelihood of human error and increase safety and reliability at the plant [10].

As Kirwan states: “Task analysis is therefore a methodology which is supported by a number of specific techniques to help the analyst collect information, organize it and then use it to make various judgments or design decisions. The application of task analysis methods provides the user with a ‘blueprint’ of human involvements in a system, building a detailed picture of that system from the human perspective” [13, p. 1].

3 The Task Analysis Library Concept

3.1 What Makes Task Analysis Difficult?

Despite the fact that there are many qualitative task analysis techniques to choose from, and that these are well documented in literature, our research found that less experienced analysts still struggle with the basics of how to apply these techniques in practice. In particular, analysts struggle with issues as simple as how to get started or what questions to ask to collect information for analysis, through to more complex issues about how to demonstrate the link between the different analysis techniques to build up an argument for the HEP quantification.

However, our research also indicated that it is not so straightforward to develop practical guidance on how to apply task analysis techniques, as often there is no “rule of thumb” for these techniques. Many of the “rules” or guidelines about what techniques to use, or the appropriate depth of analysis, etc. are heavily context-dependent and will vary for almost every HRA. Even experienced analysts will not follow the exact same format every single time, but will in fact tailor their approach to the specific needs of the analysis at hand.

Based on our findings, we considered whether it would be better to teach by example, rather than attempt to develop potentially confusing or overly complex guidance to describe when and how to apply the different techniques. An ideal situation would be for every beginner analyst to work with an experienced mentor to learn the subtleties of task analysis, but this is neither always feasible nor possible. Instead, we explored the idea of developing a task analysis library as an educational tool to capture the knowledge and experience of the analyst and to present examples of completed task analyses that less experienced analysts could refer to and learn from.

3.2 The Purpose of the Task Analysis Library

The concept of the task analysis library is a collection of qualitative task analyses of different accident scenarios showing varying levels of decomposition of tasks and the application of different task analysis techniques. The library entries are annotated by the original analyst to document how and why different techniques were applied, the analysts' own experience of apply those techniques and their thought processes along the way. The library should resemble an analyst's "diary" showing what information they collected along the way, how they organized and analyzed the information using the different task analysis techniques, and how they built an overall picture of the human tasks in the accident scenario.

In this way, it is hoped that the library will be useful as an educational tool for less experienced analysts, by providing examples to help analysts to gain a better understanding of:

- How to collate, organize, interpret and cross-reference information from different sources;
- How to decide on an appropriate level of decomposition of tasks;
- How to develop an adequate understanding of the analysis scenario, context and components, and hot to build a picture of the human operator role in that scenario; and
- What is expected when conducting a qualitative analysis for an HRA, in terms of what task analysis techniques to use and when, and what those look like in practice.

Some inexperienced analysts reported that many times they just don't know how to get started with the HRA. One of the goals of the task analysis library is to provide completed examples to help prompt analysts to get started. The idea is not for analysts to simply copy and paste from the library (although re-use of library content is discussed below), but rather to stimulate thinking and help analysts to figure out what is required for their specific case.

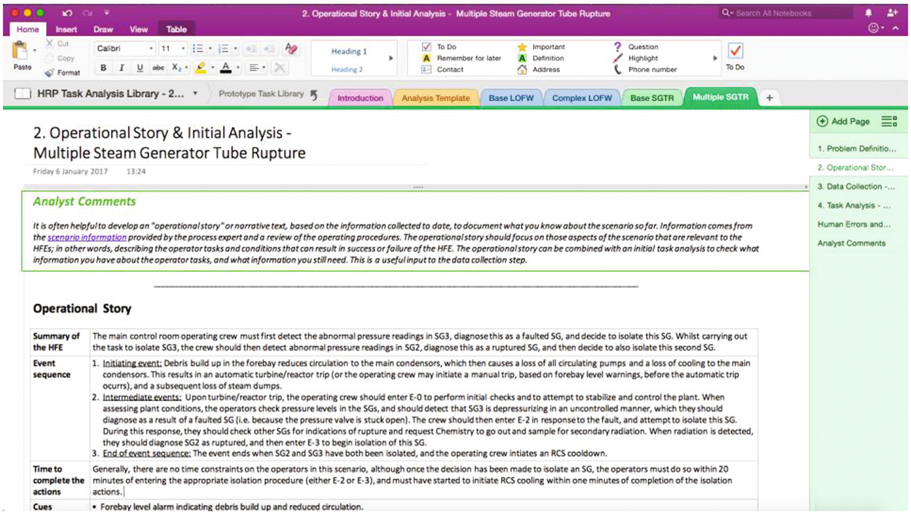
In addition to being an educational tool, a secondary goal of the task analysis library is to provide a template for analysts to expand the library by adding their own completed analyses. In this way, the analyst can gradually build up a repository of qualitative information for re-use in subsequent analyses at the same or similar plants in the future. For example, if performing a periodic review of an HRA, or if performing an analysis of a deviation from a previous HRA, rather than starting the HRA again from scratch, the analyst could use the library as a starting point to review how the previous analyst understood the scenario would unfold, the plant systems, indications and controls involved, the likely operator responses etc., and then adapt the analysis to their specific HRA as required. The expanded task analysis library may also be of interest to analysts who are unable to get access to a training simulator to perform scenario observations. Furthermore, the library becomes an even more comprehensive educational tool for new analysts at that organization in the future.

The task analysis library concept is not intended to take the place of an appropriate training or mentoring strategy; these are still considered optimal approaches for developing HRA competence. Rather, the task analysis library is seen as an additional support tool for analysts to help bridge the knowledge and experience gap in a quicker way.

3.3 The Structure of the Task Analysis Library

At the time of writing this paper, development of an initial task analysis library is underway at the Halden Reactor Project. Figure 2 shows a screen shot of the library at its current stage of development. The initial library is planned to contain four examples of completed qualitative task analyses for two major accident scenario types:

1. A Loss of Feedwater (LOFW) event; and
2. A Steam Generator Tube Rupture (SGTR) event.



2. Operational Story & Initial Analysis - Multiple Steam Generator Tube Rupture

Friday 6 January 2017 13:24

Analyst Comments

It is often helpful to develop an "operational story" or narrative text, based on the information collected to date, to document what you know about the scenario so far. Information comes from the scenario information provided by the process expert and a review of the operating procedures. The operational story should focus on those aspects of the scenario that are relevant to the HFEs; in other words, describing the operator tasks and conditions that can result in success or failure of the HFE. The operational story can be combined with an initial task analysis to check what information you have about the operator tasks, and what information you still need. This is a useful input to the data collection step.

Operational Story

Summary of the HFE	The main control room operating crew must first detect the abnormal pressure readings in SG3, diagnose this as a faulted SG, and decide to isolate this SG. Whilst carrying out the task to isolate SG3, the crew should then detect abnormal pressure readings in SG2, diagnose this as a ruptured SG, and then decide to also isolate this second SG.
Event sequence	1. Initiating event: Debris build up in the forebay reduces circulation to the main condensors, which then causes a loss of all circulating pumps and a loss of cooling to the main condensors. This results in an automatic turbine/reactor trip (or the operating crew may initiate a manual trip, based on forebay level warnings, before the automatic trip ocurrs), and a subsequent loss of steam dumps. 2. Intermediate events: Upon turbine/reactor trip, the operating crew should enter E-0 to perform initial checks and to attempt to stabilize and control the plant. When assessing plant conditions, the operators check pressure levels in the SGs, and should detect that SG3 is depressurizing in an uncontrolled manner, which they should diagnose as a result of a faulted SG (i.e. because the pressure valve is stuck open). The crew should then enter E-2 in response to the fault, and attempt to isolate this SG. During this response, they should check other SGs for indications of rupture and request Chemistry to go out and sample for secondary radiation. When radiation is detected, they should diagnose SG2 as ruptured, and then enter E-3 to begin isolation of this SG. 3. End of event sequence: The event ends when SG2 and SG3 have both been isolated, and the operating crew intiates an RCS cooldown.
Time to complete the actions	Generally, there are no time constraints on the operators in this scenario, although once the decision has been made to isolate an SG, the operators must do so within 20 minutes of entering the appropriate isolation procedure (either E-2 or E-3), and must have started to initiate RCS cooling within one minutes of completion of the isolation actions.
Cues	• Forebay level alarm indicating debris build up and reduced circulation.

Fig. 2. The task analysis library concept (under development)

The library will include a baseline and a complex version of each event, making four entries in total. Each analysis entry includes:

- A problem definition and scenario description, which includes a summary of the technical details of the scenario and the expected operator response, as well as the HFEs that were analyzed;
- An operational story and initial analysis describing the sequence of events, cues and feedback available to the operators, timing of the event, and the main operator tasks that have been identified;
- A summary of the data collection that was performed, including identification of critical steps and other important information about the scenario;
- A detailed task analysis, showing how the collected data is collated and organized, including the identification and screening of human errors and performance shaping factors; and
- Analyst comments throughout the entry, describing what was done and how, the analysts' thinking and rationale behind each step, and a summary at the end about what would be taken forward next for quantification.

The library does not include examples of quantification or calculated HEPs, as the library is intended to be independent of any particular quantification method. Development of guidance on how to apply quantification methods is also outside the scope of the Halden research project.

The task analysis library conceptis developed using Microsoft OneNote for a number of reasons, including: (i) the software is free to download and use; (ii) the software has a web interface for those who cannot download the local client; (iii) the library is stored on the cloud which means it can be accessed from anywhere that has an internet connection; (iv) the contents of the library (or parts of the library) can be password protected to prevent unauthorized access; and (v) the interface is relatively intuitive and easy-to-use.

The Microsoft OneNote platform also enables the analyst to access the library on multiple devices, which opens up the possibility of using the library template in the field on, for example, a tablet computer. For example, the analyst can document their field notes directly into OneNote, including photographs of the workspace and links to the operating procedures. The analyst also has quick access to the library examples to prompt the data collection process.

A test of the task analysis library concept is scheduled for spring 2017, to determine whether it is useful as a resource for non-experienced and less experienced HRA analysts, and to collect feedback for improvements to the concept.

4 Conclusions

There is a mismatch between what is described in HRA method guidance, and what is actually required in order to produce a useful and credible result, especially with respect to qualitative analysis. On further investigation, we identified that many of the difficulties that analysts experience with performing qualitative analysis are centered on application of task analysis techniques. However, these difficulties often relate to intangible or context-dependent aspects of task analysis, which are very difficult to write guidance for! Rather than trying to develop potentially complex and complicated guidance, we considered whether it would be better to teach by example, by providing analysts with a library of completed task analyses.

The concept of a task library to support HRA analysts is not new to the Halden Research Project. The idea is inspired by a similar initiative developed for the Petro-HRA project [9], and described in detail in [14]. However, the primary goal of the Petro-HRA task analysis library is to capture task analysis details for subsequent re-use in similar HRAs, where “each reviewer of the analysis benefits from the previous analysts’ insights while having the opportunity to add details that may have been overlooked in earlier analyses” [14].

The primary goal of our task analysis library is to provide less experienced HRA analysts with a template for how to do qualitative analysis for HRA, demonstrating what kind of data are typically collected, how these data are organized, interpreted, analyzed and used to develop an overall picture of the human operator response in a accident scenario. The task analysis library attempts to bridge the gap between what is generally considered as required for a comprehensive qualitative analysis, and what is actually described in HRA method guidance. Recent HRA publications, such as the Fire PRA guidance [7], the IDHEAS method [8] and the Petro-HRA method [9] are starting to address this gap by providing more detailed guidance on how to perform qualitative analysis and how to use qualitative results in an HRA. It is hoped that the task analysis library concept can provide a useful companion tool to this new guidance.

References

1. Forester, J., Dang, V.N., Bye, A., Lois, E., Massaiu, S., Broberg, H., Braarud, P.Ø., Boring, R., Männistö, I., Liao, H., Julius, J., Parry, G., Nelson, P.: The international HRA empirical study final report: lessons learned from comparing HRA methods predictions to HAMMLAB simulator data. HPR-373/NUREG-2127, OECD Halden Reactor Project (2013)
2. Liao, H., Forester, J., Dang, V.N., Bye, A., Lois, E., Chang, J.: Lessons learned from the US HRA empirical study. In: Probabilistic Safety Assessment and Management, PSAM 2012, Hawaii, 22–27 June 2014
3. Taylor, C.: Improving scenario analysis for HRA. In: Probabilistic Safety Assessment and Management, PSAM 2012 Hawaii, 22–27 June 2014
4. Boring, R.: How many performance shaping factors are necessary for human reliability analysis? In: Probabilistic Safety Assessment and Management, PSAM 2010, Seattle, Washington, 7–11 June 2010
5. Oxstrand, J., Kelly, D.L., Shen, S., Mosleh, A., Groth, K.M.: A model-based approach to HRA: qualitative analysis methodology. In: The 11th International Probabilistic Safety Assessment & Management Conference and ESREL 2012, Helsinki, Finland, 25–29 June 2012
6. Le Bot, P.: Human reliability data, human error and accident models – illustration through the three mile island accident analysis. Reliab. Eng. Syst. Saf. **83**, 153–157 (2004)
7. NUREG-1921: EPRI/NRC-RES Fire Human Reliability Analysis Guidelines, Final Report. EPRI Report Number 1023001, Electric Power Research Institute, California, July 2012
8. Parry, G.W., Forester, J.A., Dang, V.N., Hendrikson, S.M.L., Presley, M., Lois, E., Xing, J.: IDHEAS – a new approach for human reliability analysis. In: International Topical Meeting on Probabilistic Safety Assessment and Analysis, PSA 2013, Columbia, South Carolina, 22–26 September 2013
9. Bye, A., Lauman, K., Taylor, C., Rasmussen, M., Øie, S., van de Merwe, K., Øien, K., Boring, R., Paltrinieri, N., Wæro, I., Massaiu, S., Gould, K.: Petro-HRA, a new method for human reliability analysis in the petroleum industry. In: Probabilistic Safety Assessment and Management, PSAM 2013, Seoul, Korea, 2–7 October 2016
10. Taylor, C.: How HRA can drive plant improvement. In: 25th European Safety and Reliability Association Conference, ESREL 2015, Zurich, Switzerland, 7–10 September 2015
11. Taylor, C.: The importance of operator input to human reliability analysis. In: International Topical Meeting on Probabilistic Safety Assessment, PSA 2015, Sun Valley, Idaho, 26–30 April 2015
12. Kirwan, B.: A Guide to Practical Human Reliability Assessment. Taylor & Francis, London (1994)
13. Kirwan, B., Ainsworth, L.K.: A Guide to Task Analysis. Taylor & Francis, London (1992)
14. Boring, R.L.: A task library for petroleum applications in human reliability analysis. In: PSAM 13 Probabilistic Safety Assessment and Management, Seoul, Korea, 2–7 October 2016

Combination of Human Reliability Analysis and Task Analysis in Human Factors Engineering

Jinggong Liu(✉), Shiguang Deng, Xiufeng Tian, Xinwei Liu,
and Kunxiu Liu

China Nuclear Power Engineering Co., Ltd., No. 117 Xisanhuanbeilu,
Haidian District, Beijing, People's Republic of China
`{liujga,dengsg,tianxf,liuxwa,liukx}@cnpe.cc`

Abstract. As a risk-informed technique, Human Reliability Analysis (HRA) is widely applied in the Human Factors Engineering (HFE) activities of main control room (MCR) design, especially for the task analysis in the HFE, to identify the weakness of the design and the mechanisms of potential human errors. From the other side, task analysis also provides detailed sequences of human tasks, which can be used as the basis of the quantification of HRA. The paper shows how the HRA is combined with the task analysis, so as to improve both the qualities of HRA and task analysis in the HFE, and better direct the design activities to minimize the potential human errors.

Keywords: Human reliability analysis · Task analysis · Human factors engineering · Performance shaping factors

1 Introduction

It is required in the NUREG-0711, Rev. 3, "Human Factor Engineering Program Review Model" [1] that, more attentions need to be paid on the human actions that are critical to the safety of the operation of nuclear power plants (NPPs). These critical human actions, which may be performed during the mitigation of an accident, or performed during daily Maintenance, Test, Inspection, and Surveillance (MTIS) activities, are mainly identified from the PSA/HRA, and have to be considered in the HFE, so as to reduce the potential human errors and increase the safety of the plant operation.

Human errors may happen in any action steps of a human task for many reasons, such as improper Human-System Interfaces (HSIs) design, low operating procedure quality, unreasonable organization and management of the NPPs. As an element of the HFE, task analysis is mainly used to clarify detailed steps of human tasks and provide relevant information of these steps. However, such information only shows the content of the steps and the time spent on them, but cannot reflect the potential issues that may influence the human performance and lead to human errors. As a probabilistic method, human reliability analysis (HRA), involved in the probabilistic safety analysis (PSA), can identify the weakness of the performance of human tasks and evaluate the human

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_10

error probabilities, which provides a different way to recognize factors and mechanisms of potential human errors in each human task step.

On the other hand, task analysis also provides necessary information to HRA. The detailed steps identified and the time estimated for each step from the task analysis can be cited in the HRA as reference. Accordingly, it is of great importance to make good combination of the HRA and the task analysis in the HFE.

This paper is to establish a procedure to combine the HRA with the task analysis in the HFE, so as to identify weakness in the performance of human tasks and provide improving suggestions to utmostly avoid human errors.

2 Interaction Between HRA and Task Analysis

The NUREG-0711, Rev. 3, stipulates the interface between the HRA and the task analysis, that HRA, involved in the Treatment of Important Human Actions (IHAs), should select important actions and critical errors as inputs to the task analysis, and the task analysis, meanwhile, should provide the HRA with detailed task sequences as reference.

Specifically, the interaction between the HRA and the task analysis includes the following aspects:

- HRA indicates the risk importance of each human task, which provides a priority order of human tasks which can be served as criteria of scenario selections for task analysis.
- HRA methods provide a series of performance shaping factors (PSFs), including influence of HSI design, procedure quality, training level, organization, and other factors, which can well reflect the mechanism of human errors in each task step, so as to broaden the considerations of human performance in the task analysis.
- HRA provides assumptions of scenarios and requirements of critical actions such as time windows for the tasks, as inputs of the workload analysis in the task analysis.
- Task analysis identifies detailed sequences of human tasks in the given scenarios, and provides pertinent information of each step, which can be used as basis to evaluate the rationality of the assumptions and considerations in the HRA process.

It should be noticed that the interaction between the HRA and the task analysis is an iterative process, which means both the HRA and the task analysis results should be updated if there are any modification in any analysis results or practical designs.

3 Procedure of Combination of HRA and Task Analysis

Generally speaking, the process of the combination of HRA and task analysis can be summarized as the following steps:

1. Identification of Risk Important Human Actions (RIHAs);
2. Selection of task scenarios;
3. Confirmation of scenario strategies and task sequences;

4. Task sequence analysis;
5. Workload analysis;
6. Identification of potential human errors and mechanism analysis;
7. Verification and adjustment of HRA assumptions;
8. Summary report.

Figure 1 shows the flowchart of the combination of HRA and task analysis.

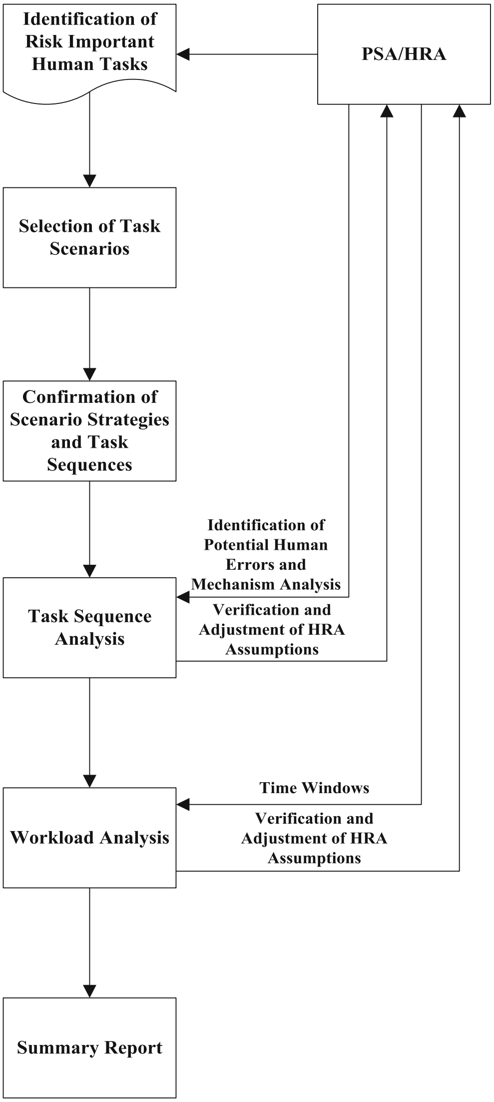


Fig. 1. Flowchart of the combination of HRA and task analysis

3.1 Identification of Risk Important Human Actions

PSA/HRA results are used to identify RIHAs, as part of the IHAs, which are served as inputs to the HFE activities. Identification criteria applied for RIHAs normally include the quantitative criteria and the qualitative criteria, where the quantitative criteria are defined from the quantification results of PSA/HRA, and the qualitative ones are given from the expert's judgment, which are beneficial complements to the quantitative criteria.

RIHAs identified from PSA/HRA are then provided as subset of the selection of task scenarios, with the form of a RIHAs list. The list should at least include the following information:

- Descriptions of each RIHA;
- Descriptions on accident scenarios related to each RIHA;
- Success criteria for each RIHA (such as critical steps and the time window);
- Other relevant information or assumptions about the RIHA (such as the applicable procedures, the persons in charge of the task).

3.2 Selection of Task Scenarios

RIHAs are required to be involved in the task scenarios and analyzed in the task analysis. For the reason that RIHAs are derived from the PSA model, the accident scenarios concerning the RIHAs are often supposed from the probabilistic view, which are difficult to be implemented in reality (e.g. component failures are often assumed as random process in the PSA model, but can only be considered as certain specific component failure mode in the deterministic analysis). Besides, some RIHAs drawn from the PSA/HRA have the same contents but only with different accident scenarios. These RIHAs may have to be considered separately in the PSA model, but do not need to be analyzed repeatedly in the task analysis. Thus, it is necessary to make a task scenario selection before the performance of task analyses on RIHAs, so that all the scenarios are well configured for task analyses, and redundant task scenarios are rationally screened out.

In addition, during the selection of task scenarios, it is also important to collect pertinent information regarding the scenarios and the tasks, with the format of a table or others, which is proposed to contain:

- Information of scenarios – brief descriptions of the task scenarios;
- Information of time window and success criteria – the available time for plant staff to recognize the accident scenario and finish the required tasks, and the criteria to judge whether the required tasks are successfully executed;
- Information of tasks – a brief description of the operator's task, including general mitigating strategy of the accident, the sequences of different task steps, and the interactions among different steps;
- Information of responsible personnel – the person who is mainly in charge of the task, such as the reactor operator or the turbine operator.

3.3 Confirmation of Scenario Strategies and Task Sequences

For each selected accident scenario in task analysis, it is necessary to further confirm the strategies of the accident treatment, and choose the suitable emergency operating procedures to respond the accident (in the case that procedures have not been finished, the procedures of predecessor or similar plants are also applicable). Moreover, for some complex accident scenarios, especially for those accidents with concurrent failures of mitigating systems, it is likely that multiple procedures need to be followed simultaneously, and switches from one procedure to another have to be performed to satisfy the requirement of accident mitigation. In this situation, it is also required to collect all the procedures regarding the accident treatment and the switch points among these procedures.

After the confirmation of scenario strategies and relevant information, further analysis need to be carried out to establish the task sequence for each scenario. The task sequence means a series of correlative tasks which have to be sequentially executed to accomplish the scenario treatment. The RIHAs identified from the PSA/HRA are obviously required to be involved in the task sequence.

3.4 Task Sequence Analysis

After the establishment of task sequence, task sequence analyses are enabled to be performed. Tasks involved in the task sequence are further divided into detailed action steps, and the steps are arranged in the order of time, so as to compose the action sequence of the task. Then the actions are associated with the relative controls, indicators and alarms, as well as the manipulating positions. All the information should be recorded, in order to facilitate the follow analyses.

In addition, at this stage, it is possible to identify potential human errors from each detailed step, and determine the probable mechanisms and consequences of human errors from the sight of HRA (such as the PSFs provided by different HRA methods). This information can also be collected as probable references for other HFE activities.

3.5 Workload Analysis

Once all the tasks and steps required for the scenario treatment are finally confirmed, it is enabled to estimate the time spent on each action, so as to calculate the nominal time for the operating staff to finish all their tasks in the scenario. The estimation is normally corresponding to the type of the human action, which is summarized from the empirical judgment and interviews with operating staff of NPPs.

After the calculation of the total nominal time for a task sequence, the workload analysis can be performed by comparing the estimated nominal time of the task sequence and the time window for the task sequence, to evaluate the workload of the operating staff to execute the required tasks. Ordinarily, the ratio of the nominal time to the time window is served as a quantitative index to evaluate whether the workload is acceptable, and a threshold is set for the ratio to make this judgment. If there are any task sequences determined with too high workload, then improvements have to be taken, from the aspects of function allocation, HSIs design, procedure development, or training program development, followed with modifications on corresponding assumptions applied in the HRA.

3.6 Identification of Potential Human Errors and Mechanism Analysis

As mentioned in Sect. 3.4, during the task sequence analysis, it is possible to identify potential human errors, determine the probable mechanisms, and evaluate the consequences of human errors for each step. This job is usually the responsibility of the HRA personnel. There have been more than 60 different PSFs employed in the existing HRA methods, which are great supplement to the considerations of the deterministic task analysis in the HFE. The mechanism analysis is used to find main reasons of the occurrence of human errors, based on the PSFs considered in HRA methods. Reference [2] summarizes typical PSFs that are commonly used in the existing HRA methods, as shown in the Table 1.

Table 1. Summary of common PSFs in HRA [2]

	PSF categories	PSF contents
1	Organization-based factors	1.1 Training Program 1.2 Corrective Action Program 1.3 Other Program 1.4 Safety Culture 1.5 Management Activities 1.5.1 Staffing 1.5.2 Task Scheduling 1.6 Workplace Adequacy 1.7 Problem Solving Resources 1.7.1 Procedures 1.7.2 Tools 1.7.3 Necessary Information
2	Team-based factors	2.1 Communication 2.2 Direct Supervision 2.3 Team Coordination 2.4 Team Cohesion 2.5 Role Awareness
3	Person-based factors	3.1 Attention 3.1.1 Attention to Task 3.1.2 Attention to Surroundings 3.2 Physical and Psychological Abilities 3.2.1 Fatigue 3.2.2 Alertness 3.3 Morale/Motivation/Attitude 3.3.1 Problem Solving Style 3.3.2 Information Use

(*continued*)

Table 1. (*continued*)

	PSF categories	PSF contents
		3.3.3 Prioritization 3.3.4 Compliance 3.4 Knowledge and Experience 3.5 Skills 3.6 Familiarity with Situation 3.7 Bias
4	Machine/Design-based factors	4.1 Human-System Interface 4.2 System Responses
5	Situation-based factors	5.1 External Environment 5.2 Hardware & Software Conditions 5.3 Task Load 5.4 Time Load 5.5 Other Loads 5.5.1 Non-task Load 5.5.2 Passive Information Load 5.6. Task Complexity
6	Stressor-based factors	6.1 Perceived Situation Severity 6.2 Perceived Situation Urgency 6.3 Perceived Decision Responsibility

3.7 Verification and Adjustment of HRA Assumptions

Since the HRA activities are performed earlier than the task analysis at the beginning to provide scenarios to the task analysis, it is inevitable to make some rational assumptions in the preliminary stage of HRA, such as the performance of the operating staff, the qualities of the HSIs and procedure, and, especially for the quantification analysis, the level of corresponding PSFs. These assumptions must be ensured and, if necessary, modified during the process of the task analysis, so that the results of the HRA are in accordance with the practical conditions.

3.8 Summary Report

After finishing all task analyses on the RIHAs, it is necessary to make a summary of the analysis results with the form of a report. The summary report should focus on the RIHAs that are determined with excessive workload by workload analyses, and the human errors together with the mechanisms identified from the action sequences of the tasks. In addition, it is also necessary to document and illustrate the HRA assumptions that are apparently inconsistent with the actual situation during task analyses in the summary report.

4 Conclusion

The paper establishes a procedure of the combination of HRA and task analysis in the HFE. In the procedure, HRA is in charge of providing RIHAs, selecting task scenarios, and identifying potential human errors with potential mechanisms from each detailed step of the task sequence, according to the perspective of commonly used PSFs in different HRA methods. Task analysis is performed to confirm strategies, construct task sequences, illustrate each detailed step of the task sequences, and evaluate the workload of the task sequences, so as to verify and make necessary adjustment of HRA assumptions.

Certainly, both HRA and task analysis are aiming at finding weakness in the HFE design, but it is no doubt that the combination of these two jobs can better accomplish this goal, for the reason that not only the deterministic opinions, but also the probabilistic ones are taken into consideration in the HFE activities.

References

1. O'Hara, J.M., Higgins, J.C., Fleger, S.A., Pieringer, P.A.: Human Factors Engineering Program Review Model (NUREG-0711, Rev. 3). Technical report, U.S Nuclear Regulatory Commission (2012)
2. Katrina, M.G.: A Data-informed Model of Performance Shaping Factors for Use in Human Reliability Analysis. Technical report, Sandia National Laboratories (2012)
3. Kirwan, B., Ainsworth L.K.: A Guide to Task Analysis: The Task Analysis Working Group. CRC Press (1992)

Human Error and Resilience in Safety-Critical Domains

From Reason and Rasmussen to Kahneman and Thaler: Styles of Thinking and Human Reliability in High Hazard Industries

Ronald W. McLeod(✉)

Ron McLeod Ltd., 16, Waterside Avenue, Glasgow G77 6TJ, Scotland, UK
ron@ronmcleod.com

Abstract. Much of the practical work conducted to minimise risk of human error and loss of human reliability, at least in oil and gas, chemicals and other process industries over the past 30 or so years, has been based on the model of basic error types known as the Generic Error Modelling System (GEMS). Over roughly the same period, psychologists and behavioural economists, have developed a rich understanding of the nature and characteristics of what, in simplified terms, are widely considered to be two styles of thinking – often referred to as "System 1" and "System 2". This paper explores the relationship between the GEMS model and what is known of the functioning of the two styles of thinking, and in particular the characteristics and biases associated with System 1 thinking. While many of the ideas behind the two styles of thinking are embedded in the GEMS model, there are some important omissions.

Keywords: Cognitive bias · System 1 thinking · Human error · GEMS · Human reliability · Human reliability analysis · Human failure modes · Performance shaping factors

1 Introduction

A frequent experience as one develops from adolescence into some form of adult maturity, is the recognition that in many situations, if you don't know what's going on, or don't understand something, you will not be the only one in that state of ignorance. If you are struggling to understand an issue, it is virtually certain others are struggling in exactly the same way. So it is often worth asking the "dumb" question.

This paper presents my "dumb" question. It may perhaps reflect a struggle only experienced by someone who has spent a career as a Human Factors specialist working in applied domains. Specialists working in an academic or a research environment may have experienced no such struggle – to them, the answer may well be clear, perhaps obvious.

The background to my question lies in the seeming difference between, on the one hand, the approach to identifying, analyzing and assessing the potential for human error that for at least the last four decades has been at the heart of so much work in human factors in major hazard industries. And on the other hand the understanding of how human beings – though perhaps not "experts" - make risk-based judgements, decisions

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_11

and choices in real-time, and the ways in which those judgements, decisions and choices can deviate from what is considered "rational" that has arisen in the field of Behavioural Economics supported by a significant part of mainstream Psychology.

So here is my "dumb" question – in fact two questions:

1. What exactly is the relationship between the two styles of thinking summarised by Daniel Kahneman [1] as "System 1" ("fast") and System 2 ("slow") and the widely used model of human error in which James Reason combined Jens Rasmussen's ideas about Skill-based, Rule-based and Knowledge-levels of performance into his Generic Error Modelling System (GEMS) [2]?
2. What is the relationship between System 1 thinking and the heuristics and biases associated with it, and the failure mechanisms and Performance Shaping Factors used in many approaches to Human Reliability Analysis (HRA)?

It is possible that these questions will not have resonance in industries such as Nuclear Power, Aviation and Air Traffic Management. Those industries have long recognised the central role that human cognition plays in system performance and reliability. Because of that, they have a history of employing cognitive engineers and Human Factors specialists with a background in cognitive science. The representation of cognition in system models and attempts to identify mechanisms of cognitive 'failures' in those industries, has been sophisticated.

But in industries including oil and gas, chemical, manufacturing, maritime, mining and even healthcare, investment in understanding and addressing the cognitive contribution to safety has not achieved the same levels. In most of those industries, the ideas expressed in the GEMS model and its modern variants continue to have great currency and remain widely used in safety analysis and incident investigations.

My "dumb" questions also may not have resonance with those in the Naturalistic Decision Making (NDM) community, who focus on the intuition and expertise in 'real-life' decision making situations and see little place for the kind of heuristics and biases associated with System 1 thinking. Kahneman and Gary Klein [3] have reviewed this area and reached some fundamental agreements. A great many operational situations with the potential for major accidents however do not meet the three requirements Kahneman and Klein concluded are necessary for the development of intuitive judgement by experts: (i) an environment that is sufficiently regular to be predictable, (ii) an opportunity to learn those regularities through prolonged practice and (iii) an environment that provides feedback that is both meaningful and available quickly enough for the individuals to be able to learn what works and what does not. A clear example is the situation surrounding the events leading up to the explosion and fire on the Deepwater Horizon drilling platform in 2010 and the fatalities, environmental, social and financial consequences that followed (see [4] for a review of Human and Organisational Factors associated with the incident).

1.1 An Example

To provide context for the discussion, it is worth considering an incident. On 22 January 2014, a railway lookout walked into the path of an oncoming train. Figure 1

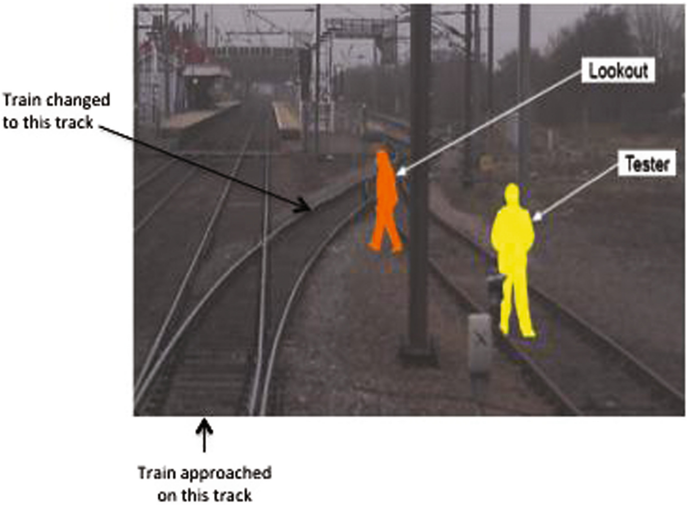


Fig. 1. Video image taken from approaching train. From [8].

illustrates the situation a few moments prior to the incident. The train driver sounded his horn, and the lookout raised an arm in acknowledgement. But it was not until about a second before he was struck and killed that the lookout turned and saw the oncoming train. (This incident is also discussed in [6, 7]).

In their investigation of this incident, the RAIB [8] concluded that the most likely explanation for the actions of the lookout was that he believed the train was going into platform 1 (in the middle of Fig. 1), and therefore that it was safe to move to the position he did: *"..the lookout was looking towards the train until about one second before the train started to deviate from the down main line onto the first crossover. At that point he turned his head away from the train and the purposeful manner in which he did so is consistent with him believing that he knew the path of the train, which at that point still looked to be straight into platform 1"* [6, p. 18]. Conclusions focused on competence assurance, and the need for front line managers to have time to implement procedures. There were also recommendations to improve site safety discipline and vigilance and planning of safe systems of work.

But the deeper question, that the investigation did not attempt to answer, is how is it that the lookout could come to hold such an erroneous belief? A belief that led a trained and experienced operator, who fully understood the risks, and had no apparent desire to cause harm to himself or anyone else, to deliberately walk in front of an oncoming train? What was he thinking? Or, more to the point, how was he thinking?

1.2 System 1 and System 2 Thinking

In 2011, Daniel Kahneman published his book "Thinking, fast and slow" [1]. I am going to use Kahneman's overview of the more than 40 years of research that forms the basis of his description of System 1 and System 2 thinking as the basis for this discussion. Some psychologists take issue with Kaheman's simplification of the two systems, though for the present purpose, his descriptions are more than adequate.

Thinking and decision making then (at least by non-expert decision makers) is proposed in a simplified sense as comprising two distinct systems, or styles, of mental activity: what are referred to as 'System 1' (fast) and 'System 2' (slow) thinking. System 1 is fast, intuitive and efficient and works through a process known as "associative coherence". Faced with a problem, judgement or decision, if System 1, drawing on a near instantaneous mental network of associations in which ideas or feelings trigger other ideas and feelings, is able to find a solution that feels comfortable, or coherent, then it will offer it to consciousness as a solution. System 1 is 'always on': we cannot turn it off. And it works automatically, requiring no effort or conscious control. Most of the time, System 1 works perfectly well. It is because of its speed and efficiency that we are able to think, act and perform in ways that would simply not be possible if the brain had slowly and carefully to seek, process and think through all of the information and options available to it.

But from the point of view of the levels of human reliability that are now demanded in safety critical industries, System 1 thinking has characteristics that are far from desirable. It does not recognize ambiguity, does not see doubt, and does not question or check. If the mental network can quickly produce an interpretation of the world or an answer to a problem that feels comfortable, it will take it. *'The measure of success for System 1 is the coherence of the story it manages to create. The amount and quality of the data on which the story is based are largely irrelevant. When information is scarce, which is a common occurrence, System 1 operates as a system for jumping to conclusions.'* [1, p. 79]. Conclusions that, sometimes, can have disastrous consequences.

System 2, by contrast, is slow, lazy and inefficient. But it is careful and rational. It takes conscious effort to turn it on and demands continuous attention: it is disrupted if attention is withdrawn. System 2 looks for evidence, reasons with it, takes the time to check, and questions assumptions. It is aware of doubt, and sees ambiguity where there is more than one possible interpretation of events or answers. Switching between System 1 and System 2 takes effort, especially if we are under time pressure.

From the Behavioural Economics perspective, Richard Thaler describes the two systems in the context of how self-control influences behaviour and decision making using the metaphor of a "Planner" (equivalent to System 2) and a "Doer" (equivalent to System 1). *"... at any point in time an individual consists of two selves. There is a forward-looking "planner" who has good intentions and cares about the future, and a devil-may-care "doer" who lives for the present"* [9, p. 104].

Drawing on a number of incidents as case studies McLeod [10, 11] has considered some of the potential operational implications of System 1 thinking on safety critical front-line operations. McLeod [11] also discusses the potential impact of system 1 biases on judgements and decisions about risk that are made away from the front line using risk assessment matrices (RAMs).

1.3 The Generic Error Modelling System

The Generic Error Modelling System (GEMS) was first described in James Reason's 1990 book 'Human Error' [2]. Figure 2 illustrates the dynamics of the model. GEMS provides a conceptual framework based around three generic error types that draw on Jens Rasmussen's classification of human performance into three levels; Skill-based slips and lapses, Rule-based mistakes, and Knowledge-based mistakes. The model summarises the error mechanisms operating at these three levels of performance based around two core assumptions;

1. Skill-based (SB) performance, and the error mechanisms associated with it, is assumed to precede the detection of a problem, and;
2. Rule-based (RB) and Knowledge-based (KB) performance (and the errors associated with them) are assumed to occur after a problem has been detected.

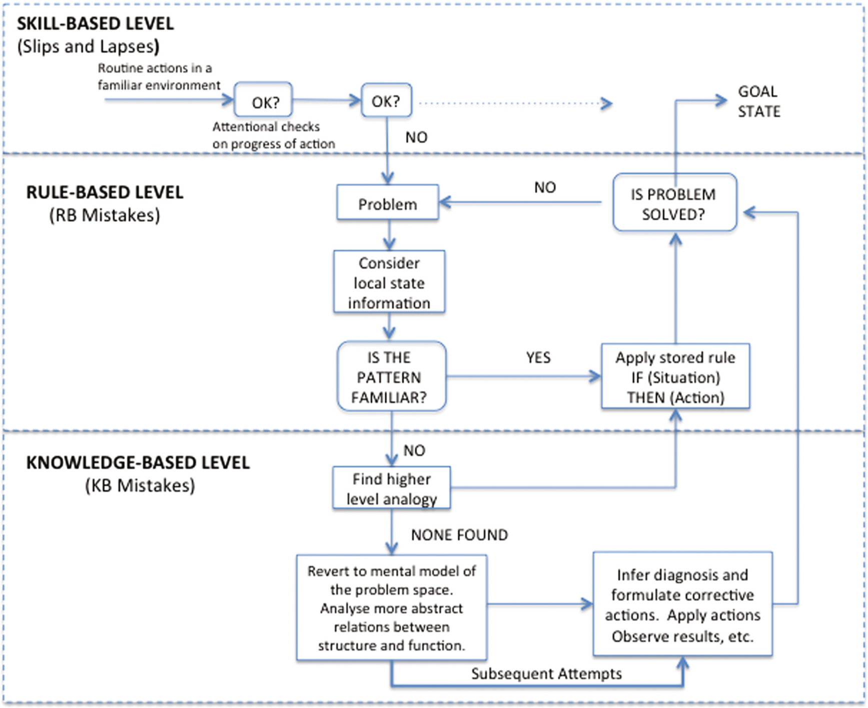


Fig. 2. Outline of the dynamics of the Generic Error Modelling System (GEMS) [2].

GEMS also considers how switching between the three performance levels occurs, and the mechmisms that can interfere with optimal switching.

Generically, GEMS considers errors to be associated with one of two types of failure: (i) failure of attentional processes to monitor the progress of pre-programmed behaviour sequences (both checking that actions are running according to plan, and whether a current plan is still adequate), or (ii) failure to solve problems detected during the monitoring process.

The three basic error types are distinguished in terms of at least eight dimensions ranging from the type of activity (whether or not the individual is involved in problem solving at the time an error occurred) to the extent to which each of the three error types is likely to be affected by some form of change from normal practice.

Finally, GEMS describes thirty-five potential failure modes that can be associated with the three levels of performance. These include inattention and overattention at the SB level, misapplication of good rules, or application of bad rules at the RB level, and limitations of the mental workspace and overconfidence, to problems due to the complexity of the task at the KB level.

2 GEMS and System 1 Thinking

The GEMS model is grounded in knowledge about a broad range of aspects of cognition and information processing. Much it draws on what at the time the model was developed was recent thinking and research into the ways that human judgement and decision making departs from an idealized view of people as economically optimal and rational decision makers: what were sometimes labeled “Econs” [9]. Indeed, most of the KB failure modes in GEMS, and many of those developed to support HRA since (see for example [12, 13]), draw heavily on research into biases and heuristics, including the work of Kahneman and his collaborators among many others. While the language used, and many of the assumptions and descriptions of the underlying psychological processes can be very different, it is clear that there is no real conflict: there is no question that the GEMS model and its modern variants and spin-offs recognize the importance of bias and reflect many – though, importantly, not all - of the characteristics of the two styles of thinking.

At least three clear conclusions can be drawn;

1. The idea of two systems of thinking is deeply embedded in, and fundamental to, the GEMS model, both in terms of error types and failure modes.
2. Similarly, the principle that people have a strong innate drive to seek to avoid cognitive strain is fundamental to GEMS.
3. The *“.simplifying shortcuts of intuitive thinking”* associated with System 1 can manifest themselves in errors at least at the RB and KB levels, and possibly also at the SB level.

2.1 Which Level of Error Does System 1 Thinking Produce?

Superficially, it is in Reason’s description of failure modes at the KB level (including ‘Confirmation Bias’, ‘Overconfidence’, ‘Biased Reviewing’, the ‘Halo effect’, and ‘Problems with Causality’) that the GEMS model appears to come closest to capturing System 1 thinking. For example, the KB failure mode ‘Out of sight out of mind’ is equivalent to both the availability heuristic and what Kahneman refers to as “what you see is all there is” (WYSIATI): the powerful tendency of System 1 to jump to conclusions based on limited evidence without considering what is missing.

It might therefore be concluded that the errors associated with System 1 thinking manifest themselves in GEMS primarily as KB mistakes. Such a conclusion however would be overly simplistic: the impact of System 1 thinking goes much deeper. It certainly impacts RB performance and the associated error types and could contribute to SB slips and lapses. Indeed, the following quote might suggest that System 1 thinking exerts its impact by producing errors at the SB and RB levels and that errors at the KB level arise from System 2 thinking: *“Errors at the SB and RB levels occur while behaviour is under the control of largely automatic units within the knowledge base. KB errors, on the other hand, happen when the individual has ‘run-out’ of applicable problem-solving routines and is forced to resort to attentional processing within the conscious workspace”* [2, p. 57].

The phrase *“..attentional processing within the conscious workspace”* sounds very like System 2 thinking. Also; *“The key feature of GEMS is the assertion that, when confronted with a problem, human beings are strongly biased to search for and find a prepackaged solution at the RB level before resorting to the far more effortful KB level, even when the latter is demanded at the outset”* [2, p. 65].

These quotes suggest, in contrast to what is implied in the KB level failure modes, that it is performance at the RB level that relies most heavily on System 1, and that System 2 is equivalent to the KB level. Similarly, Reason’s comment on how errors can occur in switching from RB to SB levels (see the link between “Is problem solved” and “Goal State” on Fig. 2): *“There will be powerful cognitive and affective forces conspiring to encourage the problem solver to accept inadequate or incomplete solutions as being satisfactory..”*, [2, p. 67] again sounds very similar to Kahneman’s System 1 heuristic “what you see is all there is” (WYSIAT).

Equating System 1 with errors at the RB level therefore is also overly simplistic. The reality is that the fingerprint of System 1 can be found at both RB and KB levels of performance and the associated GEMS error types. It may, indeed, also exert an impact at the SB level.

2.2 Is Anything Missing from GEMS?

There are however a number of aspects of how System 1 is thought to work, that can be powerful drivers of seemingly irrational decisions and actions that do not seem to be well reflected in the GEMS model. These include, among others: (i) the Substitution heuristic, and (ii) Framing and Loss Aversion. Space here only allows consideration of the first of these. The potential impact that Framing and Loss Aversion could have on human reliability in industrial operations is considered in [8].

One of the mechanisms System 1 uses to help in its search for coherence, and to avoid exerting effort, is what is referred to as “Substitution”. According to Kahneman: *“If a satisfactory answer to a hard question is not found quickly, System 1 will find a related question that is easier and will answer it”* [1, p. 97]. And it will do so without any conscious awareness of having substituted an easier question.

The need to be mindful, and to maintain constant awareness of risk has been much discussed as one of the key characteristics of High Reliability Organisations [14]. Organisations involved in safety critical industries at least expect and assume that everyone involved will be aware of the risks and hazards associated with their activities. One of the safety management system controls that is perhaps most widely shared across industries is variously termed a pre-job safety review, Job Hazard Analysis (JHA), Job Risk Assessment (JRA) or similar. The expectation is that before carrying out any work known to be hazardous or critical, the individuals involved will take the time to ensure they are aware of the hazards and risks involved, and that all of the barriers or other controls expected to be in place actually are in place and functional.

Consider for a moment what that simple expectation actually means in terms of the psychological demands on the people involved. People who, it must be assumed, will have the necessary competence and experience and who will be in a fit state to carry out the work. Who will be familiar with the equipment and work systems involved and may have carried out the same or similar operations successfully in the past, perhaps many times. And people who, statistically, are unlikely to have any direct personal experience of activities that had led to fatalities, serious injury or major damage or environmental loss. The expectations behind conducting a pre-job safety review/JHA, is that such people, by looking at the work situation and discussing and considering what is to be done, will correctly identify the hazards and risks and will be able to determine whether all the necessary controls are in place. The kind of mental judgements and decisions that need to be made to meet these expectations must include things like;

- Do we know what the major hazards associated with this activity are?
- Do we understand what could go wrong?
- Are all of the measures needed to control risk in place, and functional?
- Do we have all of the tools, resources, information and knowledge needed?
- Do my colleagues understand their roles and responsibilities and are they competent, capable and in a fit state to work safely?

If the answer to any of those questions – or indeed any number of others – was "No", the team would be expected to step back from the work and not continue until the issues were resolved. But those are hard questions. And if System 1 is involved, those hard questions will be substituted for easier ones. Perhaps questions like;

- Are we familiar with the equipment?
- Can we see anything unexpected or obviously not right?
- Can we remember or have we heard about a serious incident doing this job?
- Do I know and trust my colleagues?

Those questions are much easier for System 1 to answer. But they are not the questions that a pre-job safety review/ JHA expects and needs answered.

Pre-job safety reviews is just one example that illustrates how the process of Substitution has the potential to play a central role in leading people to make erroneous judgements and decisions in high-risk situations – including not being aware of the risks immediately in front of them. There seems nothing in the GEMS model that adequately reflects the importance of Substitution in producing error types, or that recognizes the kind of failure modes that may be associated with it.

3 Issue Arising

In terms of the practice of Human Factors in safety critical industries, there are at least three questions that arise from reflection on the relationship between GEMS and the two styles of thinking;

1. Does looking at human performance and reliability in major hazard industries through a System 1 ‘lens’ add any value?
2. Do the “failure modes” or “proximate causes” that are widely used in HRA take adequate account of what is known about the characteristics of System 1 thinking and how it can lead to erroneous judgements, decisions, and actions?
3. Is the way that Performance Shaping (or Influencing) Factors (PSFs or PIFs) are used in HRA adequate to capture the ways the immediate context of work might encourage System 1 thinking in critical situations?

3.1 Looking Through a System 1 “Lens”

Does looking at human performance and reliability through a System 1 “lens” adds any value? Value in the analysis and assessment of human error potential, in the way incidents are investigated and the learning that should follow, or in the actions that can be taken to avoid situations where people may not think or act in the ways that are expected and relied on to ensure safety. I think it does, in all three cases.

To return to the incident discussed earlier in which the rail lookout walked in front of a train and was killed (see Fig. 1). That incident has all of the characteristics of someone who was reasoning and making decisions using System 1. Someone who jumped to a conclusion about the future path of the train into the station: a conclusion about which he had no doubt, and that formed the basis for his subsequent actions and the fatal consequences that followed. If he had any doubt at all, he would surely have at least looked in the direction of the train before approaching the track. But System 1, Kahenman tells us, does not have doubt.

The “fast” and “slow” thinking perspective lends itself to a deep understanding of how and why the individuals involved may have come to make the judgements they did about risk – including failing to be aware of risk - and to take the decisions and act the way they did, in the context they found themselves at the time. It readily supports Sydney Decker’s principle of seeking to understand local rationality [15]; of trying to get inside the operators head and understand how what they did must have made sense to them at the time they made decisions and acted. McLeod [10] has demonstrated in some detail how looking through a System 1 “lens” can provide insight into the local rationality that may have led to major incidents.

3.2 Failure Modes and Proximate Causes

The differences between GEMS and the styles of thinking perspective perhaps come into clearest focus when considering the failure and error modes associated not only with GEMs, but with many of the approaches to HRA that are in widespread use.

In 2012 the US Office of Nuclear Regulatory Research published probably the most thorough and wide ranging review of the psychological foundations of HRA [12] since the classic report by Swain & Guttman in 1983 [5]. Based around five macrocognitive functions identified as being central to nuclear power operations, the NRR review sought to identify the psychological mechanisms that could lead to human error, and the "proximate causes" that can cause one or more of those macrocognitive functions to fail. A total of thirteen proximate causes and seventy one different mechanisms were identified. The types of bias associated with System 1 thinking (including confirmation bias, availability, anchoring and overconfidence) feature prominently among the underlying mechanisms.

However, even the NRR review does not capture the impact and immediacy with which System 1 processes can lead to erroneous decisions. To take one example, for the macrocognitive function '*Understanding and Sensemaking'* one of the three proximate causes of failure was identified as *"Incorrect data"*. One of its' mechanisms is *"Information in the environment is not complete, correct or accurate, or otherwise sufficient to create understanding of the situation"*. The example is used of an incident where a valve stuck open, though the control room display showed the valve to be closed: *"Without information from the plant indicators or procedures to guide them, operators did not understand the nature of the situation and made several inappropriate knowledge-based decisions"* [12, p. 186].

This is precisely the situation where, according to Kahneman, System 1 would use the "what you see is all there is" (WYSIATI) heuristic. If System 2 is not involved, System 1 will jump to a conclusion based on the information that is available, whether or not that information is sufficient to make a good decision.

Labeling this as a KB error is consistent with the GEMS model. However, recognising the likelihood that the operators System 1 thinking processes may have led them to jump to an erroneous conclusion is equally valid. Indeed, I would argue, of more value and providing deeper insight into understanding the risks involved when operators, in a situation where their thought processes may well be dominated by System 1, lack the information they need to diagnose a problem. This is very different indeed from a situation where the operators know, through interface design, experience or some form of task support, that they are lacking information or that information may not be valid, but they make the best decision they can (hopefully using System 2 thinking) in the circumstances.

3.3 Performance Shaping Factors

The use of Performance Shaping Factors, (PSFs, also referred to as Performance Influencing factors, PIFs) is central to every approach to HRA. The aim is to identify those aspects of the circumstances of performance that could change – increase or reduce - the likelihood that an error would be made compared with the situation where no PSFs were present. To take one example, the recently published 'PetroHRA' method [13] developed for use in the petroleum industry, uses nine PSFs (ranging from Time, Threat Stress and Task Complexity, to Attitudes to Safety, Teamwork and the Physical Working Environment).

To what extent is the use of PSFs in HRA as it is currently practiced, capable of capturing and modeling the situational factors that would make the operation of System 1 biases more or less likely? If you posed the question, "what are the situational factors most likely to trigger System 1 type errors of judgement at the moment when a critical decision needs to be made?" you might come up with answers such as;

- How difficult is the specific question or judgement at that moment? (If the question is hard, System 1, through Substitution, will answer an easier question).
- How much of the information necessary to make a good decision is available to the individual at that moment? (System 1 will to jump to a conclusion with the information available, even if it is incomplete or ambiguous).
- Are the relative risks associated with the situation at that moment framed in such a way that the individual could view one possible outcome as a gain and another one as a loss? (If so, could loss aversion influence the decision?).
- Is the individual emotionally engaged with or committed to a plan or course of action at that moment? (Such engagement is likely to lead to commitment bias).
- Could the individual believe, or could they be motivated to want or need to believe, that elements of the task world will be in a particular state at that moment? (If so, confirmation bias is likely).

What is striking, while at the same time perhaps being rather obvious, is that the contextual factors likely to cause heuristics and biases to lead to erroneous outcomes operate in real-time, inside the head of the individual who is making the judgement or decision. They are not factors that are readily identified from the level of description of PSFs that are commonly used in HRA. PSFs likely to lead to System 1 reasoning errors need to be tightly coupled both to the specific situation as the individual experiences it and believes it to be, at the moment when the decision is being made. This is reflected in the need to include the phrase "at that moment" in each of the items in the list above. And it is consistent with the emphasis Dekker [15] and others have placed on the importance of local rationality: of understanding how and why the decisions an individual made and the actions they took must have made sense to them at the time.

In GEMS, Reason explored how the immediate environment can trigger error types through what he termed "cognitive underspecification": *"When cognitive operations are underspecified, they tend to contextually appropriate high-frequency responses"* [2, p. 97]. From the point of view of finding means of mitigating risk from human error arising from System 1 thinking, I have argued that there is a compelling need to find ways of avoiding such cognitive-underspecification in work systems: *"There is a need to develop approaches that can be integrated into the design of work systems that are effective in breaking into System 1 thinking, and forcing people to adopt a System 2 style of thinking at the critical moments when they are interpreting information and making decisions in real-time"* [10, p. 392]. In HRA terms, PSFs are needed that reflect the existence of cognitive underspecification in the context of task performance and whether there are contextual factors likely to stimulate System 2 thinking.

At the extreme, this reflection on PSFs and System 1 reasoning errors leads to a troubling conclusion: that, because they are not sufficiently tightly coupled to the immediate context of thought, PSFs as they are currently conceived are not capable of adequately capturing the situational factors likely to trigger the heuristics and biases associated with System 1 thinking that can lead to errors in real-time, safety critical tasks. Indeed, other than by trying to simulate a very specific situation, it may be virtually impossible to capture or model that impact.

It is worth noting in passing that one of the PSFs that is most widely used in HRA – that of fatigue or lack of alertness – is likely to impact very directly on the likelihood of people making System 1 reasoning errors in real-time operational decision making. Perhaps the most immediate effect of fatigue is to reduce the fatigued persons willingness or ability to exert effort. System 2 however, by definition, requires effort. Although there seems to be little or no direct research evidence available, you do not need to be a psychologist to speculate that one of the likely consequences of fatigue will be to make the individual more prone to the kind of reasoning errors associated with System 1 thinking.

4 Conclusions

Safety management relies on trained and competent people in a fit state to work (not fatigued, stressed, etc.); people working in a culture that places a high value on safety and that empowers them to intervene or stop work when things are not as expected; and where work systems are designed, laid out and organised to support the tasks people are relied on to perform to a high level of reliability under a wide variety of situations. Critically, safety management at the sharp-end relies on and expects people to pay attention to detail, to check and avoid jumping to conclusions, to avoid complacency, and to take care and be sensitive to weak signals of the unexpected. As far as human reliability is concerned, it is the extent to which System 1 thinking is prone to bias and irrationality, its tendency to jump to conclusions, not to have doubt and not to see ambiguity that can represent such significant risk.

It is clear that many of the features and characteristics of the two styles of thinking, including many of the heuristics and biases associated with System 1 are deeply embedded in the theory, mechanisms and processes underlying the GEMS model. There are however significant areas where both GEMS and current approaches to HRA do not adequately capture the powerful heuristics associated with System 1 that can lead to systematic and widespread errors of judgement and decision-making.

References

1. Kahneman, D.: Thinking, Fast and Slow. Allen Lane, London (2011)
2. Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)

3. Kahneman, D., Klein, G.: Conditions for intuitive expertise: a failure to disagree. Am. Psychol. **64**(6), 515–526 (2009)
4. CSB: Investigation report Volume 3: Drilling rig explosion and fire at the Macondo well. Report no 2010-10-I-OS: Investigation report Volume 3. Chemical Safety Board (2016)
5. Swain, A.D., Guttman, H.E.: Handbook of human reliability analysis with emphasis on nuclear power plant applications. Final report, NUREG/CR-1278, USNRC (1983)
6. McLeod, R.W.: Implications of styles of thinking for risk awareness and decision-making in safety critical operations. Cognitia **22**(3), Human Factors and Ergonomics Society (2016)
7. McLeod, R.W.: Human factors in barrier management: hard truths and challenges. Process Saf. Environ. Perform (In Press). IChemE
8. RAIB: Fatal accident involving a track worker near Newark North gate station, 22 January 2014. Report 01/2015, Rail Accident Investigation Branch (2015)
9. Thaler, R.: Misbehaving: The Making of Behavioural Economics. Allen Lane, London (2015)
10. McLeod, R.W.: Designing for Human Reliability: Human Factors Engineering for the Oil, Gas and Process Industries. Gulf Professional Publishing, Houston (2015)
11. McLeod, R.W.: The impact of styles of thinking and cognitive bias on how people assess risk and make real world decisions in oil and gas companies. Oil and Gas Facilities, Society of Petroleum Engineers (2016)
12. Office of Nuclear Regulatory Research: Building a Psychological Foundation for Human Reliability Analysis. NUREG-2114, INL/EXT-11-23898 (2012)
13. IFE: The PetroHRA guideline. IFE/HR/E-2017/001, Institute for Energy Technology (2017)
14. Weick, K.E., Sutcliffe, K.M.: Managing the Unexpected: Resilient Performance in an Age of Uncertainty, 2nd edn. Jossey-Bass, San Francisco (2007)
15. Dekker, S.: The Field Guide to Understanding Human Error. Ashgate, Farnham (2006)

Epic Human Failure on June 30, 2013

Fred J. Schoeffler[1](✉) and Lance Honda[2]

[1] Sheff LLC, Pine, USA
[2] Prineville, USA

Abstract. Nineteen Prescott Fire Department, Granite Mountain Hot Shot (GMHS) wildland firefighters and supervisors (WFF), perished on the June 2013 Yarnell Hill Fire (YHF) in Arizona. The firefighters left their *Safety Zone* during forecast, outflow winds, triggering explosive fire behavior in drought-stressed chaparral. Why would an experienced WFF Crew, leave 'good black' and travel downslope through a brush-filled chimney, contrary to their training and experience? An organized *Serious Accident Investigation Team* (SAIT) found, "… no indication of negligence, reckless actions, or violations of policy or protocol." Despite this, many WFF professionals deemed the catastrophe, "… the final, fatal link, in a long chain of bad decisions with good outcomes." This paper is a theoretical and realistic examination of plausible, faulty, human decisions with prior good outcomes; internal and external impacts, influencing the GMHS; and two explanations for this catastrophe: *Individual Blame Logic* and *Organizational Function Logic*, and proposed preventive mitigations.

Keywords: Wildland fire · Hot Shot · Human Failure · Drift into failure

1 Introduction

Per the *National Wildfire Coordinating Group* (NWCG), "Wildland firefighting (WF) is a high-risk occupation, evidenced each year by deaths or injuries in the line of duty" [1]. One way the NWCG recognized to help reduce these WFF fatalities is to "… identify factors responsible for past fatalities, … [to] … mitigate those factors in future fire seasons" [1]. This current article presents essential details and inferences about the June 2013 Yarnell Hill Fire (YHF), which resulted in the disastrous outcome of 19 WFF fatalities. The authors' main goal is to provide a theoretical and realistic examination of the following subsections:

1. Wildland Firefighting Rules
2. Environmental Influences
3. Human Failure Theory
4. Organizational Cultures and Influences
5. Conclusion with Recommendations

The authors' motivation is twofold: (1) to identify WFF Decision Errors specific to the YHF to recognize the dynamics involved in the adverse outcomes of this fatal event, and (2) to apply this understanding for training, procedural, and/or systemic change recommendations to prevent future WF disasters. As difficult as this is, the

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_12

following paper attempts to move beyond the desire to blame the fatalities on a specific offender(s), also known as *Individual Blame Logic* (IBL), but to seek further understanding of the systemic causes triggering the outcomes, i.e., examining the entire WF organization, known as *Organizational Function Logic* (OFL) [2].

The content of this paper solely reflects the views of the authors, both retired Hot Shot Crew Superintendents, not those of any current or former Agency or Department. Moreover, the authors are responsible for the inferences, suggestive evidence, facts, and accuracy of the information presented herein. Some analyzed content may contain subject matter judged by some to be graphic, disturbing, and/or offensive.

2 Wildland Firefighting Rules

All WFFs are trained with specific rules, crucial to follow to ensure good direction, leadership, preparedness, and safety. The Standard Firefighting Orders, organized in a deliberate and sequential way, are to be implemented and applied systematically in all fire situations [3]. Abiding by the WFF Rules promotes good decisions and outcomes [3–7].

The most critical WFF lists of rules listed in our Incident Response Pocket Guide (IRPG) [8] consist of:

1. Standard Firefighting Orders
2. Eighteen Watch Out Situations
3. Downhill Checklist
4. Lookouts - Communications - Escape Routes - Safety Zones (LCES)
5. Common Denominators of Fire Behavior on Tragedy/ Near-Miss Fires
6. Wildland Urban Interface (WUI) Watch Outs

Regarding Entrapment Avoidance, a USFS Risk Management pamphlet states: “If firefighters follow the Standard Firefighting Orders and are alerted to the 18 Watch Out Situations, much of the risk of firefighting can be reduced” [3]. Memorizing, understanding, and following the *Ten Standard Fire Fighting Orders*; and likewise memorizing, recognizing, and mitigating the *18 Watch Out Situations*; and the other WFF Rules, are responsible for saving tens of thousands of WFF lives each-and-every fire season [7]. It is common knowledge and practice in the WF community that the WFF Rules do work and that all firefighters must know and apply them [3–7, 12].

3 Environmental Influences

3.1 Fuels, Fire Weather, and Topography

Firefighters discuss Wildland Fires in terms of fuels (what is burning), weather (what are the influences), and topography (where is it burning). The YH Fire occurred in central Arizona in rugged terrain covered in dense, decadent chaparral brush that had not burned since 1966, creating an explosive fuel bed with extremely high rates of spread and extreme resistance to control. It was critically hot and dry for weeks.

As with the YHF, the strongest and most variable winds occur during thunderstorms, and generate extreme downdrafts, micro-bursts, outflows, and gust fronts, which adversely affect fire behavior, thus exacerbating and seriously complicating WFF choices and safety [9].

4 Human Failure Theory

Human Failure frequently has significant consequences. The prescriptive Ten Standard Fire Orders, created in the 1950s, and in subsequent years the cautionary Watch Out Situations [10] are perceptive resolutions against wildland fire fatalities, addressing risk management, grouping them based on their importance in the following logical sequence: (1) Fire Behavior, (2) Fireline Safety, (3) Organizational Control, and (4) Fight Fire [3, 8, 10]. Annually, thousands of WFFs and Supervisors validate the WFF Rules during required fire refreshers and trainings. Almost every WFF agrees that the YHF tragedy would have been impossible for 19 men to have died accordingly had they held to these tried-and-trued WFF Rules [3, 7]. Arguing against the SAIT conclusions disclosed in the Serious Accident Investigation Report (SAIR), the authors examined the Human Failure associations of how, why, and when the GMHS considered their actions acceptable risk(s) for years, unsuspectingly and steadily heading toward a *drift into failure* [7, 9, 11].

4.1 Individual Blame and Organizational Factors

Catino (2008) established two distinct, possible approaches for explaining incident origin and dynamics: (1) Individual Blame Logic (IBL) and (2) Organizational Function Logic (OFL). Shown as two distinct reasons generating different outcomes, IBL suits societal demands to identify accident cause(s) and transgressor(s). Conversely, OFL is an organizational and functional approach, aimed at identifying the factors within the system supporting event occurrence. In the OFL method, expectations are similar events cannot recur or infrequently occur once these influences are removed [2].

The IBL method seeks out careless, inattentive individuals who are liable. In complex organizational systems, reprimanding an employee for an accident without examining the system deficiencies may entail inadvertently transferring the risk(s) to future employees [2]. The OFL emphasizes the avoidance of individual blame, however, it is dangerous to overlook legitimate individual responsibility. The collective approach may risk concealing accountability and avoiding necessary questions of where responsibility lies. It is possible to distort the emphasis in favor of wider organizational factors, avoiding individual fault(s), even when that is where it resides [2]. Clearly, both IBL and OFL infractions were present regarding the GMHS, based on a steady drift into failure from 2009 until 2013 [7, 11]. These same logics can be applied to all other WF fatality fires with similar conclusions.

Far-and-away the most critically devious and potentially treacherous decision and outcome pattern is the Bad Decisions with Good Outcomes. One can get away with this combination without any consequences whatsoever for an entire career, indeed an entire lifetime. It is professionally safe to assert that this has been a fatal combination on most, if not all, of the fatal wildland fires, where firefighters were killed by fire, throughout the history of wildland firefighting, excluding those from other environmental deaths, where WFF were killed by lightning, rocks, trees, and the like.

Southwestern New Mexico HS and writer Hannah Coolidge (2015) recounts her Superintendent, "… talk[ing] about 'bad decision/good outcome' scenarios—how it's easy, once you've developed bad firefighting habits, to forget how dangerous those habits are after engaging in them 'repeatedly without negative consequences'" [4].

In the WFF realm, *"Bad Decisions with Good Outcomes"* (Fig. 1) is also referred to as '*The Rule of 99*' and the '*Normalization of Deviance*,' coined by researcher Dianne Vaughan examining the Challenger Space Shuttle disaster [13]. The authors allege ongoing and recurring Bad Decisions with Good Outcomes for years as well as the subtle influencing Fire Department attitude toward structures, a priority over WFF safety, which likely swayed those ultimately responsible for the YHF fatalities [7].

Decisions	***Decisions***	***Decisions***	***Decisions***
GOOD	**GOOD**	**BAD**	**BAD**
GOOD	**BAD**	**GOOD**	**BAD**
Outcomes	***Outcomes***	***Outcomes***	***Outcomes***

Fig. 1. Decisions and outcomes matrix

The Prescott City Attorney (PCA) offered the following account of the survivor GMHS lookout's story, related to him by the PFD Wildland BC, who disagreed with the account [14]. While moving vehicles with the Blue Ridge HS (BRHS), the GMHS lookout allegedly overheard radio traffic between DIVS A and the GMHS supervisor, with 17 Crew members, atop a ridge in the black. In the radio call, DIVS A told the GMHS supervisor to leave "the black," which was safe, and join him at the BSR. The GMHS supervisor protested, saying such a move would be dangerous. The radio exchange turned into a dispute [14].

"My understanding of the argument between DIVS A and GMHS was that GMHS supervisor did not want to go down," said the PCA [14]. Per the PCA's account, the GMHS supervisor objected until DIVS A gave him a direct order to descend. The GMHS supervisor acceded to the command to relocate to the BSR. He told DIVS A that he thought it was a bad idea. During one of the final radio transmissions, the GMHS supervisor told DIVS A the Crew was not going to make it [14, 15]. Due to a scarcity of actual recordings, these GMHS Crew Net radio '*Discussing Our Options*' [14] excerpts are alleged and thus hearsay dialogue, where the 'Arizona Rule 803, Hearsay exceptions' applies [16], allowing it as suggestive evidence in this paper [7].

John Hopkins University researchers found, "By narrowing attention, … attention shifts from vision to audition caused increased activity in auditory cortex and decreased activity in visual cortex and vice versa, reflecting the effects of attention on sensory representations." The experiment was designed to create tunnel vision, but a completely unexpected event occurred. While *vision* was being tunneled, performance of the *audible* control center decreased" [17]. The researchers further found that tunneled vision leads to diminished hearing. Tunneled hearing led to diminished vision. The researchers concluded that a person intently listening to audible cues, like a radio or cell phone, could have diminished visual performance. In some cases, when the stress is severe enough, the hearing receptors in the brain may shut off completely, referred to as *auditory exclusion* [17]. This clearly relates to wildland firefighting.

Thus, the GMHS would primarily "*see*" the weather and fire behavior that they were focused on, however, their own brains may have sabotaged or delayed their ability to perceive and react to threats from those recognized hazards or even from the focus they were directing their attention to, typical of all humans engaged in this type of encounter [17]. Numerous cell phone and radio conversations occurred during the YH Fire, likely distracting them from truly '*seeing*' the emerging weather and fire behavior hazards and reevaluating priorities contributing to their steady drift into failure [11, 17]. The firefighters were aware of the forecast weather that morning; the risk factor was high enough to make them think twice, but they deferred to their own judgment, unknowingly falling victim to distractions. The weather forecast warned of considerable thunderstorm outflow wind danger, but the GMHS seemingly discounted those warnings, and left their Safety Zone at the worst possible time [7]. Arizona is well known for its dynamic, sometimes tragic large fires; late June is considered extremely hazardous, i.e. the Dude Fire (1990) when 6 WFF died, where severe outflow winds also set the stage for potential deathtraps, and where harsh Human Failure also won out.

During an October 2013 YH Fire Site Visit, a Southwest HS Crew Superintendent stated during the Integration Phase, where participants can share their emotions, conclusions, and such: "*this was the final, fatal link, in a long chain of bad decisions with good outcomes, we saw this coming for years*" [7] and about 8 other HS Superintendents spoke up and stated they had all attempted unsuccessfully over the years through peer pressure to get the GMHS to alter their unsafe actions [7]. So then, what happened on that fateful afternoon to distort these experienced wildland fire supervisors' individual and collective judgments to the degree they would put themselves, and their Crew, clearly in harm's way? Did they all perish in a predictable, and therefore, avoidable death-by-fire incident? Indeed, per the conclusion of these authors, they did. However, this was not an "*accident*" as some have continued to falsely believe and advocate [7, 9, 12]. Strict compliance with the WFF Rules bolsters WFF safety and for those whom they were are ultimately responsible [3–7].

4.2 Abilene Paradox and Groupthink

The Abilene Paradox [18] is based on a parable about a family in Texas in July on a hot, muggy day. Someone in the family suggests they take a trip to Abilene even though they all know their car has no air conditioning. Throughout the trip, they are all agreeing what a wonderful time they are having even though deep inside their hearts, they are really hating it [18]. The Abilene Paradox succinctly means: "Go Along to Get Along" and "Don't Rock the Boat" [18]. During a July 2013 News Conference at the

YHF Fatality Site, the PFD Wildland Fire BC commented: "*I would have followed them down there blindfolded*" [7, 19, 23]. "*They ... stuck together ... they saw and felt the same way ...*" [7, 19, 23]. Despite attempting to protect the integrity of his men and their decisions and justify the GMHS actions on June 30, 2013, these comments and others strongly suggest both the Abilene Paradox and Groupthink, both very hazardous attitudes [8].

Although the GMHS knew the stakes were high that day, none of them went to work on June 30, 2013 planning on injuring themselves or others. None of them thought this could happen. Likewise, they never considered they would be unnecessarily risking their lives or leaving family members brokenhearted for the rest of their lives. Unwittingly, they were merely a large and confident group with a herd mentality. *Groupthink,* occurs when a group makes faulty decisions because group pressures lead to a deterioration of "mental efficiency, reality testing, and moral judgment" [19]. Groups harmfully affected by Groupthink ignore safer alternatives such as occurred on the YHF when the GMHS decided to leave their Safety Zone [19]. Avalanche survivor Megan Michelson said once the plan to ski Tunnel Creek was made and they left the back-country gate it was difficult to think of removing herself from the group, [20]. There were broken conversations of how to manage the run but not a full group conversation on doing it safely [20]. These same '*broken conversations*' likely occurred amongst the GMHS in the black and along the route to the Boulder Springs Ranch (BSR), bolstering and intensifying the devious, insidious power of Groupthink [9, 20].

4.3 The Morality of Obeying Stupid Orders

In the WFF realm, one is told to obey orders unless they are one or more of the following: (1) unsafe, (2) illegal, (3) unethical, or (4) immoral. Vietnam Veteran Reed argues to obey orders where the overall mission big-picture benefits are sufficient to warrant the risks. "If the superior won't back down, the lower leader has a moral decision to make. If the lower leader thinks the mission is too dangerous for its benefits, he should resist the order to the point of refusing to carry it out" [21]. This is exactly what the Acting GMHS Superintendent was doing, however, he was using '*mitigating or hinting speech,*' defined as "any attempt to downplay or sugarcoat the meaning of what is being said," likely in deference to authority, instead of direct speech and actions to hold true to their ultimate obligation of maintaining the safety and welfare of those they supervise [22]. Gladwell described this in detail regarding a fatal aircraft mishap when the Junior Pilot conceded to the Chief Pilot regarding escalating inclement weather, rather than being more direct and persuasive in his safety concerns [22].

4.4 Public Decisions

"A factor that binds people to their actions is '*going public*' when someone participates in and is *identified publicly with a decision*, that person will resolve inconsistencies to produce attitudes consistent with that choice" [13]. The authors contend that the GMHS '*public decision*' to stay put in the black and not assist the Yarnell Structure Protection

Specialist and to send the BRHS instead [7, 13] were overridden by what the authors define as the stronger '*micro-level public decisions*' during their discreet Crew Net "*discussing our options*" radio conversations [7, 9, 13, 14].

5 Organizational Culture and Influences

5.1 Human Factor Barriers to Situation Awareness

The IRPG [8] lists several Hazardous Attitudes in the overarching '*Human Factor Barriers to Situation Awareness*' section. They include:

1. *Invulnerable* - That can't happen to us
2. *Anti-authority* - Disregard of the team effort
3. *Impulsive* - Do something even if it's wrong
4. *Macho* - Trying to impress or prove something
5. *Groupthink* - Afraid to speak up or disagree

The authors contend these hazardous attitudes exist in varying degrees in all WF fatalities where they are killed by fire. They reigned fatal that day, manifested as perilous attitudes, decision making errors, and engagements, ultimately resulting in their deaths [6, 7]. The authors and other WFF often expand on the Anti-authority attitude, to include: "*The rules don't apply to us*" and "*Don't tell us what to do*." The authors argue that all the hazardous attitudes applied that day because they had gotten so used to bad decisions with good outcomes for years, and this was merely normal work to them [6, 7]. The insidious Groupthink mindset, discussed in more detail above, prevailed that fatal day, and very likely a major cause of many Wildland Fire Human Failures.

5.2 PFD Attitude and Influence

The PFD Wildland Battalion Chief (WBC) literally considered these GMHS men to be his sons; he was virtually in shock, attempting to defend their fatal actions as well as the Fire Department in a July 2013 New Conference at the YH Fire Fatality Site. He essentially held structure protection as a higher priority than individual firefighter safety when he stated: '*no WFF is satisfied sitting in a Safety Zone while structures were being threatened*' [23]. In addition, he used the Fallacies of Equivocation and False Analogy [7] by maintaining what the GMHS did was identical to firemen running into burning buildings [7, 23]. Indeed, municipal/structural firemen have much heavier personal protective clothing, SCBA oxygen tanks, and much more water to apply to the fire compared to WFF. All these are strong indicators of mixed and/or conflicting values contributing to confusion, frustration, and uncertainty, setting up for potential human failures on any wildfire, mainly those with threatened structures, as on the YHF and many others, and likely to continue with future fires unless WFF truly learn the costly lessons of these recurring tragedies [11].

5.3 The Marine Corps Viewpoint and Influence

The three GMHS Marines garnered immense respect from their Crew as well as considerable sway over them. The Marine ethos is one of a selfless group culture, where the group is held in a higher regard than the individual [24]. A feature story on the GMHS Marines stated, they "*went into the toughest fight they had ever faced with no second-guesses. They had a mission to accomplish—protect the community of Yarnell—and just like their time* in *the Corps, they were willing to lay down their lives to achieve that goal*" [25]. Indeed, honorable and laudable traits - until June 30, 2013. Truly, this is a rather dramatic statement, and hopefully merely intended to soothe the grief and pain of family, friends, and loved ones.

A Military.com video interview with the PFD WBC reveals: "*at about 4:00 … I clicked back on their frequency, and heard a couple things, … that they need to move, that the fire was coming back on their position*" [26]. This contradicts the SAIT SAIR stating there was a "*gap of over 30 minutes in the information available for the GMHS* [9]." There was compliant GMHS communications at times, however, when they were on the move leaving the good black, the SAIT SAIR alleges there was none [9].

It is also noteworthy the PFD WBC listened in on their Crew Net frequency, which suggests he heard some '*Discussing our options*' discreet radio conversations [14], talking about whether to stay in the good, safe black or head down through the unburned fuel in chutes and chimneys toward the BSR. It makes no sense. The GMHS were perfectly safe, with the best vantage point of anyone on the YH Fire, save Air Attack.

5.4 High Reliability Organizations (HRO)

HRO's, such as aircraft carrier and nuclear power plant operations, are preoccupied with all failures, especially small ones. Insignificant things that go wrong are often early warning signals of deepening trouble and give insight into the health of the entire system. If you catch problems before they grow bigger, you have more possible ways to deal with them. Instead, we tend to overlook our failures, suggesting we are incompetent, focusing instead on successes, suggesting we are competent [11, 27]. For many WFF, this apparently is accepted guidance in theory, and not realistically, due to the recurring WFF fatalities for the same fatal causes based on bad decisions.

The primary author was the Operations Specialist for a 1996 Fire Shelter Deployment Investigation that occurred on the Coconino NF outside Flagstaff, AZ [28]. The Human Factors Specialist briefing began with, "*The first thing we are going to do is establish a conclusion, then find the facts to fit that conclusion.*" The question was broached: 'A*ren't we supposed to follow the facts to lead us to a conclusion?*' The HF Specialist reiterated, we would establish a conclusion, then find the facts to fit it. The discussion concluded with: *"Then we could write anything we wanted to*" [7, 28].

The YHF Serious Accident Investigation Team (SAIT) followed that format of first establishing a conclusion when they discovered "*no indication of negligence, reckless actions, or violations of policy or protocol*" [7, 9]. Stated in the affirmative, this means they did it all right. Yet 19 men died. It is impossible to do everything right on a wildland fire and 19 men burn to death in one fell swoop. That ‘conclusion’ smacks of Orwellian Doublespeak and even worse, Doublethink [29]. “To respect words and their conveyance of reality is to show respect to the very foundation of reality. To manipulate words is to seek to *manipulate truth* and to instead *choose falsity* and illusion over reality. The manipulation of words is itself a violent act. …” [30]. The authors confidently argue, this ongoing, unsettling truth manipulation [30] has occurred with all wildfire SAIT Reports, Reviews, and the like where WFF were killed by fire [7].

Borrowing from Avalanche Fatality Human Factors, the “*Expert Halo*” heuristic comes to mind, in which the experienced believe that their expertise will keep them safe [31]. Several Individual, Organizational, and Social Factors likely contributed to and influenced the PFD, including the emphasis on Structure Protection, setting up WFF safety goal conflicts [11]. Avalanche human factors researcher McCammon writes that the *Default* or *consistency heuristic* concludes that when a venture is still an abstract notion, there are likely discussions about the various conditions, the pros/cons, and deliberate assessments of the risks of proceeding take place. But once the initial decision is made, deliberate calculations stop, that thought takes on deceptive power [31]. People generally have a powerful bias for sticking with what they already have, not switching course, and they let their minds default to what is given or what has already been decided. They rely on stay-the-course impulses all the time, often with deceptively satisfactory results [31, 32]. This was likely what happened on June 30, 2013, and other fatal WF, and cogently describes the bad decisions with good outcomes process.

6 Conclusion

The paper provided general information for in-depth discussions, education, and training. The authors and most WFF supervisors believe that WFF must do the following to practice responsible leadership and safe WFF practices:

1. Commit to memory, understanding, and following the Ten Standard Fire Fighting Orders as well as the 18 Watch Out Situations;
2. Memorize, understand, and follow the directives in the Downhill Checklist;
3. Memorize, know, and mitigate the Common Denominators of Tragedy Fires, including Human Factors.

As difficult as it may be, all WFF must move beyond blaming fatalities on an offender(s), and instead we must seek to understand the systemic causes triggering the outcomes, as concluded by Reason (1997) [33]. To reduce the frequency of WFF fatalities, a complete understanding of the human failures that led to these outcomes, including the human failures role, is critical. Continued identification of WFF Decision Errors and complete comprehension of the dynamics involved in both positive and negative outcomes are necessary. Furthermore, to truly improve WFF safety and performance, we must continue to know and strictly follow the basic WFF Rules discussed above and apply the following proposals regarding improved future training, existing procedural checklists, and overall support systems. Paraphrasing heavily from the municipal Fresno F.D. (2015) Cortland Incident Findings and Recommendations, some suggestions generated from that persuasive investigation and this current article include:

1. Ensure effective operational Leadership, Human Factors, and Group Dynamics training for all WFF and Supervisors;
2. Effectively evaluate and cultivate competent, experienced, worthy WFF leaders [34];
3. Reconsider and substitute the flawed practice of simply accepting the WFF injury and death causes as just being part of the job, or the cost of doing business [34];
4. Disapprove of freelancing and independent action, i.e., when individuals or groups work independently and commit to tasks without the expressed knowledge and/or consent of their supervisor [34];
5. Effect positive change to avoid future WFF fatalities by requiring all WFF at all levels to memorize, recognize, apply, and mitigate the WFF Rules on every WF assignment, e.g. 10 Standard Firefighting Orders, 18 Watch Out Situations, LCES, Common Denominators of Fatality Fires, and the Downhill Line Construction Checklist [8], and
6. To effectively change a culture, this transition will succeed only when followed by, supported by, and mandated with a "top down" mentality [34], i.e., managers, supervisors, and firefighters working collaboratively, "*talking the talk* <u>*and*</u> *walking the walk.*"

While leaders at all levels of the organization must make their expectations clear, trust and accountability by all those within the chain of command will be critical to any overall success. The desired change must be clearly communicated, and those in positions of authority at all levels must be held accountable to set a good example, in addition to leading, supporting, enforcing, and creating these expectations [34].

Acknowledgements. The authors wish to express their appreciation to: Dr. Martha Underwood, Patricia Burch, and Jill Cobb for their proficient knowledge and skills in managing this paper; Dr. Ted Putnam for his Human Factors insight; sources that are mandated to be uncooperative, due to Department and/or Agency 'YHF direction,' fear, threats, and/or intimidations; and sources that know a lot, keenly choosing to remain silent. To those who lost family, friends, and loved ones on the YHF, the authors and most WFF think about those young men daily, and this is a means toward healing and optimistically, averting similar future wildfire tragedies.

References

1. Mangan, R.: Wildland firefighter fatalities in the United States: 1990–2006. NWCG, PMS 841. Boise, ID. National Wildfire Coordinating Group, Safety and Health Working Team, National Interagency Fire Center, 28 p. (2007)
2. Catino, M.: A review of literature: individual blame vs. organizational function logics in accident analysis. J. Contingencies Crisis Manag. **16**(1), 53–62 (2008)
3. U.S.D.A. Forest Service, Fire and Aviation Management: Risk Management (2017)
4. Coolidge, H.: Rules and Risk in Wildland Firefighting. Hypocrite Reader **52** (2015). http://hypocritereader.com/52/rules-and-risk

5. Holmstrom, M.: Common denominators on tragedy fires – updated for a new (human) fire environment. Wildfire Magazine, pp. 26–34. International Association of Wildland Fire (2016)
6. Orasanu, J., Martin, L., Davison, J.: Errors in aviation decision making: bad decisions or bad luck? In: Fourth Conference on Naturalistic Decision Making, 8 p. The Airlie Conference Center, Warrenton (1998)
7. Schoeffler, F.J.: Human Factors Influenced the 30 June 2013 Yarnell Hill Fire Fatalities. Central Arizona Wildfire Response Team, RT-130 Refresher. Power Point (2016)
8. National Wildfire Coordinating Group (NWCG), Incident Response Pocket Guide (IRPG). PMS 461, NFES 001077 (2014)
9. Yarnell Hill Fire Serious Accident Investigation Team (SAIT): Yarnell Hill Fire Serious Accident Investigation Report (SAIR), 122 p. (2013)
10. Ziegler, J.A.: The story behind an organizational list: a genealogy of Wildland Firefighters' 10 standard fire orders. Commun. Monogr. **74**(4), 415–442 (2007)
11. Dekker, S.: Safety Differently: Human Factors for a New Era, 2nd edn. (2015)
12. Stanton, B.: Western Wildfire Fatalities: Lessons Learned from Tragedy in Yarnell, AZ. Times-News, Idaho Falls (2013)
13. Vaughan, D.: The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA. Univ. of Chicago Press (1996)
14. Anglen, R., Wagner, D., Sanchez, Y.W.: Yarnell fire: new account of hotshot deaths. Arizona Republic (2015)
15. Sanchez, Y.W., Wagner, D.: Yarnell Hill Fire lawsuits settle for $670,000, reforms (2015). azcentral.com
16. Westlaw: Arizona Rules of Evidence, Article VII, Rule 803. Exceptions to the Rule Against Hearsay – Regardless of Whether the Declarant is Available as a Witness. Arizona Court Rules (2015)
17. Shomstein, S., Yantis, S.: Control of attention shifts between vision and audition in human cortex. J. Neurosci. **24**, 10702–10706 (2004)
18. Harvey, P., Martinko, M.J.: An empirical examination of the role of attributions in psychological entitlement and its outcomes. J. Organ. Behav. **30**(4), 459–476 (2008)
19. Janis, I.: Victims of Groupthink: A Psychological Study of Foreign Policy Decisions and Disasters. Houghton Mifflin, Boston (1972)
20. Michelson, M.: Tunnel Vision. Outside Magazine, 10 p. (2012)
21. Reed, J.T.: The Morality of Obeying Stupid Orders (2015). johntreed.com
22. Gladwell, M.: Outliers: The Story of Success. Little Brown, New York (2008)
23. Dougherty, J.: Granite Mountain Hot Shot Deployment Site News Conference, Parts 1 & 2. YouTube (2013)
24. Grice, M.: Lt. Col. USMC (retired): Imbuing Culture – How the Marines Do It. The Decisive Leadership Group (2013)
25. Anderson, C.: Marine Veterans Among 19 Firefighters Killed (2013). Military.com
26. Underwood, C.: Marine Vet Killed Fighting Wildfire (2013). Military.com, video
27. Weick, K.E., Sutcliffe, K.M.: Managing the Unexpected: Resilient Performance in an Age of Uncertainty, 2nd edn. (2007)
28. U.S. Forest Service: Hochderffer Hills Fire Shelter Deployment Investigation, Coconino National Forest, 31 p. (1996)
29. Orwell, G.: Secker and Warburg, London, Printed in the U.S. by Donnelley and Sons Company, pp. 741–925 (1949)
30. Cummins, M.: Words: Pieper's "Abuse of Language, Abuse of Power". The Alternative Path (2011)

31. McCammon, I.: Heuristic traps in recreational avalanche accidents: evidence and implications. Avalanche News **68**, 10 p. (2004)
32. Kelley, B.: Goal Setting and Risk Perception: An Accident Analysis of the Tunnel Creek Avalanche (2016). sportgevity.com
33. Reason, J.T.: Managing the Risks of Organizational Accidents. Ashgate Publishing, Aldershot (1997)
34. Fresno Fire Department: Cortland Incident. Multi-agency Serious Accident Review Team, SART Investigation Report. 274 p. (2015)

A Systematic Framework for Root-Cause Analysis of the Aliso Canyon Gas Leak Using the AcciMap Methodology

Maryam Tabibzadeh[1(✉)], Simone Stavros[2], Mikhil S. Ashtekar[1], and Najmedin Meshkati[2]

[1] Department of Manufacturing Systems Engineering and Management,
California State University,
Northridge, 18111 Nordhoff Street, Northridge, CA 91330, USA
`maryam.tabibzadeh@csun.edu`
[2] Daniel J. Epstein Department of Industrial and Systems Engineering,
University of Southern California, 3715 McClintock Avenue,
Los Angeles, CA 90089, USA

Abstract. According to the US Energy Information Administration [1], the natural gas industry supports 33% of electricity generation in the US. Despite this critical role, the importance of safety and safety culture in the natural gas industry has not been adequately highlighted. The absence of strict regulations and lack of attention towards precautionary measures have allowed the industry to persevere with insufficient urgency for implementing innovative technologies and safety-first protocols. On October 23, 2015, the Aliso Canyon natural gas accident highlighted how the lack of regulatory oversight in a low probability, high consequence industry could have such impactful and unpredictable repercussions. This paper analyzes the concatenation of events that led to the Aliso Canyon gas leak. It adopts the AcciMap methodology, which was originally introduced by Rasmussen in 1997 as an accident investigation framework, to conduct a systematic root-cause analysis and capture different involved socio-technical factors that contributed to the leak.

Keywords: Root-Cause analysis · Risk management · Jens rasmussen · Accimap · Accident investigation · Aliso canyon gas leak · Natural gas

1 Introduction

1.1 Overview

Southern California Gas (SoCal Gas) Company, which owns 115 wells in the Santa Susana Mountains of Aliso Canyon in Porter Ranch, CA, has been at the center of a major environmental scandal that has shed light on the natural gas industry and safety culture in its entirety. SoCal Gas, a subsidiary of utility company Sempra Energy, is being held responsible for inadequate operations that ultimately led to a four month long natural gas leak, beginning in October 2015, that has affected the community, the company, the natural gas industry, national and state regulations and the environment

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_13

in a detrimental way. It is estimated that the leak emitted 97,100 metric tons of methane, the equivalent of 2.1 million metric tons of carbon dioxide into the atmosphere [2]. This is more air pollution than 440,000 cars emit in a single year.

This paper aims not only to discover what happened that led up to the accident at the Aliso Canyon, but why it happened and how it could have been prevented. Moreover, this paper does not aim to find culpability in SoCal Gas management or operators, but rather to suggest a safer process-oriented solution on how to perform with proper conduct.

The Columbia Accident Investigation Board (CAIB) [3, p. 6], in their analysis of the Columbia Space Shuttle accident, declared: “complex systems almost always fail in complex ways, and we believe it would be wrong to reduce the complexities and weaknesses associated with the system to some simple explanation”. They strived to show that accidents do not occur only from the final interaction of the failure, but more often, the entire process is culpable for weakening a situation. This paper analyzes the progression of inadequate protocols that ultimately led to the Aliso Canyon accident.

In this paper, we propose to investigate and analyze the events that led up to the Aliso Canyon gas leak. We utilizes the AcciMap methodology, which was originally proposed by Jens Rasmussen in 1997 [4], to analyze the accident and the interaction of decision-making in multiple levels of a socio-technical system. By using the AcciMap methodology, it is possible to take the findings from this specific case study and apply it to the entire industry in an effort to enforce preventative measures and promote a stricter safety culture with proactive risk management.

1.2 Background

Beginning October 23, 2015, an undermined well allowed for a massive natural gas leak that continuously leaked from a gas storage facility reserve for four months. The well, Standard Sesnon 25 (SS-25), is 61 years old, 8750 feet deep and was once used as an oil storage field until it was drained in the 70 s. The well previously had a safety valve until it was removed in 1979 and was never replaced [5]. Well SS-25 is part of one of the nation’s largest containment locations, with the capacity to store 86 billion cubic feet of natural gas. The Aliso Canyon reservoir stores enough gas for distribution to nearly 22 million customers in the LA area and supports 17 power plants in the LA basin [6].

It is believed that SoCal Gas was pumping methane into the reservoir at beyond secure limits through the well’s main pump as well as the concrete outer casing. This increased demand in pressure possibly caused extreme metal fatigue that weakened the system and resulted in the concrete casing being undermined. The well, which has an inch-and-a-half thick concrete casing that surrounds the half-inch thick steel gas pipe, was being worked to the upper limits of 4500 lb per square inch (psi).

1.3 Kill Procedure Attempts

SoCal Gas mobilized internal crew and equipment to address the well failure. There had been multiple attempts to stop the leak using well kill procedures that involved pumping mud and brine down the well. Six attempts were conducted with minimal research of how the entire system would react; these hastily attempted were an ill-advised effort to slow the leaking. These abatement attempts further weakened the well and increased the possibility of a massive blowout. Boots and Coots, an expert well control service that offers services such as blowout response procedures and well kill operations, were called in to facilitate a proper kill procedure [7, p. 19]. Following, a relief well began construction on December 4, 2015. This well is approximately 1500 feet away from Well SS-25 at an angle, with the goal of hitting the main well at its capstone. The relief well was designed to vent the gas to relieve pressure from the main well, while cement would then be poured down the main well to seal it off. On February 11, 2016, after 4 months of leaking, the relief well pierced Well SS-25 8,500 feet below the ground's surface through the concrete casing and workers started injecting a mud-like compound.

2 Rasmussen's Risk Management Framework and AcciMap Methodology

There have been several developed methodologies to better understand and analyze accidents. Some examples of these methodologies include the Systems-Theoretic Accident Model and Processes (STAMP) by Leveson [8], Reason's model of organizational accidents [9] and Rasmussen's AcciMap approach [4]. Rasmussen's AcciMap approach is particularly useful for this purpose as it models different contributing factors of an accident, and their interactions, in a causal diagram.

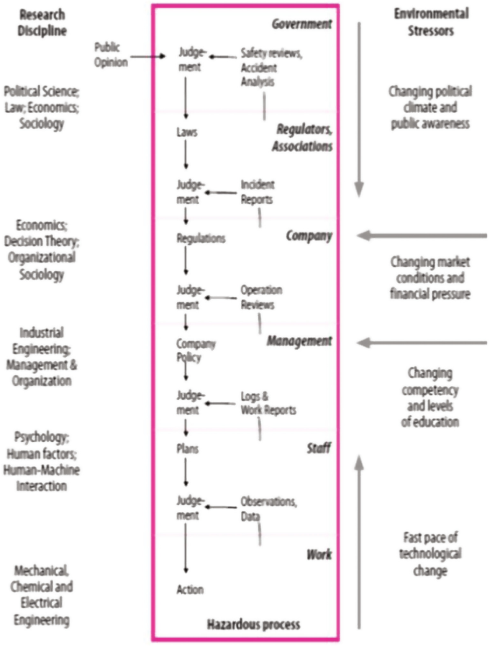


Fig. 1. Rasmussen's risk management framework [11]

Rasmussen has introduced a 6-layer, hierarchical framework (Fig. 1), known as risk management framework, with each level representing a main group of involved decision-makers, players or stakeholders in a studied system [4]. These six levels, from top to bottom, are: government, regulators and associations, company, management, staff, and work. Analysis using this framework not only considers the activities of players in each level but more importantly, the interactions between them, which take the form of decisions propagating downward and information propagating upward [10, 12].

The AcciMap methodology was developed by Professor Jens Rasmussen [4] in conjunction with his 6-layer risk management framework, which was illustrated in Fig. 1. This methodology captures the associated socio-technical factors of an accident within an integrated framework and analyzes the contribution of those factors in causing the accident. This graphical representation is useful in structuring the analyses of hazardous work systems and in identifying the interactions between different levels of decision-makers, which shape the landscape in which accidents may "unfold" themselves [11].

It is noteworthy that AcciMap is part of a broader proactive risk management process to develop risk assessment strategies from generalizing the analysis of previous accidents [10]. In general, analysis of past accidents within the stated framework can define patterns of hazards within an industrial sector. Such analysis can lead to the definition of preconditions for safe operations, which is a main focus of proactive risk management systems.

In the context of the natural gas industry, to our knowledge, the AcciMap methodology has been only applied by Hopkins [12] to analyze the explosion at the Esso Gas Plant in Longford, Australia. The scope of that study was different than the Aliso Canyon Accident; that was a gas plant while our study is related to a gas storage facility. Therefore, our study can be safely considered as the first systemic investigation that also uses the powerful AcciMap framework to analyze a major recent natural gas leak, the Aliso Canyon accident.

In this paper, the AcciMap methodology has been used to investigate and explain how certain managerial decisions, organizational processes and other contributing factors lead to an accident the scale of the one seen at Aliso Canyon. Studying this case using the AcciMap methodology will contribute to improving the industry's understanding on how human factors attributed to this accident.

Creating an AcciMap can help regulators, law makers and natural gas companies understand the interaction and interdependency of various socio-technical systems; it illustrates that not one independent factor or failure leads to the accident in its entirety, rather it was likely a compilation of various mistakes added under the burden of financial and high demand pressures within a competitive market.

3 The AcciMap Framework of the Aliso Canyon Gas Leak

In this section, the AcciMap framework has been utilized for the analysis of the Aliso Canyon gas leak, which occurred on October 23, 2015, in the SoCal Gas storage facility in Porter Ranch, CA. There have been some adjustments to the illustrated risk management framework in Fig. 1 to make the AcciMap specific to the context of our analyzed problem.

Our developed AcciMap in this paper, which has been shown in Fig. 2, consists of 6 main layers. In this AcciMap, the first five layers of the framework are: government and regulators; parent company (Sempra Energy); SoCal Gas Company; technical and operational management and crew; and physical events, processes and conditions. Finally, the sixth layer is the outcome, which is the Aliso Canyon gas leak. In Fig. 2, each level has been depicted by a separate color code in order to highlight the impact of

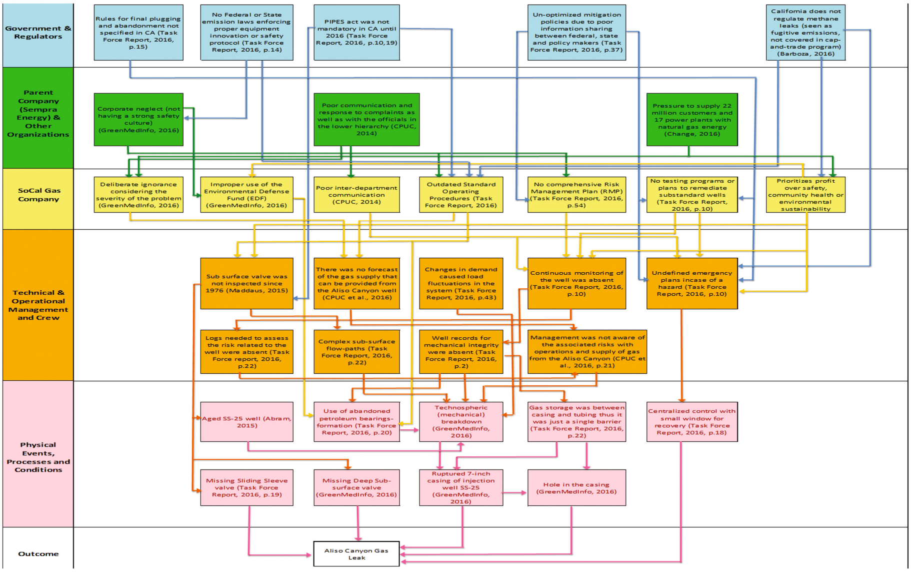


Fig. 2. The developed AcciMap framework for the analysis of the Aliso Canyon accident (http://www-bcf.usc.edu/~meshkati/AlisoCanyon_AcciMap.pdf)

the different layer components on each factor/box in the AcciMap. In addition, the main source of reference for each captured contributing factor has been cited in its box.

The developed AcciMap framework illustrates the contributing factors to the Aliso Canyon accident from each of the stated layers. It also depicts the interactions of different factors and involved decision-makers and key players in all these layers, which contributed to the accident. The following sub-sections provide more detail regarding the involved contributing factors in each of the stated layers of the AcciMap framework.

3.1 Government and Regulators

The first influential level of the AcciMap is government and regulators. A multitude of key governmental factors that attributed to this accident were found. First, nationally, there were not stringent enough laws enforced by the Pipeline and Hazardous Materials Safety Administration (PHMSA) prior to the accident. Additionally, the state of California does not regulate methane emissions because they are seen as fugitive emissions that are not regulated under the clean air act.

In June 2016, Congress passed the Protecting our Infrastructure of Pipelines and Enhancing Safety (PIPES) Act of 2016, signed into law by President Obama. The law implements standards for operation, environmental protection and integrity management; considers economic impacts of the regulations on customers; and ensures there is no significant economic impact on the end users. This Act was not mandatory in CA till 2016 [7, p. 10, 19].

Well SS-25's transition from being an old oil reserve to a storage unit meant that there were less stringent regulations in comparison to the ones enforced for newer facility. However, even up-to-date regulation standards at the time of the accident were more lenient than they are today.

3.2 Parent Company (Sempra Energy) and Other Organizations

SoCal Gas's parent company Sempra Energy did not have sufficient organizational sociology within the company and therefor this level experienced tradeoffs between safety and profit. The company has always tried to put forth an image of environmental sustainability and community outreach yet repetitive history and dangerous work procedures show otherwise. The culture of the company is blaming individuals and doing the bare minimum for safety measures. The utility company tries to do its best while still providing reliable energy to its customers. The pressure to supply 22 million customers and 17 power plants with natural gas energy pushed upper management to prioritize unsafe supply to meet increasing demands [6]. The company's neglect for safety culture was a starting point for what ultimately led to the Aliso Canyon accident [13].

3.3 The SoCal Gas Company

SoCal Gas's management decisions allude to lack of leadership, which in turn affects the dependent following levels, such as the staff that looks up to them for guidance. Management sets the safety culture and enforces proper protocols to be followed by the employees. Their responsibility is to be held accountable for their actions even if their employees are the ones in direct contact with the technology. The employees will act within the components of the safety control structure. SoCal Gas management made the decision to ignore possible technological gaps in their system when previous, smaller-scale accidents could have been indicating a larger issue [7, p. 62]. Within the organization, there were interdepartmental communication issues that did not allow for the proper flow of information.

Additionally, no comprehensive risk management plan was established prior to the accident, making mitigation difficult and prolonging the kill procedure timeline [7, p. 54]. There were also no testing programs or plans in place to remediate substandard wells [7, p. 2]. In 2008, the British Geological Survey (BGS) [14, p. 127] iterated how important having mitigation and remediating risk plans in place prior to accidents are to bring a system back under control as quickly and safely as possible. The United States did not have parallel legislations to allow SoCal Gas to have a timely mitigation plan at the time of the accident.

A "Noise and Temperature Survey" record from 1992 shows that SoCal Gas was aware of a possible leak in well SS-25 [13]. The Flow-Log survey record states "check for potential leakage past shoe as high as 8150 [feet]" under the 'reason for survey' section [13], exemplifying SoCal Gas' lack of urgency to improve their systems and disregard for safety prevention. Actions like these show how companies are quick to blame individual incidents for errors rather than taking the time to make a cultural change within the company. Although it may be a longer solution, taking the time to set a tone of equality, openness and creativity, in the end save the company money and lowers the possibility of risks because safety is not forgone. This accident could have been prevented if only there was a stronger emphasis on safety culture and a well-established preventative risk assessment system. The key traits of a positive safety culture include leadership, problem identification, personal accountability and continuous improvements. When these traits are compromised, lives, environmental sustainability, health and safety are being jeopardized for the sake of making a profit.

3.4 Technical and Operational Management and Crew

The staff might have not been aware of or in control of the fact that the technology they were using was not state of the art because the lack of questioning attitude emphasized in the company culture. With no baseline federal or state regulation to compare standards to and little leadership from within the company, there was little motivation among operators to work within the boundary of functionally acceptable performance. Management never fully comprehended or relayed the associated risks with operation and supply of gas from the Alison reservoir to their employees.

Considering the age of well SS-25, precautions should have been taken to bridge the gap between venerable infrastructure and safety measures. The original Downhole Safety Valves (DHSV), which are designed to shutoff flow to the surface when off-normal conditions are observed, were often replaced when the reservoir was being converted to a gas facility. During the 1970s, when Aliso Canyon was undergoing these changes, the DHSVs were often replaced with subsurface sliding sleeve valves, which are meant to be used during normal well operations. The Task Force's investigation shows that 60 of the 115 wells did not have any indication of DHSV being installed [7, p. 19]. Well SS-25 subsurface valve last recorded inspection was in 1979 [5]. The BGS [14, p. 128] suggested regular sonar logging runs in underground gas storage wells to assist in monitoring and detecting leaks, but Aliso Canyon only installed infrared thermal imaging cameras after the accident was contained [15].

During the accident, no continuous monitoring on the wells was put in place on the complex sub-surface flow paths used in this system [7, p. 22]. Further, changes in the load put on the wells fluctuated constantly due to changes in demand, and upper legal limits were often ignored to assure natural gas was meeting demand [7, p. 59]. The company also had lenient requirements for record keeping. Logs needed to assess possible risks associated to the wells and records for mechanical integrity were both absent [7, p. 21, 22].

3.5 Physical Events, Processes and Conditions

This level in the AcciMap works as a final layer of defense against accidents. The flow of events depends on all the preceding interdisciplinary sectors. The work of pumping the gas can be maintained when there is not a burden to pump the gas at a rate faster than the infrastructure allows, however greed and time pressures work against this safety precaution. The energy industry is highly interdependent with tightly linked operations, which magnify each failure through every segment of the company. The concatenation of having an aged well using an abandoned petroleum bearing formation allowed for the mechanical breakdown of the system to go unnoticed until it was too late to be stopped [7, p. 20]. The missing sliding sleeve valve and deep subsurface valve made the crack in the 7-inch tubing possible, and the small window for recovery did nothing to help stop this accident from playing out [13]. These accumulated factors have a compounding effect that in the end, cost the utility companies more than any safe precautionary investments would have been in the first place, and have additional external cost of community detriment and environmental degradation.

4 Conclusion

4.1 Model Analysis

There are different sets of learning points from the analysis of the developed AcciMap model described in the previous section for the investigation of the Aliso Canyon gas leak. A very important characteristic of the AcciMap approach is placing the events and conditions that finally released the accident into the necessary context to enable the understanding that how and why the accident occurred. This characteristic avoids the unfair blame of the frontline operators, since it provides a big-picture perspective and background on where those events and conditions come from and what the sources of operators' decisions are.

This concept in the context of analyzing the Aliso Canyon accident is equivalent to not only considering the immediate physical events and conditions or the decisions and actions made by the crew and technical management on the day of the accident as the contributing causes of that tragedy, but also to investigating the role and the contribution of factors in higher levels of the company or the external elements. In another word, AcciMap enables analysts to identify high-level contributing factors relating to organizational, governmental and regulatory practices as well as direct contributing causes of the outcome, by investigating all these stated factors within the scope of the illustrated levels in Fig. 2. For instance, referring to Fig. 2, one of the direct contributing factors to the Aliso Canyon gas leak was a hole in the casing (Refer to the Physical Events, Processes and Conditions level). Using the AcciMap, we can trace all the other causes that led up to the creation of a hole in the casing.

Following the AcciMap, the existence of no Federal or State emission laws enforcing proper equipment innovation or safety protocol (refer to the first level) contributed in some ways to Sempra Energy's (as the SoCal Gas's parent company) corporate neglect and its lack of a strong safety culture (refer to the second level). This factor resulted in having no established risk management plans prior to the accident (refer to the third level), which made mitigation difficult and prolonged the kill procedure timeline. The existence of no specific risk management plan contributed to not continuously monitoring the well operations (refer to the fourth level), which resulted in the absence of well records for mechanical integrity (another factor in the fourth level). This factor in conjunction with complex sub-surface flow paths as well as load fluctuations in the system, which were often beyond legal upper limits, caused by changes in demand (two other factors in the fourth level) led to technospheric (mechanical) breakdown in the system (refer to the fifth level), including a ruptured 7-inch casing of injection well SS-25 (another factor in the fifth level), which itself contributed to the above-mentioned hole in the casing. Through examination of the relationships between each factor as seen in this example, the AcciMap becomes a powerful tool for accident analysis and tracking the contributing factors of the accident in different analyzed level.

Another important advantage of the AcciMap is highlighting the interactions, communication and interoperability within and between the captured levels of the framework, which each of them represents a group of decision-makers or players in the context of a studied problem. This way, it is possible to analyze and identify ineffective communication and interoperations in each level and between levels, which according to many references, they themselves are root causes of several accidents.

An accident of this scale was highly likely at a site like this considering the limits the company was pushing this well to with high gas demands and outdated technology. This alludes to the fact that human factors and insufficient safety culture are the leading contributing factors to accidents like these, a trend cross cutting through many other industries as well. This accident could have been avoided and we hope to illuminate a system that creates a better procedure for preventing and preparing for such disasters. This paper complements the findings and recommendations of the Interagency Task Force Report [7] on Natural Gas Storage Safety on how accidents in this industry can be prevented. The government report states that “while incidents at U.S. underground natural gas storage facilities are rare, the potential consequences of those incidents can be significant and require additional actions to ensure safe and reliable operation over the long term” [7, p. 3].

As stated before, there have been some taken improving actions by the government, regulatory bodies and the utility companies since the occurrence of the Aliso Canyon accident. We are able to state those actions in this paper due to space limitation. The next section provides some further recommendations suggested by the Task Force Report [7], which are aligned with the analysis of our developed AcciMap framework, as well as some additional recommendations provided by the authors.

4.2 Recommendations

Moving forward from this accident, the question is how can we prevent this from happening again? One main recommendation of this study is to instill human factors characteristics within the utility companies to move towards a culture that focuses on quality without compromising quantity.

The culture of the utility companies comes from the top and flows down, setting the tone for how workers will conduct their profession. Having a strong leader with safety values and actions is the best way to set an example for operators to conduct their work with integrity. By fixing the culture that the company is based on, the solution would be cross cutting into worker safety, community health and environmental sustainability.

The natural gas industry, federal and state regulation and local agencies must work together in a preventative safety culture to lower the chance of future leaks [7]. Gas storage facilities across the US ought to conduct risk assessments, develop and implement transition plans that address high-risk infrastructures and apply procedures to maintain safety and reliability throughout all facilities until the transition towards modern well design standards are recognized [7, p. 60]. The Task Force Report [7, p. 1] focuses on three areas of study: well integrity, public health and environmental effects, and energy reliability concerns in the case of future leaks. Regarding well integrity, the key recommendations include emphasizing new well designs that prevent single point of failure accidents that cause leaks and uncontrolled flow, and wells that do not assure this satisfaction and do not follow this design should be phased out of service. Well integrity should also assure that operators follow risk management protocols that include monitoring programs, well integrity evaluation, leakage surveys, mechanical integrity tests and conservative assessment intervals.

Reducing public health and environmental effects from natural gas storage leaks is an integral part on how recommendations should be enacted to assure the best outcome for the public. The Task Force Report [7, p. 2] recommends some key steps to prevent and mitigate the impact of future leaks. First, if leaks, such as the one at the Aliso Canyon, require a response from multiple jurisdictions, a unified command should be defined early. Leaders from each agency should coalesce to provide a clear communication channel between agencies, and with the company and the public. This will help move all stakeholders towards a communal goal of controlling the leak and addressing possible health risks. In addition, state and local agencies should establish emergency air monitoring protocols.

The final key area of focus is energy reliability. The US' 400 facilities has the capacity to store four trillion cubic feet of natural gas, which would run one-third of the nation's electricity [16]. When large accidents such as the one at the Aliso Canyon occur, millions of people are indirectly affected by the natural gas emission. The probability of electricity shortages in Southern California was increased long after the leak began and ended. The natural gas storage industry is responsible for providing energy to households and businesses year-round, and this is increasingly important during high demand times such as the winter. The Task Force Report [7, p. 75] recommends strengthening planned and coordinated efforts to lower possible future impacts of prolonged disruptions in infrastructure. It also suggests including optional back-up strategies to reduce reliability risk in the case of supply failures [7, p. 76].

These key factors interrelate and affect the way that the Task Force Report [7] suggests certain recommendations as well as how companies should react to such regulations. The findings of the Task Force Report [7] build on the recommendations suggested in the 2015 Quadrennial Energy Review [17]. The overarching goal behind the findings of all agencies emphasizes the urgency to replace, expand and modernize natural gas transmission, storage and distribution infrastructure [7]. The report does not suggest lowering the use of natural gas as an energy source, as these facilities are a large component of providing the US with electricity. Additionally, companies need to be held accountable for ensuring the safety of workers. Prevention and safety culture are crucial to the company and the public, but operators at the front line of defense during an accident must be aware of proper conduct that assure no harm is caused.

These recommendations give a comprehensive protocol on how technology should be updated, environmental and health effects should be minimized, and energy reliability should be ensured. However, by using the AcciMap methodology the recommendations go beyond the operator working at the time of the accident, or the steps that can be taken after an accident. The AcciMap allows us to see how every level in the methodology impacts and concatenates on creating the environment in which an accident can happen or be prevented. The root cause is not a single point of failure on the operator's end. It is the executives creating the safety culture in which risk management protocol is defined and implemented. Management needs to provide the training, improve the monitoring and data collecting technology and give the tools to the operators to succeed in this safety-critical industry.

To go beyond the recommendations of the Task Force, steps need to be put in place to ensure there are continuous improvements and updates that coincide with technological advances. New protocols and operator training should be updated when new federal standards are issued to assure compliance with legislation and improvement in safety culture. As updates are being implemented, proper transition plans and guidelines must be outlined to assure no discrepancy between types of systems and no inconsistencies within the company.

Another characteristic is that these recommendations are being adapted in a way that are properly addressing the issues that have led to accidents in the past. By using Rasmussen's AcciMap methodology, the analysis of an accident such as the Aliso Canyon can better equip the natural gas industry in how to be compliant in the future. These systems established with the help of utilizing frameworks such as the AcciMap should be robust enough to address new scenarios that could succumb to possible vulnerabilities we have never seen before. Systems should be put in place that have the flexibility to address engineering factors, physical disruptions such as weather, technological complications and human errors to assure it can stand up to unforeseen possible failures. There are some procedures that are only capable of handling planned changes, called Management of Change (MOC) procedures. Companies that do not have proper safety protocols often adopt MOC procedures. This is while these procedures are not capable of handling unplanned changes such as changes in human behavior over time or changes in the environment, which increases the risk of accidents [18]. Rasmussen believed that high-risk or unsafe systems were often caused by factors such as economic or political pressure [4]. To modernize these systems, there must be organizational changes with newly directed goals that adapt to changing markets, competitive pressures and governmental oversight and newly introduce legislations [18].

Finally, there should be a systematic framework on how companies can mitigate the environmental effects they have caused that align with a baseline national regulations. This will assure that there is a price to pay beyond lowering energy reliability and endangering the community. The framework will assure the environmental mitigation plan is comparable to the damage from the accident, as well as the financial burden being placed within the company and not on their customers.

References

1. US Energy Information Administration, April 2016. General format. https://www.eia.gov/tools/faqs/faq.cfm?id=427&t=3. Accessed 12 Dec 2016
2. Rich, N.: The Invisible Catastrophe. The New York Times Magazine, March 2016. https://www.nytimes.com/2016/04/03/magazine/the-invisible-catastrophe.html?_r=1. Accessed 2 Mar 2017
3. Columbia Accident Investigation Board: Report Volume I. National Aeronautics & Space Administration and the Government Printing Office, Washington, D.C., August 2003
4. Rasmussen, J.: Risk management in a dynamic society: a modeling problem. Saf. Sci. **27**(2), 183–213 (1997)
5. Barragan, B.: Disastrously Leaking Gas Well in Porter Ranch Hasn't Had a Safety Valve Since 1979, January 2016. http://la.curbed.com/2016/1/4/10850008/porter-ranch-leaking-gas-well-safety-valve-lawsuitss. Accessed 25 Dec 2016
6. CPUC et al.: Aliso Canyon Risk Assessment Technical Report. California Public Utilities Commission (CPUC), CA Energy Commission, CA Independent System Operator, LA Department of Water and Power, Southern California Gas Company, April, 2016. http://www.energy.ca.gov/2016_energypolicy/documents/2016-04-08_joint_agency_workshop/Aliso_Canyon_Risk_Assessment_Technical_Report.pdf. Accessed 28 Dec 2016
7. Task Force Report: Ensuring Safe and Reliable Underground Natural Gas Storage: Final Report of the Interagency Task Force on Natural Gas Storage Safety. General format, October 2016. https://energy.gov/sites/prod/files/2016/10/f33/Ensuring%20Safe%20and%20Reliable%20Underground%20Natural%20Gas%20Storage%20-%20Final%20Report.pdf. Accessed 20 Nov 2016
8. Leveson, N.: A new accident model for engineering safer systems. Saf. Sci. **42**(4), 237–270 (2004)
9. Reason, J.: Managing the Risks of Organizational Accidents. Ashgate Publishing Limited, Hampshire (1997)
10. Branford, K.: Seeing the big picture of mishaps: applying the AcciMap approach to analyze system accidents. Aviat. Psychol. Appl. Hum. Factors **1**(1), 31–33 (2011)
11. Rasmussen, J., Svedung, I.: Proactive Risk Management in a Dynamic Society, 1st edn. Raddningsverket, Risk and Environmental Department, Swedish Rescue Services Agency, Karlstad (2000)
12. Hopkins, A.: Lessons from Longford: the ESSO Gas Plant Explosion. CCH, Sydney (2000)
13. GreenMedInfo Research Group: A Special Report on the SoCalGas Natural Gas Leak at Aliso Canyon, January 2016. http://www.greenmedinfo.com/blog/special-report-socalgas-natural-gas-leak-aliso-canyon. Accessed 28 Dec 2016
14. British Geological Survey (BGS): An appraisal of underground gas storage technologies and incidents, for the development of risk assessment methodology. Health and Safety Executive 2008 (2008)
15. SoCal Gas Company: Aliso Canyon Storage Facility Fitness-for-Service Overview. SoCal Gas Community Awareness (2017). https://www.socalgas.com/1443740344669/Z_N17G0004A_Aliso-Cyn-Fitness-for-Service_Package_Final.pdf. Accessed 2 Mar 2017

16. The Pipeline and Hazardous Materials Safety Administration (PHMSA): Pipeline Safety: Safety of Underground Natural Gas Storage Facilities. General format, December 2016. https://www.federalregister.gov/documents/2016/12/19/2016-30045/pipeline-safety-safety-of-underground-natural-gas-storage-facilities. Accessed 28 Dec 2016
17. Quadrennial Energy Review:. Energy Transmission, Storage, and Distribution Infrastructure. The White House, April 2015
18. Leveson, N.: Rasmussen’s legacy: a paradigm change in engineering for safety. Appl. Ergon. **59**(Part B), 581–591 (2017)

Permit-to-Work Systems as a Health and Safety Risk Control Strategy in Mining: A Prospective Study in Resilience Engineering

Manikam Pillay[1(✉)] and Michael Tuck[2]

[1] School of Health Sciences, The University of Newcastle, Callaghan, Australia
Manikam.Pillay@newcastle.edu.au
[2] School of Engineering and Information Technology, Federation University, Ballarat, Australia

Abstract. Mining is an important contributor to the social and economic fabric of our society. However, it is also considered to be one of the most dangerous industries. Compared to manufacturing, mining is generally regarded as a more complex industry to work in, creating additional challenges for policy makers, researchers and practitioners. This paper first discusses the state of mining health and safety in Australia, followed by an examination of some of the complexities that characterizes the industry. Next one contemporary approach, permit-to-work systems (PTW), is introduced, followed by a review of the literature relating to its use as a health and safety risk control strategy. This is followed by a discussion of Resilience engineering (RE) as an innovation in health and safety management, and a case made for investigating RE as a safety management strategy using PTW systems. The paper concludes by suggesting a pragmatism research framework and two organizational theories upon which such research can be advanced.

Keywords: Mining safety · Permit-to-work · Organisational resilience · Resilience engineering · Gap between work-as-imagined and work-as-performed

1 State of Health and Safety in Australian Mining

Mining is a major contributor to national income, investments, exports and government revenues in Australia. The use of mined products by nations worldwide is extensive and includes electrical generation, production of cement, steel, agricultural lime, commercial and residential building products, asphalt, and medicines, as well as countless household, electronic, and other manufactured products. For this reason the industry has been a key driver of higher living standards in Australia over the last decade, contributing to 18% of the nominal gross domestic product [1] and providing employment to some 187000 workers [2]. However, mining has also been one of the most work hazardous work environments [3]. In Australia for example, the most recent statistics suggesting 112 workers have died in the industry in the last twelve years; a fatality rate of 3.99 fatalities per 100000 employees. This is more than twice that in comparison with other industries such as manufacturing [4]. Figure 1 illustrates trends

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_14

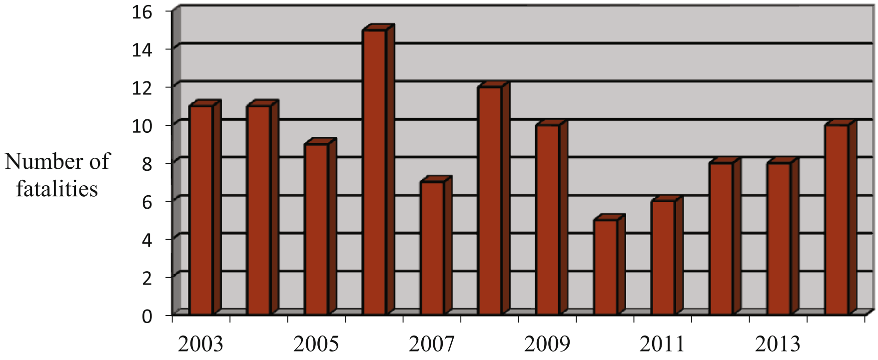


Fig. 1. Fatalities in the Australian mining industry 2003-2014 [4]

in fatalities experienced in the Australian mining industry from 2003 to 2014. While there were some improvements in performance until 2010, 10 workers died in the industry in 2014, which is the same as a decade ago.

The fact that safety performance of the industry has plateaued has previously been identified [5]. What makes it more disappointing is that in spite of a range of different safety initiatives, and a realisation that deaths in the industry are both foreseeable and preventable, workers continue to be killed and injured in the same, old ways [6]. This is also suggestive that some of the contemporary approaches used for improving safety have not been successful [7], and efforts to address this remain a key challenge for policy makers, researchers and practitioners. There is therefore, a need for improved and innovative strategies to move the industry off this plateau if the aspirations of zero harm are to be achieved [5].

1.1 The Complexities of Mining

There are number of things that occur make it different to a contemporary industry such as manufacturing. Perrow [8] was among one of the first to point this out, suggesting that mining operations could be regarded as complex but loosely coupled systems. A number of things can make the industry complex, including:

- deposit type, rock strength, depth, thickness, inclination, roof, floor strata, narrow veins, steeply inclined deposits, and deposits at great depths
- surface mining practices which include site preparation, overburden drilling and blasting, loading and hauling overburden, drilling and blasting deposits, loading and hauling ore, and reclamation of site [9]
- a high-percentage of contracting and sub-contracting worker arrangements [10]
- mix of highly skilled and semi-skilled operators operating both old and newer mining equipment, some of which can be sophisticated and beyond the cognitive reach of operators

What this means is that improving safety in mining can be more difficult compared to a manufacturing facility, and existing contemporary approaches may not be sufficient in driving safety improvements any further than what has been achieved. Many of the contemporary approaches are over fifty years old and, while they may have been adequate for the period they were developed, they are inadequate for present day organisational systems and operations. One such contemporary approach includes permit-to-work systems which are commonly used in the mining industry.

2 Permit-to-Work (PTW) Systems

PTW systems have a long history of association with safety management. The Health and Safety Executive defines a PTW as: "a formal written system used to control certain types of non-routine work, usually maintenance, that are identified as hazardous. The term 'permit-to-work' or 'permit' refer to the certificate or form that is used as part of the overall system of work. The permit is a written document that authorises certain people to carry out specific work at a specific time and which sets out the hazards associated with the work and the precautions to be taken" [11]. Thus PTW systems incorporate both a written documents and a series of rules describing and circumscribing safe methods of working.

However, the *purposes* which PTWs seek to achieve have been suggested to be more diverse and complex, and can include:

i. ensuring that hazards associated with any project and/or work being undertaken have been adequately considered,
ii. ensuring appropriate precautions and risk control measures have been put in place, and
iii. facilitating communication between the different stakeholders involved in the project and/or work [12].

Another view suggests that PTWs perform at least three different distinct functions, viz;

i. aid in the identification of hazards, together with concomitant precautions which need to be take,
ii. assist in coordinating and imposing the precautions, and
iii. provide a written record of what was done, by whom, and why [13].

PTW systems are also an essential component of safe systems of work, together with safety rules and work instructions [14].

2.1 Literature Review on PTW Systems

While PTW systems have been part of the safety management practice for a long time, there has been little published research on this topic. Four articles which we were able to locate from our search and selection strategy are reviewed below.

Booth and Butler [15] discuss a case study where procedures associated with PTW systems were radically overhauled and implemented at Shell U.K. Exploration and Production company operating in the offshore environment. The impetus for this, which took over 18 months, was based on the findings from the *Piper Alpha* Disaster. The most critical maintenance problem identified in *Piper Alpha* was a failure of the PTW system, resulting in Pressure Safety Valve number 504 being removed and replaced with by a blind flange without proper tagging [16]. The overhauled procedures included a greater emphasis on the control and coordination of dangerous operations, maintenance of effective plant and process isolations, defined authorities and responsibilities, extensive training and competence assurance, which were further reinforced by strict auditing and quality improvement processes. The actions discussed in this article are largely a reactive approach to safety management, and is common in most industries.

Scott [12] reported the findings a number of PTW surveys undertaken in 137 small and medium-sized chemical plants. This research revealed that a significant numbers of companies still had inadequate systems in many, with over 60% of the companies surveyed failing to audit their systems. Further interviews revealed that there was no or very little systematic approach to the permits, there was a lack of expertise in designing of permit forms with more than 50% of those surveyed copying their permits from elsewhere. In addition, use and precautions specified in permits were not monitored, hand-back procedures were not always effective, and training on PTW was inadequate.

Another survey, published in 1995, identified a number of areas where current PTWs are inadequate, the type and format of PTWs varied widely across the spectrum of plants investigated, most of these plants used at least 3 different forms to cover a variety of jobs, while a smaller group of plants used as many as 10 different forms [17]. At the lower end of the scale many companies ‘copied’ the permits of other companies without paying adequate regard to their appropriateness to their own needs; while at the upper end of the scale very specific permit forms led to a confusion of paperwork and loss of efficiency. The authors contend this finding was a symptom of a more general confusion in companies over precisely which jobs should be covered by a PTW and which should not. This survey also revealed that:

- in many companies there was disagreement between management, foremen and spotters over the general applicability of PTWs and the extent to which these had to be followed in every detail,
- fitters were unclear about the extent of their freedom to vary their work in light of developing knowledge of the maintenance situation, and the extent to which they were bound to follow the plan,
- it was frequently difficult to locate authorized issuers when a permit was actually needed, with the result that permits are commonly issued at a specific time in the morning with actual commencement of the work being left till some time later in the day or, in some cases, until several days later,
- similar confusions also existed over responsibilities for the sign-offs.

Another study conducted in 1996 identified three main weaknesses in PTWs. The first was that these were uninformative, in that they assumed that those who issued them were competent in identifying all hazards [13]. This is a fallacy, a myth of man being the fountain of all knowledge [18]. This is not necessarily the case, modern workplaces were more complex and the hazards and threats were not always visible. Moreover, while most skilled operators can identify normal, everyday hazards of the work being done, that is not necessarily the case with unexampled threats [19]. Second, they lacked clarity, in that most PTWs included a combination of lists, which the permit issuers checked and ticked off in against; these probably made sense to the issuer but not to anyone else [13]. This, in some ways, is to do with the myth that all workers are literate [18]. The third was that many of these were inflexible, which could suit those organisations and organisations which were compliance driven, but generally caused tension in those organisations where flexibility in business processes were necessary. The authors proposed that many of these weaknesses could be addressed by moving away from paper-based to computerised PTW systems [13].

The brief review suggests a number of gaps in research. First, all these studies were conducted in the chemical industry, operating both on-shore and offshore. There is an absence of research published research on PTW from mining. Second, two of these included evaluations [12, 17], while the other two looked at improvements to process aspects of PTWs [13, 15]. The third, and biggest gap, was the absence of any conceptual or theoretical model to guide the research process. Hence a theoretical framework which integrates the key ideas and concepts between PTW and safety does not exist.

2.2 Assumptions and Myths Associated with PTS

PTW systems are generally associated with safe systems of work, similar to safety rules and procedures [14]. In most cases these are prescribed by engineers and handed out to workers, with the assumption that they will follow these to the latter when doing work. However, people do not necessarily follow procedures all the time, or to the latter [20], and violations of procedures and rules are common in industry [21]. Moreover, some violations were sometimes necessary for achieving safe operations [22]. Routine violations are an example, where operators or teams regularly achieve objectives by means which differ from prescribed procedures [23]. In this case, the violations are often so deeply embedded into the daily practice that they are no longer identified as being con-compliant, and eventuate following the path of least effort, *laissez-faire* management style and poorly-designed procedures [24]. Another example of this occurs with exceptional violations when operators or teams performing an action in a context identified as a non-routine one, are required to make adjustments which results in a departure from the prescribed practice [23]. The aim here is generally to solve a problem that is not identified by the initial procedures [24]. These are also likely to occur if the rules or procedures were incomplete. Schein [25] contended that workers learnt that no matter how clearly the rules were specified, the world was (to some degree) unpredictable, and this led them to learn and adapt the rules to suit context. Subsequently these adaptations become part of the organisation's normal ways of working. According to Hollnagel [26] the planning and allocation of resources to get work done generally assumed that:

a. inputs to the work and resources required were regular and predictable,
b. the demands and resources were within limits,
c. working conditions always remained within limits, and
d. outputs complied with the expectations or norms.

In practice however, these conditions were not always fulfilled, but employees always adjusted their practices to adapt, improvise and get the work done [27]. This also required them to make a number of trade-offs to achieve not only safety, but other goals as well [28, 29]. Deviations from written procedures and violation of safety rules are part of these trade-offs. Moreover, because procedures and rules always required an interpretation between what was prescribed (assumed) and conducted (performed), gaps were likely to be created between the two.

What was important about this gap (*between work as imagined by management and work as actually performed by workers*) is that it is also a key indicator of resilience engineering (RE) [30–32], a recent innovation in safety management. So RE offers an opportunity to understand the links between PTW and safety.

3 Resilience Engineering

RE has been associated with safety management arena since 2003, although the idea of resilience has been part of the organisational and risk management literature as early as the 1990s [33]. The key concepts, ideas and principles associated with RE have been investigated in selected industries such as aviation, healthcare, nuclear and petrochemical; while a small number of studies have also been published from electricity distribution and railway domains [34]. Most of these have been suggested to be complex and tightly coupled organisational systems [8]. However, a number of authors have argued that there is scope for RE in traditional industries such as construction [32, 35, 36] and mining [7]. This, in essence, is an opportunity to explore the links between PTW, RE and safety in these industries. A central research question that one could ask is, “Do PTW systems enhance or hinder RE as a safety management strategy?”

Answering this research question empirically requires one to use an appropriate conceptual and theoretical framework to collect, analyze and interpret the data. The next section briefly discusses one such framework.

3.1 A Conceptual Framework for Investigating RE

An important facet of RE involves getting an understanding of the *gap between work as imagined and work as performed* [30]. RE itself is aimed at enhancing the ability of organisational systems to adapt or absorb disturbances, disruptions and change [34]. A prerequisite for this to occur without having a significant impact on operations is the need to create processes that are not only flexible, but also efficient and effective. This also relates to the prescription of work procedures and the rules that accompanied them (as perceived by the designers, engineers, planners and/or managers) and the undertaking of that work (by skilled operators and tradesman) [37]. Prescriptions, which also represented management intentions, acted as a guide to standardized ways of working.

These, according to the authors when communicated to different levels of the organisation, came as assigned responsibilities, specific objectives, norms, standard operating procedures or tasks. The authors also contend however, that workers did not generally accept prescription as a course of action but saw this as a constraint to their normal ways of work, causing them to derive their own interpretation of how the prescription was to be applied [37]. This transformation of prescription resulted from an interpretation, in mute dialectic with their accumulated experience, motivation, and particular circumstances of the moment and this guided their ultimate actions. As a consequence, instead of following procedures workers participated with co-workers and supervisors in making some sense of their work, interpreting situations and ultimately in attenuating variability through cooperative actions [38]. The essence of safe practice, from the point of view of those actually performing work, was not about having in-depth knowledge of, or compliance to, prescribed procedures; but the ability to perceive and distinguish between different types of situations and act accordingly when there was no clear path to follow [39]. The authors contend that this expertise developed through participation in collective actions involving continuous loops of repetitions-distinctions-descriptions. An expanded version of this model has been proposed for examining RE through safe work method statements in the construction industry [32]. This modified prescriptions-repetitions-distinctions-descriptions (M-PRDD) model is a useful and pragmatic model for examining whether PTW systems enhance or hinder RE as a safety management strategy.

3.2 Organizational Theories

Apart from a conceptual framework, academic research endeavours also require one to build on any existing theories in seeking to develop an understanding of whether PTW systems enhance or impede RE as a safety management strategy. We propose that an understanding of systems and social construction of safety are two organizational theories which provide a good starting point for such research.

Systems Theory. A system is a set of interrelated and interdependent parts arranged in a manner that functions as a whole to achieve a common purpose [40]. Two broad typologies of organisations are common, closed and open. Closed organisations are mechanistic, highly specialised with rigid departmentalisation, a narrow span of control, high formalisation, limited information network, and little participation in decision-making by employees [41]. Open organisations, one the other hand, were more organic, highly adaptive and flexible, little formalization of how work was done, and a greater reliance on innovation and creativity to achieve outcomes [41]. There is a growing recognition that organisations are complex adaptive systems which are fluid and flexible [42]. The very essence of the system lies in the interaction between the parts and the overall behaviour that emerges from these interactions [43]. It is these interactions that give rise to emergent behaviour that lead to operations being safe or unsafe. Getting a deep insight into these interactions is important in understanding how safety in mining organisations is achieved. Developments in systems research also suggests organisations are socio-technical systems (STS), with safety itself an emergent property of such systems [44]. The authors proposed a STS framework that could be useful for decomposing the dynamics of risk management in these types of systems, and this can be useful for advancing research of PTW, safety management and RE.

Social Construction of Safety. In discussing the systems view of organisations it was suggested that safety was an emergent property of a system. This idea of safety is closely linked with the notion that safety is a social construct [45], a process of discovery [33]. Gherardi and Nicolini [46] are also active proponents of this school of thinking, suggesting that safety arose from a constellation of interconnected practices, was socially constructed, innovated and transmitted to new members [46]. Proponents of RE such as Woods and Cook argued that safety was created as workers interacted with the hazardous processes inherent in the field of activity, amidst multiple demands, limited tools and resources [38].

4 Conclusion

This paper has argued that, while mining is an important contributor to the Australian society and, it has a poor safety record. Previous attempts to improve this performance using contemporary approaches resulted in some improvements, but this is far from acceptable. PTW systems are one such approach, but the paucity of empirical research and the lack of a conceptual and theoretical model in the limited studies published on this strategy mean there the utility or otherwise of PTW systems as a safety management strategy is unknown. This paper also presented some assumptions and myths surrounding the PTWs, as a safety procedure and/or rule, followed by an introduction to RE as an innovation in safety management. A central question of “Do PTW systems enhance or hinder RE as a safety management strategy?” was then asked. A conceptual framework for investigating the links between PTW, RE and safety has also been proposed, based on previous published research using the notion of the gap between work-as-imagined and work-as-performed. It has also been argued that socio-technical systems and social construction of safety are two organizational theories on which these research questions can be investigated. The next steps will include developing an appropriate range of examination, exploration and measurement instruments and tools for conducting pilot studies and full-scale investigations through a series of case studies in mining organisations.

References

1. Minerals Council of Australia. http://www.minerals.org.au/corporate/about_the_minerals_industry
2. Australian Bureau of Statistics: Mining operations, Australia 2013-2014. Canberra (2015)
3. Groves, W.A., Kecojevic, V.J., Komljenovic, D.: Analysis of fatalities and injuries involving mining equipment. J. Saf. Res. **38**, 461–470 (2007)
4. Safe Work Australia: Work-related traumatic injury fatalities, Australia 2014. Canberra (2015)

5. Galvin, J.M.: Occupational health and safety acts – performance and prosecution in the Australian minerals industry. Min. Technol. **114**, 251–256 (2005)
6. Ash, G., Betts, N.: Safety in tunneling and underground construction-an Australian perspective. In: (Re) claiming the Underground Space: Proceedings of the ITA World Tunneling Congress 2003, pp. 25–35. CRC Press (2003)
7. Pillay, M., Borys, D., Else, D., Tuck, M.: Safety culture and resilience engineering: theory and application in improving gold mining safety. In: Gravity Gold 2010, pp. 129–140. AusIMM (2010)
8. Perrow, C.: Normal Accidents: Living with High Risk Technology. Princeton University Press, Princeton (1999)
9. National Research Council: Evolutionary and revolutionary technologies for mining. The National Academies Press (2002)
10. Kenny, A., Bezuidenhout, A.: Contracting, complexity and control: an overview of the changing nature of subcontracting in the South African mining industry. J. S. Afr. Inst. Min. Metall. **99**, 185–191 (1999)
11. Health and Safety Executive: Setting up and running a permit-to-work system: How to do it. HMSO, London (1996)
12. Scott, S.: Management of safety-permit-to-work systems. In: Symposium of Major Hazards Onshore and Offshore, pp. 171–182. Institution of Chemical Engineers (1992)
13. Iliffe, R.E., Chung, P.W.H., Kletz, T.A.: More effective permit-to-work systems. Process Saf. Environ. Prot. **77**, 69–76 (1999)
14. Boyle, T.: Health and Safety: Risk Management. IOSH Services Limited, Leicestershire (2012)
15. Booth, M., Butler, J.D.: A new approach to permit to work systems offshore. Saf. Sci. **15**, 309–326 (1992)
16. Pate-Cornell, M.E.: Learning from the Piper Alpha accident: a postmodern analysis of technical and organizational factors. Risk Anal. **13**, 215–232 (1993)
17. Riley, D.B., Weyman, A.K.: A survey of permit-to-work systems and forms in small to medium sized chemical plant. In: ICHEME Symposium Series No. 139, pp. 367–380 (1995)
18. Pillay, M., Borys, D.: Towards an exploration of safe work method statements as a safety management strategy in construction: background and theory. In: Seventh International Conference on Construction in the 21st Century (CITC-VII) (2013)
19. Epstein, S.: Unexampled events, resilience, and PRA. In: Hollnagel, E., Nemeth, C., Dekker, S.W.A. (eds.) Resilience Engineering Perspective. Remaining Sensitive to the Possibility of Failure, vol. 1, pp. 49–62. Ashgate, Aldershot (2008)
20. Dekker, S.: Failure to adapt or adaptations that fail: contrasting models on procedures and safety. Appl. Ergon. **34**, 233–238 (2003)
21. Reason, J.: In Search of Safety. The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries. Ashgate, Surrey (2008)
22. Reason, J., Parker, D., Lawton, R.: Organizational controls and safety: the varieties of rule-related behaviour. J. Occup. Organ. Psychol. **71**, 289–304 (1998)
23. Reason, J.T.: Human Error. Cambridge University Press, Cambridge (1990)
24. Besnard, D., Greathead, D.: A cognitive approach to safe violations. Cogn. Technol. Work **5**, 272–282 (2003)
25. Schein, E.H.: Three cultures of management: the key to organizational learning. Sloan Manag. Rev. **38**, 9–20 (1996)
26. Hollnagel, E.: Barriers and Accident Prevention. Ashgate, Hampshire (2004)
27. Grøtan, T.O., Størseth, F., Rø, M.H., Skjerve, A.B.: Resilience, adaptation and improvisation: increasing resilience by organising for successful improvisation. In: Third Symposium on Resilience Engineering. MinesParis Tech. (2008)

28. Hollnagel, E.: The ETTO Principle: Efficiency-Thoroughness Trade-Off-Why Things That Go Right Sometimes Go Wrong. Ashgate Publishing Company, Burlington (2009)
29. Cooke, R.I., Render, M.L., Woods, D.D.: Gaps in the continuity of care and progress on patient safety. Br. Med. J. **320**, 791–794 (2000)
30. Dekker, S.W.A.: Resilience engineering: chronicling the emergence of confused consensus. In: Hollnagel, E., Woods, D.D., Leveson, N. (eds.) Resilience Engineering Concepts and Precepts, pp. 77–92. Ashgate Publishing Ltd., Aldershot (2006)
31. Hollnagel, E., Woods, D.D.: Epilogue: resilience engineering precepts. In: Hollnagel, E., Woods, D.D., Leveson, N. (eds.) Resilience Engineering Concepts and Precepts, pp. 347–358. Ashgate Publishing, Aldershot (2006)
32. Pillay, M.: Construction safety management, safety rules, and gaps between work-as-imagined and work-as-performed: an integrative review. Int. J. Risk Assess. Manage. (under review)
33. Wildavsky, A.: Searching for Safety. Transaction Books, New Brunswick (1988)
34. Pillay, M.: Resilience engineering: a state-of-the-art survey of an emerging paradigm for organisational health and safety management. In: Azeres, P. (ed.) Advances in Safety Management and Human Factors, vol. 491, pp. 211–222. Springer, Switzerland (2016)
35. Saurin, T.A., Sanches, R.C.: Lean construction and resilience engineering: complementary perspectives on variability. In: 22nd Annual Conference of the International Group for Lean Construction, IGLC, pp. 61–71 (2014)
36. Schafer, D., Abdelhamid, T.S., Mitropoulos, P.T., Howell, G.A.: Resilience engineering: a new paradigm for safety in lean construction systems. In: 16th Annual Conference of the International Group for Lean Construction, IGLC (2009)
37. Nathanael, D., Marmaras, N.: Work practices and prescription: a key issue for organizational resilience. In: Hollnagel, E., Nemeth, C., Dekker, S. (eds.) Resilience Engineering Perspectives. Remaining Sensitive to the Possibility of Failure, vol. 1, pp. 101–118. Ashgate, Aldershot (2008)
38. Woods, D.D., Cook, R.I.: Nine steps to move forward from error. Cogn. Technol. Work **4**, 137–144 (2002)
39. Nathanael, D., Marmaras, N.: On the development of work practices: a constructivist model. Theor. Issues Ergon. Sci. **9**, 359–382 (2008)
40. Friedman, B.D., Allen, K.N.: Systems theory. In: Brandell, J.R. (ed.) Theory and Practice in Clinical Social Work, pp. 3–20. The Free Press, New York (2011)
41. Robbins, S.P., Bergman, R., Stagg, I., Coulter, M.: Foundations of Management. Prentice Hall, Frenchs Forest (2006)
42. Schneider, M., Somers, M.: Organizations as complex adaptive systems: implications for complexity theory for leadership research. Leadesrhip Q. **17**, 351–365 (2006)
43. Ottino, J.: Complex systems. AIChE J. **49**, 292–299 (2003)
44. Rasmussen, J., Svedung, I.: Proactive Risk Management in a Dynamic Society. Swedish Rescue Services Agency, Karlstad, Sweden (2000)
45. Rochlin, G.I.: Safety as a social construct. Ergonomics **42**, 1549–1560 (1999)
46. Gherardi, S., Nicolini, D.: Learning in a constellation of interconnected practices: canon or dissonance? J. Manag. Stud. **39**, 419–436 (2002)

Modeling Framework of Team Contexts for the Design of Laboratory Experiments and Team Training

Taro Kanno[1(✉)], Satoru Inoue[2], Daisuke Karikawa[3], and Dingding Chao[1]

[1] The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
`{kanno,chao}@sys.t.u-tokyo.ac.jp`
[2] Electronic Navigation Research Insititute, 7-42-23 Jindaiji-Higashi-Machi, Chofu, Tokyo, Japan
`s-inoue@mpat.go.jp`
[3] Tohoku University, 6-6-11 Aramaki Aza Aoba, Aoba-ku, Sendai, Miyagi, Japan
`dasisuke.karikawa@tohoku.ac.jp`

Abstract. This paper presents a modeling framework of team structure that describes major elements of team settings and conditions and the relationships between them. These are the elements of human, task, resource, expertise, authority, tools and devices, and place. The relationships between these elements can capture and summarize important aspects of team structure such as the distribution and sharing of objects and functions to each team member and the physical environment. This paper provides details of the proposed modeling framework and discusses how to assess and quantify the similarity between a naturalistic team setting and a simulated or game-like setting.

Keywords: Team architecture · Mesocognition · Similarity assessment · Design of experiments and training

1 Introduction

Practical human factors studies emphasize the importance of macrocognition, namely practitioners' cognitive processes and behaviors in naturalistic rather than laboratory environments [1, 2]. In addition, resilience engineering stresses the need to develop an understanding of the "task as done" at the sharp end [3, 4]. It is recommended to conduct such a macrocognitive study to obtain suggestions for the practical improvement of the work environment and onsite conditions. However, there are many obstacles in conducting macrocognitive studies. For example, it is usually difficult for university researchers to find new collaborators and access to the work field. Even if we have or still find such collaborators through luck, many constraints and limitations are found in intervening with the field workers and recording data or in the experimental design that manipulates various parameters. In that case, we must give up pure macrocognitive studies without an alternative, then reluctantly implement non-macrocognitive approaches such as laboratory experiments using student participants and game-like tasks.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_15

The question here is whether "It is really possible to make good recommendations to improve the work field based on the data obtained from such studies?" Naturally, strong macrocognitivists will deny such approaches; however, it will be useful to design a non-naturalistic experimental setting that retains essential similarities from which can be derived useful recommendations or implications. It would also be useful if the degree of such a similarity or difference from actual work settings can be assessed objectively and quantitatively. This method and principle will provide guidelines for designing an appropriate experiment that considers actual work settings and experimental purposes, and facilitate accumulation and integration of the data and knowledge obtained from well-formulated experiments. This holds true for the design of training settings. If it is possible to manipulate or at least quantify the degree of similarity, not fidelity, of training settings to actual work settings, we can design and assess an entire training program systematically based on this. We refer to experimental and training settings that are non-naturalistic but with some similarity as "meso-cognitive" settings, and to the cognitive process and behavior entailed in these settings as "meso-cognition."

In this study, for the first step to theorize meso-cognition or meso-cognitive settings, we tried to develop a modeling framework to describe the characteristics of real team settings.

2 Related Studies

To the authors' knowledge, not much attention has focused on relevant issues around meso-cognition; however, numerous studies serve as useful references for building the modeling framework, which is introduced in this sections.

2.1 Team Categorization

Many studies provide taxonomy or classification schemes of teams. These aim to help better understand team effectiveness in different domains and contexts. For example, Devine [5] reviewed numerous classic taxonomies of teams and team tasks in the psychology and management literatures, and provided the seven major dimensions of team context listed in Table 1. He also provided an integrative taxonomy comprising 14 team types classified according to these attributes. Furthermore, Nonose [6] extended Devine's taxonomy by offering the nine additional dimensions shown in Table 2 to capture more characteristics of team context.

If these context dimensions are exclusive and exhaustive and if the value of each dimension can be quantified somehow, team context can be defined and represented as a multidimensional vector. Once team contexts are vectored, differences or similarities between two team contexts can be provided by the cosine value of the two vectors, for example. Similarly, different team contexts can be categorized using a vector clustering method. This approach is simple and rigid; however, it will not work with the dimensions available to date, as they are not exclusive and exhaustive. For example, several dimensions, such as communication mode, physical ability requirements, and hardware dependence, are heavily dependent on task contents and work conditions.

Table 1. Seven major dimensions of teams [5]

Contextual dimensions	Meaning
1. Fundamental work cycle	Smallest meaningful unit of collective activity for the team
2. Physical ability requirements	Types of human characteristics needed to be accomplish the team task
3. Temporal duration	Length of time for which the team exists
4. Task structure	Degree to which member actions relate to outcomes in an understandable and predictable way
5. Active resistance	Presence of human opposition directly trying to prevent the team from accomplishing its goal
6. Hardware dependence	Degree to which team activities are constrained by technological resources in the form of communication systems
7. Health risk	Likelihoodand severity of human injury associated with errors in team performance

Table 2. Additional dimensions [6]

Contextual dimensions	Meaning
1. Spatial distribution	Location and deployment of team members
2. Temporal distribution	Degree to which team tasks must be done in a synchronous or asynchronous manner
3. Dependency	Degree to which team members are dependent to each other
4. Mode of communication	Face-to-face or communication via tools and computers
5. Expertise	Degree of variety of expertise required
6. Stability	Degree to which team members consist of the same persons
7. Scale	Number of team members
8. Fluidity of task	Degree to which task process is fixed or fluid
9. Physical task	Degree to which tasks require physical actions

Rather task contents and work conditions are important factors that determines team contexts and behaviors. Therefore, it is necessary to focus on these two factors and their relationships with other dimensions and variables.

2.2 Task Categorization

There are also numerous studies on task taxonomy. Steiner [7] provided one of the earliest categorizations, which focuses on the relationship between individual member performance and group performance, namely disjunctive, conjunctive, additive, and discretionary tasks. McGrath [8] also conducted well known early work, in which he defined eight types of group tasks: (1) generating plans, (2) generating ideas, (3) solving problems with a correct answer, (4) decision making, (5) resolving conflicting viewpoints, (6) resolving conflicts of interest, (7) resolving conflicts of power, and (8) performances.

One of the limitations in these earlier taxonomies of team task taxonomies is that a team sometimes performs multiple tasks in different categories; thus, this taxonomy cannot be used to categorize or comprehensively describe a team or team context. Another problem is that these taxonomies focus only on the content of team tasks, but as mentioned in the previous subsection the relationships with other dimensions must also be considered when describing team contexts comprehensively.

2.3 Team Structure

The review and discussion thus far suggests that to build a framework for team context or settings for meso-cognitive studies, a good starting point is to distinguish independent dimensions and their resultant dimensions or variables, and to capture the relationship between them. This is a system modeling approach to describe the structure of teams. A widely known structural model of human system is the PCANS model proposed by Krackhardt and Carley [9], which describes organizational structure in terms of the relationships between humans, tasks, and resources. PCANS represents acronyms of the important relationships between these three elements, as listed in Table 3.

While the PCANS model was originally developed to describe organizational structure, it is also applicable in describing team structure, because the fundamental elements of these two human systems are mostly common. However, the spatial and temporal factors listed by Nonose have more impact on team context and team behaviors, and should be considered and incorporated.

Table 3. PCANS relationships (slightly modified from [9])

	Human	Task	Resource
Human	Network: This covers various social relationships among members	Assignment: A member has a responsibility to accomplish certain tasks	Skill: This represents member’s accessibility and capability to handle resources such as equipments and tools
Task	–	Precedence: This represents a temporal ordering of tasks, such as means-ends relationships and procedures	Commitment: An agent must commit certain resources to a task to accomplish it. This represents the relationship between a task and required resources
Resource	–	–	–

3 Building a Framework of Team Context

Based on the review and discussion in the previous section, we developed a design policy to build a framework of team context as follows.

1. Separate and distinguish structural aspects from content, in particular task content.
2. Separate and distinguish physical factors—spatial-temporal-substantial factors—from conceptual or behavioral factors.
3. Adopt an extended PCANS model to describe team structure

3.1 Modeling Framework of Team Structure

In this subsection, a prototype of the modeling framework of team structure extended from the PCANS model is proposed. This modeling framework consists of seven major elements and describes team structure in terms of these elements and the relationships between them. The overview of the modeling framework is shown in Table 4. Seven elements were selected to capture the structure of a team, namely, human, task, tools and devices, resource, expertise, authority, and place elements.

Table 4. Proposed modeling framework

	Human	Task	Tool device	Resource	Expertise	Authority	Place
Human	Social relations	Assignment	Assignment Accessibility	Accessibility	Competence Ability	Role Status	Work station Work post
Task		Process Procedure	Employment	Materials Input	Required abilities (for execution)	Qualification (for execution)	Work area
Tool device			Substitutability Dependency	Driving force Input	Required abilities (for use)	Qualification (for use)	Installation location
Resource				Substitutability Dependency	Required abilities (to handle)	Qualification (for handling)	Place of use or storage
Expertise					Substitutability Dependency	Requirement (to be authorized)	–
Authority							–
Place							Work environment

There are two main extensions from the PCANS model. One extension is to detail the resource element of the PCAN model as the four items of tool and devices, resource, expertise, and authority. In the proposed model, resource refers to physical and tangible materials. The intention behind this categorization is to distinguish between physical and conceptual factors: the former two are physical factors and the latter two are conceptual and non-tangible. The other extension is to introduce place to capture the spatial characteristics of team context. It is obvious that the spatial distribution of model elements such as team members and devices becomes a constraint on team behaviors and interactions. Thus, it is important to replicate or imitate it when building a meso-cognitive context. The spatial distribution of the model elements is represented in the last column of Table 4. In addition, this modeling framework can capture another important type of distribution, namely the distribution of model elements to team members. This distribution is represented in the second to sixth columns in the first row in Table 4. It is also obvious that this type of distribution becomes constraints on or reasons for team interactions and coordination.

3.2 Discussion

Currently, we have not yet constructed detailed models of each element. However, except for the task element, it seems relatively easy to describe the details. For example, it is easy to identify which type of tools and devices as well as resources are used in an actual team context because they are usually tangible and visible. In addition, once these elements are identified, it is easy to find relevant expertise and authorities for their use and handling. Task, and especially its content, is the most critical element in a team context; therefore, modeling task content is the next critical step. The temporal aspects not considered in the current framework will be captured in the model of task content, because almost all temporal factors can be associated with task content, such as duration, timing, and time constraints.

One of the advantages of focusing on team structure, elements, and the relationships between them is that once these elements and their relationships are identified, then the structure can be represented in the form of a network. Graph theory provides various quantitative characteristics of network structure such as density, size, average degree, an average path length, which enables discussion of the differences and similarities between two team structures objectively and quantitatively.

4 Conclusions

This paper presented a design policy for building a modeling framework of team context that aims to provide a theoretical base for designing team contexts similar to actual ones. This paper also provided a modeling framework of team structure comprising seven elements: human, task, tools and devices, resource, expertise, authority, and place. This framework can capture the spatial distribution of team context elements and the distribution of these to team members. The proposed design policy is expected to offer new insights into and directions to the modeling of team context.

Acknowledgments. This work was partly supported by JSPS KAKENHI Grant Number 16H03131.

References

1. Letsky, M.P., Warner, N.W., Fiore, S.M., Smith, C.A.P.: Macrocognition in Teams: Theories and Methodologies. Ashgate, Aldershot (2008)
2. Klein, G., Ross, K.G., Moon, B.M., Klein, D.E., Hoffman, R.R., Hollnagel, R.: Macrocognition. IEEE Intell. Syst. **18**(3), 81–85 (2003)
3. Hollnagel, E.: Safety-I & Safety-II. CRC Press, Boca Raton (2014)
4. Nemeth, P.C., Hollnagel, E.: Resilience Engineering in Practice, Vol. 2: Becoming Resilient. Ashgate, CRC Press, Aldershot (2014)
5. Devine, J.D.: A review and integration of classification systems relevant to teams in organization. Group Dyn. Theory Res. Pract. **6**(4), 291–310 (2002)
6. Nonose, K.: Study on team cognition based on meta-cognition, Doctoral thesis, School of Engineering. The University of Tokyo (2010)
7. Steiner, I.: Group Process and Productivity. Academic Press, New York (1972)
8. McGrath, J.E.: Groups: Interaction and Performance. Prentice Hall Inc., Upper Saddle River (1984)
9. Krackhardt, D., Carley, K.M.: A PCANS model of structure in organizations. In: Proceedings of 1998 International Symposium on Command and Control Research and Technology, pp. 113–119 (1998)

The Impact of Human Errors on the Estimation of Uncertainty of Measurements in Water Monitoring

Ewa Kmiecik(✉)

AGH – University of Science and Technology,
Mickiewicza 30 Avenue, 30-059 Kraków, Poland
`ewa.kmiecik@agh.edu.pl`

Abstract. The main purpose of a physical act of measurement is to enable decisions to be made. In case of an assessment of the chemical status of groundwater body, or assessment of suitability of water for drinking purposes, or possibility of discharges sewage into surface waters, the measurements of physicochemical parameters of water are an indispensable first step. The reliability of the mentioned above decisions heavily depends on knowing the uncertainty of the measurement results. If the uncertainty of measurements is underestimated, for example because the human errors are not taken into account, then erroneous decisions can be made that can have in some cases substantial financial consequences. In this work there are presented examples of human error identification and estimation in measurements made during water monitoring on the base of duplicate control samples (empirical approach) with the use of control charts method.

Keywords: Human errors · Water monitoring · Duplicate control samples

1 Introduction

The main purpose of a physical act of measurement is to enable decisions to be made. In case of water monitoring for the assessment of the chemical status of groundwater body, or assessment of suitability of water for drinking purposes, or possibility of discharges sewage into surface waters, the measurements of physicochemical parameters of water are an indispensable first step.

Errors occur throughout the whole process of water quality monitoring. It is estimated that about 30% of the errors are generated in the sampling process, 60% are mistakes made during processing and preparation of samples for analysis, and the errors of the analytical measurement alone do not generally exceed 10% [1].

Because of the manner of occurrence in the results, the errors are divided into: random, systematic and gross (dubious results, outstanding values). This classification of errors and their analysis is devoted to many publications, i.a. [2–4].

In turn, due to the source of the origination errors can be divided into: instrumental (deriving from the imperfections of measuring instruments and their individual components), methodical (the causes of which lie within the measurement itself, or even the

© Springer International Publishing AG 2018

R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_16

principle of measurement, and are not possible to eliminate) and personal (operating, human factor). Personal errors are caused by the imperfection of the personnel performing the measurements (sampling/analysts) and depend on individual personal characteristics.

Human activity is never free from errors, so human errors are not predictable and can't be taken into account in advance. However, they can be detected and eliminated, or at least significantly reduced, by training staff, decoupling methods from the human factor, and in extreme cases by replacing specific staff by another, e.g. with a higher level of education or better qualified.

Human error in chemical analysis is defined as any action or lack thereof that leads to exceeding the tolerances of the conditions required for the normative work of the chemical analytical (measuring/testing) system with which the human interacts [5].

There are two groups of human errors: errors of commission and errors of omission [6]. Errors of commission (mistakes and violations) are inappropriate actions resulting in something other than what was intended. Errors of omission (lapses and slips) are inactions contributing to a deviation from the intended path or outcome. Details on those classification and methods of human errors quantification can be found in the literature [5–16].

Accredited laboratories should be able to control human errors, classify and quantify them, and be able to develop preventive actions. Study of human errors is required i.a. by the US Food and Drug Administration, the UK Medicines and Healthcare products Regulatory Agency and by other regulators, as a part of quality risk assessment [12].

Human error should not be confused with measurement error. It may cause measurement error and influence measurement uncertainty. In this work there are presented examples of human error identification and estimation in measurements made during water monitoring on the base of duplicate control samples (empirical approach) with the use of control charts method. Presented examples are based on data collected during different projects realized by accredited hydrogeochemical laboratory at the AGH-UST, covering different water matrices (groundwater and mining waters), various samplers (analysts) with different level of experience.

2 Methodology

The detailed analysis of human errors was done for physicochemical parameters which are analyzed in situ (electrical conductivity EC, pH) or in the laboratory (chlorides), but with the use of titration method, in which the human factor plays an important role (human dependent method).

Table 1 summarizes the methodology of analysis. Relative expanded uncertainty is declared by the laboratory uncertainity, estimated with the uncertainty budget method, taking into account all the factors affecting the final results of analysis (random and systematic). It is usually higher than uncertainty estimated on the base of duplicate control samples. It also does not include additional estimation of human errors.

pH and *EC* are the two most important indicators determining the physicochemical conditions of the migration of substances in natural waters, including toxic substances.

Table 1. Field analytical methods characteristics.

Parameter	Analytical method	No. of standard	Limit of quantification	Relative expanded uncertainty [%] (k = 2, 95%)
pH	Potentiometric	PN-EN ISO 10523:2012	2 [-]	5.5
EC	Conductometric	PN–EN 27888:1999	2 µS/cm	9.0
Chlorides	Titration	PN-ISO 9297:1994	1 mg/L	4.9

pH belongs to these characteristics of water, which should be measured directly in the field during water sampling. Measurements made on the sample transported to the laboratory are unreliable and, as a result of transport and storage processes, often differ significantly from field testing [17].

pH measurement has been made for more than 100 years [18]. It is probably the most frequently performed measurement in analytical chemistry today. However, the quality of pH measurement results is still a subject for discussion and improvement. In any pH measurement, human error may influence reliability of the measurement results [12]. For example 34 scenarios of human error in pH measurement of groundwater were discussed and quantified [6].

Electrical conductivity is the most commonly measured general water parameter. This is due to the close correlation of EC with the amount of dissolved substances in water and the correlation with the content of a number of major constituents and characteristics of water (chlorides, sulfates, general hardness). Virtually any change in the amount of dissolved substances in water causes a change in EC [17].

Chlorides are the most common form of chlorine in nature, and they are among the major constituents of natural waters. They are a very important indicator of pollution of fresh groundwater. This is due to the presence of elevated chloride concentrations in most urban sewage, natural fertilizers, slurry, leachate from municipal landfills, and many industrial wastes.

In practice, the most commonly used chlorine determination methods are: titrimetric method (Mohr) and ion chromatography. In the titrimetric method, human factor plays important role.

There are several types of errors influencing the result of the titrimetric analysis. This method is human sensitive and human errors may concern i.a. misjudging the color of the indicator near the end point (different sensitivity to colors) or misreading the volume (a parallax error). The detailed discussion about errors generated in titration method can be found in the literature [19].

The reliability of decisions made on the base of physicochemical analyses results heavily depends on knowing the uncertainty of the measurement results. If the uncertainty of measurements is underestimated, for example because the human errors are not taken into account, then erroneous decisions can be made that can have in some cases substantial financial consequences. Therefore, it is essential that effective procedures are available for estimating the uncertainties arising from all parts of the measurement process, including human factor.

There are two broad approaches to the estimation of uncertainty. One, empirical ('top down'), uses replication of the whole measurement procedure as a way of direct estimation of the uncertainty of the final result of the measurement. The second, modelling approach ('bottom up'), aims to quantify all the sources of uncertainty individually, and then uses a model to combine them into overall uncertainty characterizing the final result. An empirical approach makes it easy to estimate the uncertainty of measurement using control samples – duplicate control samples. Details of how they are collected and processed can be found in the literature, i.e. [20–27].

To develop the data obtained from the duplicate samples, individual measurements control charts (charts of differences) are used. Examples of application of this technique for developing groundwater monitoring data are described in [28–30] and in the 7870 series of standards.

In this work control charts were done using PS Imago solution (based on IBM SPSS Statistics – www.psimago.pl). Measurement uncertainty was estimated with the use of U-Ranova software [31].

3 Examples of Human Errors Identification

3.1 Example 1: Raba River Basin Monitoring

During groundwater monitoring of the Raba River Basin project 8 pairs of duplicate control samples were taken on the same day [32], by one accredited sampler. pH and EC were measured in situ by one sampler, chlorides were determined at the laboratory by one analyst.

Estimated on the base of duplicate samples results of measurement uncertainty for pH and EC are lowered than declared by the laboratory ($U'_{pH} = 0.75\%$, $U'_{EC} = 0.6\%$, $k = 2$, 95%) except chlorides for which measurement uncertainty is higher than declared ($U'_{Cl} = 15\%$).

The differences between results for normal and duplicate samples were plotted on individual measurement control chart (Fig. 1).

Each point plotted on the chart represents an individual process measurement, difference of results. The central line (CL) is the average of the data. The two dashed-lines are the upper and lower control limits (UCL and LCL, ±3 standard deviations (3 s) of the mean). When an analytical process is within control, approximately 99.7% of all values are found to be within ±3 standard deviations (3 s) of the mean. As only 0.3%, or 3 out of 1000 points, will fall outside the ±3 s limits, any value outside of ±3 s is considered to be associated with a significant error condition.

No point or sequential signals were identified on the charts for pH and EC (Fig. 1) and therefore differences between these results can be regarded as stable. Chart for chlorides shows no point signals, but decreasing tendency, what could be the result of an analyst error. Section 3.3 demonstrates the results of competence test of analysts performing titrations, which are systematically carried out in the laboratory, including all analysts.

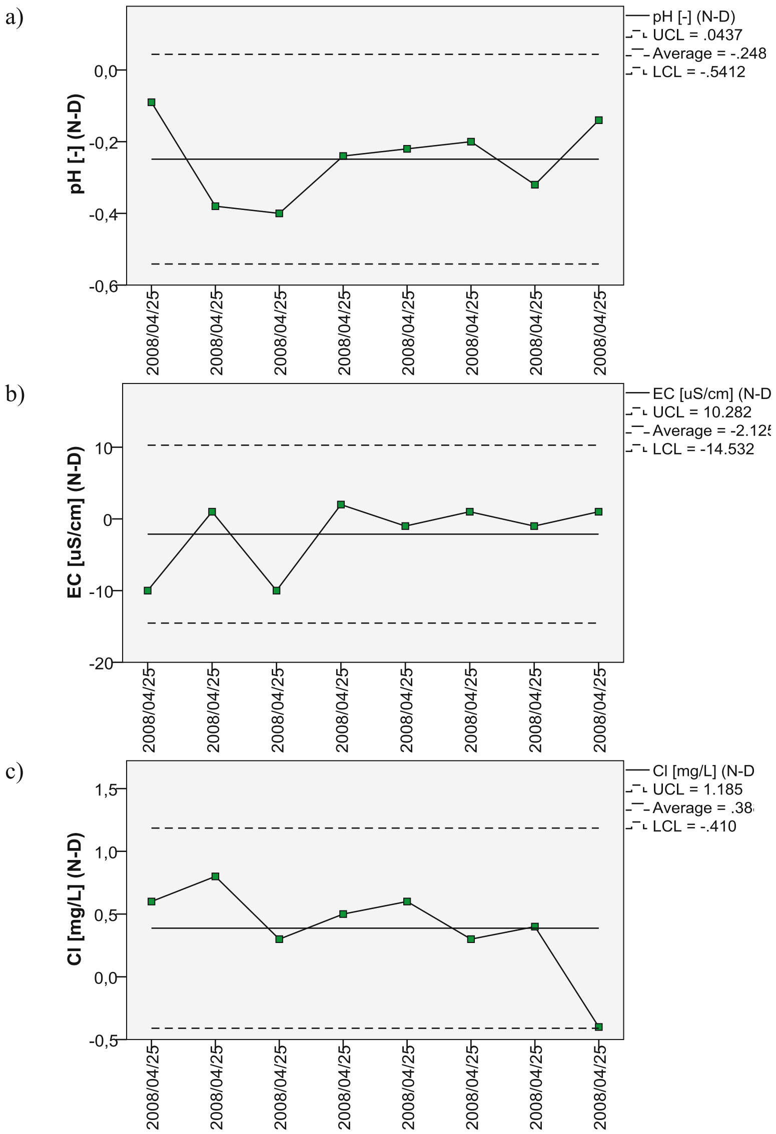


Fig. 1. Individual measurement control charts for results differences in normal and duplicate samples: (a) pH, (b) EC, (c) chlorides concentration.

3.2 Example 2: Mining Water Monitoring

In mining water monitoring project [33, 34] 11 pairs of duplicate control samples were taken on the same day, by one accredited sampler. Chlorides were determined in two laboratories (A and B) with different methods: titration (A – 1, B – 2) and ion chromatography (B – 3). Measurement uncertainty U'_{Cl} estimated on the base of duplicate control samples results was, respectively: 6.51% (1), 1.23% (2), 4.11% (3) (k = 2, 95%). Estimation for method (2) seems unrealistic. During the discussion it was explained that the QA/QC program was done in this case separately from normal analyses. The analyst knew that he was doing analyses of the control samples and tried to achieve the best results (most compatible for subsequent samples).

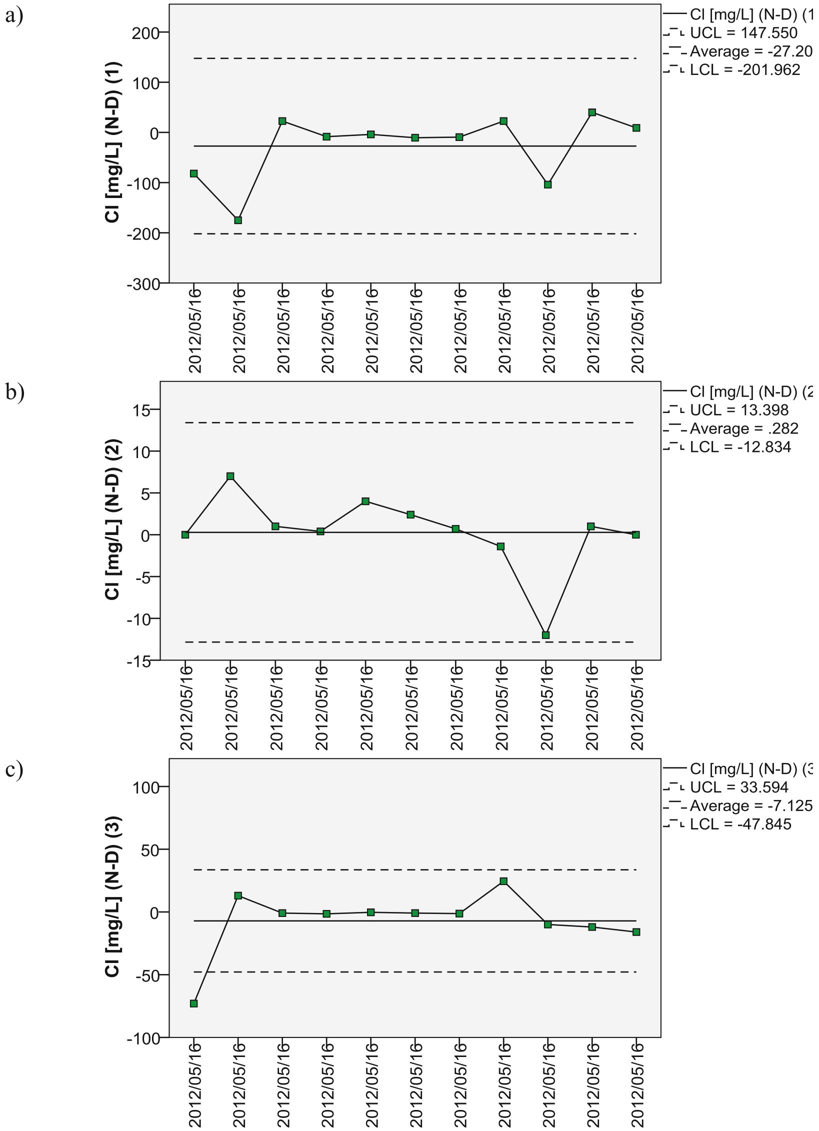


Fig. 2. Individual measurement control charts for chlorides concentration differences in normal and duplicate samples: (a) laboratory A, titration method, (b) laboratory B, titration method, (c) laboratory B, IC method.

The differences between results for normal and duplicate samples for different methods were plotted on individual measurement control charts (Fig. 2).

Differences in the results of chloride determinations obtained by titrimetric method in laboratory A are higher than in laboratory B (these are suspiciously small). The results can be regarded as stable, with the exception of the one point signal on the chart for the results obtained by ion chromatography.

3.3 Example 3: Testing Analysts

Hydrogeochemical laboratory systematically carries out different competence tests of analysts performing analyses. In the tests are included all analysts.

In this section two examples are presented – one for pH and EC measurements in real groundwater samples, second for chlorides determination in reference material.

The pH and EC measurements test was carried out in groundwater which has stable chemical composition. Two newly trained samplers performed pH and EC measurements at short time intervals, using the same calibrated equipment. The results are presented on charts of differences (Fig. 3).

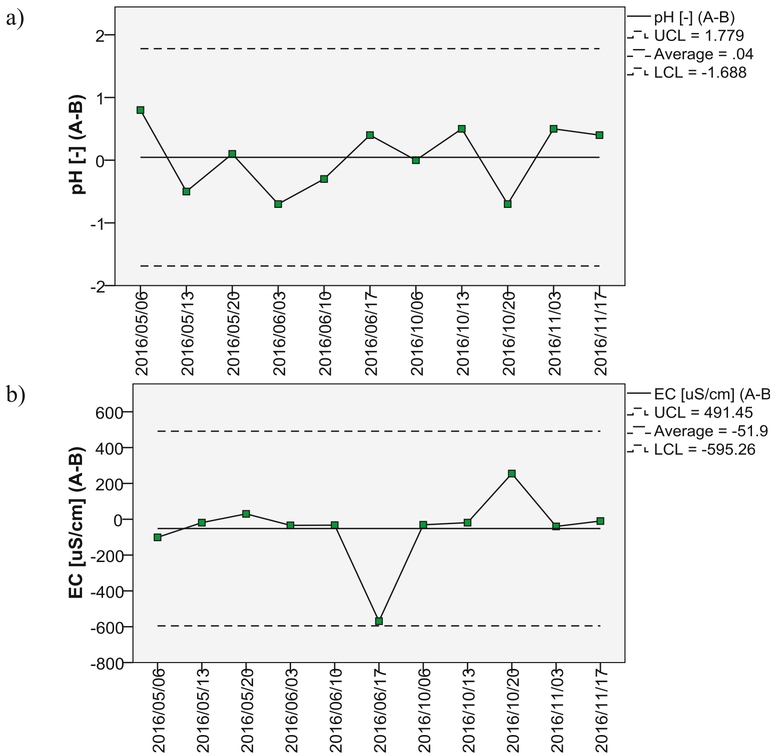


Fig. 3. Individual measurement control charts for results differences between samplers: (a) pH, (b) EC.

On the difference chart for pH measurements, the points oscillate around zero. There is a great match of results. On the difference chart for EC the greater variety of results is observed. Sequence plot for EC measurements shows that the both samplers got quite large spread of results (Fig. 4). The quality manager directed both samplers to additional internal training in the measurement of field parameters.

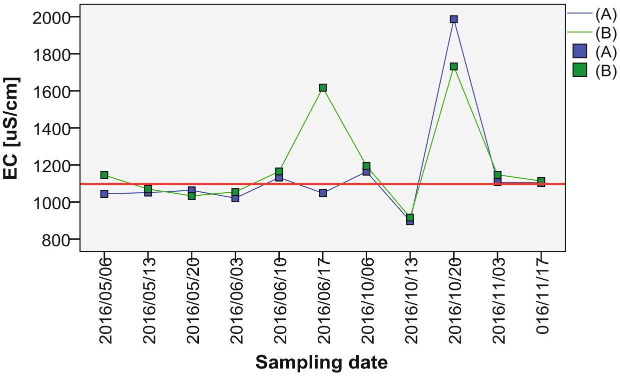


Fig. 4. Sequence plot for EC measurements (two samplers A and B).

During the next test three analysts were titrating the reference material. The experiment was conducted on the same day, every few dozen minutes. Titrant consumption was recorded without conversion the results for chlorides concentration (Fig. 5). Target consumption should be approximately 3.45 mL.

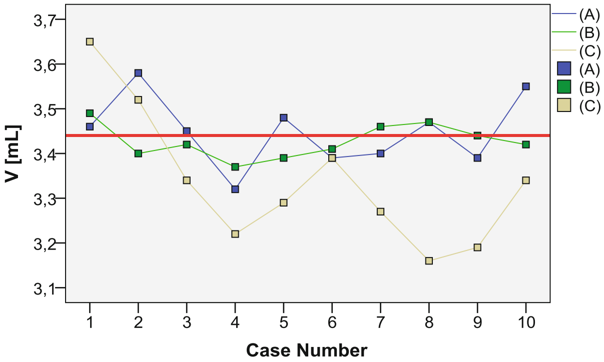


Fig. 5. Sequence plot for titration test (three analysts: A, B and C).

Sequence plot shows that the one analyst got decreasing trend of results (Fig. 5). It was caused by the fatigue of the analyst's eyesight during the day. The quality manager suggested that the analyst should perform more frequently instrumental analyzes that did not require color detection, and occasionally participate in titrimetric studies.

4 Summary

Sample preparation, equipment problems and human error are the top three categories of main causes of errors in chemical analysis [7]. Measurement methods are the better, the less dependent on the human factor, but in some cases this factor can't be excluded. When human beings are involved in some of the operations, a risk of human error remains even if we will minimize possible errors by the laboratory quality system. So it is very important to remember that human errors can have significant contribution to the measurement uncertainty. As such, it should be included in the uncertainty budget.

Human errors are most often caused by unqualified personnel. But we should also remember, that in the case of experienced samplers/analysts there is a risk of fall in routine, which can also generate errors.

The results of water testing always depends on the entire process, starting with the proper design of the monitoring point/monitoring network, correct sampling, proper preservation, storage and transportation of samples to the laboratory, sample processing and analysis of sample chemical composition using specific analytical methods. At each of these stages, the impact of the human factor is important. The impact of the sampler change or the methodology of analysis change on the uncertainty of determinations in water monitoring (groundwater, geothermal water and drinking water) has been presented in many publications, i.a. [35–38].

The ultimate purpose of testing the quality of waters is always a decision on e.g. the status of these waters, their suitability for a particular purpose or the effectiveness of the treatment method used. Our awareness of the quality of the result on the basis of which assessment is conducted and knowledge of its parameters, including e.g. uncertainty with the estimation of human errors, contributes to the credibility of the decision made.

Acknowledgements. The paper has been prepared under the AGH-UST statutory research grant No. 11.11.140.797.

References

1. Nielsen, D.M. (ed.): Practical Handbook of Ground-water Monitoring. Lewis Publishers, CRC Press, Boca Raton (1991)
2. Bulska, E.: Metrologia chemiczna (Chemical metrology). Malamut, Warszawa (2008)
3. Miller, J.N., Miller, J.C.: Statistics and Chemometrics for Analytical Chemistry. Prentice Hall, Harlow (2005)
4. Szczepańska, J., Kmiecik, E.: Statystyczna kontrola jakości danych w monitoringu wód podziemnych (Statistical quality control of data in the groundwater monitoring). AGH Publisher, Kraków (1998)
5. Kuselman, I., Pennecchi, F., Fajgelj, A., Karpov, Y.: Human errors and reliability of test results in analytical chemistry. Accredit. Qual. Assur. **18**, 3–9 (2013)
6. Kuselman, I., Pennecchi, F.: IUPAC/CITAC guide: classification, modeling and quantification of human errors in a chemical analytical laboratory (IUPAC technical report). Pure Appl. Chem. **88**(5), 477–515 (2016)

7. AMCTB No 56. What causes most errors in chemical analysis? Analytical Methods Committee. RSC Publishing. doi:10.1039/c3ay90035e
8. Ellison, S.L.R., Hardcastle, W.A.: Causes of error in analytical chemistry: results of a web-based survey of proficiency testing participants. Accredit. Qual. Assur. **17**, 453–464 (2012)
9. Hellier, E., Edworthy, J., Lee, A.: An analysis of human error in the analytical measurement task in Chemistry. Int. J. Cogn. Ergon. **5**(4), 445–458 (2001)
10. Joe, J.C., Boring, R.L.: Individual differences in human reliability analysis. INL/CON-14-31940. PREPRINT. PSAM **12** (2014)
11. Kuselman, I., Fajgelj, A.: Human errors and out of specification test results. Chem. Int. **35** (3), 30–31 (2013)
12. Kuselman, I., Kardash, E., Bashkansky, E., Pennecchi, F., Ellison, S.L.R., Ginsbury, K., Epstein, M., Fajgelj, A., Karpov, Y.: House-of-security approach to measurement in analytical chemistry: quantification of human error using expert judgments. Accredit. Qual. Assur. **18**, 459–467 (2013)
13. Kuselman, I., Goldshlag, P., Pennecchi, F.: Scenarios of human errors and their quantification in multi-residue analysis of pesticides in fruits and vegetables. Accredit. Qual. Assur. **19**, 361–369 (2014)
14. Kuselman, I., Pennecchi, F., Epstein, M., Fajgelj, A., Ellison, S.L.R.: Monte Carlo simulation of expert judgments on human errors in chemical analysis – a case study of ICP-MS. Talanta **130**, 462–469 (2014)
15. Kuselman, I., Pennecchi, F.: Human errors and measurement uncertainty. Metrologia **52**, 238–243 (2015)
16. Kuselman, I., Pennecchi, F., Bich, W., Hibbert, D.B.: Human being as a part of measuring system influencing measurement results. Accredit. Qual. Assur. **21**, 421–424 (2016)
17. Witczak, S., Kania, J., Kmiecik, E.: Katalog wybranych fizycznych i chemicznych wskaźników zanieczyszczeń wód podziemnych i metod ich oznaczania. (Guidebook on selected physical and chemical indicators of groundwater contamination and methods of their determination). Biblioteka Monitoringu Środowiska, Warsaw (2013). http://www.gios.gov.pl/images/dokumenty/raporty/ebook2_20130422.pdf. Accessed Mar 2017
18. Camões, M.F.: The quality of pH measurements 100 years after its definition. Accredit. Qual. Assur. **14**, 521–523 (2009)
19. Titration errors, http://www.titrations.info/titration-errors. Accessed Mar 2017
20. Kmiecik, E., Tomaszewska, B., Wątor, K., Bodzek, M., Rajca, M., Tyszer, M.: Implementation of QA/QC program in research related to the membrane processes used in geothermal water treatment. Desalination Water Treat (in press). doi:10.5004/dwt.2017.20604
21. Nordtest: Handbook for calculation of measurement uncertainty in environmental laboratories. Report TR537, Espoo, Finland (2003)
22. Nordtest: Uncertainty from sampling – a Nordtest handbook for sampling planners on sampling quality assurance and uncertainty estimation (based upon the Eurachem international guide estimation of measurement uncertainty arising from sampling). Report TR604, Oslo, Norway (2007)
23. Ramsey, M.H.: Sampling as a source of measurement uncertainty: techniques for quantification and comparison with analytical sources. J. Anal. Atom. Spectrom. **13**, 97–104 (1998)
24. Ramsey, M.H., Thompson, M.: Uncertainty form sampling in the context of fitness for purpose. Accredit. Qual. Assur. **12**, 503–513 (2007)

25. Ramsey, M.H., Thompson, M., Hale, M.: Objective evaluation of the precision requirements for geochemical analysis using robust analysis of variance. J. Geochem. Explor. **44**, 23–36 (1992)
26. Thompson, M., Howarth, R.J.: Duplicate analysis in practice – Part 1. Theoretical approach and estimation of analytical reproducibility. Part 2. Examination of proposed methods and examples of its use. Analyst **101**(1206), 690–698, 699–709 (1976)
27. Thompson, M., Coles, B.J., Douglas, J.K.: Quality control of sampling: proof of concept. Analyst **127**, 174–177 (2002)
28. Szczepańska, J., Kmiecik, E.: Ocena stanu chemicznego wód podziemnych w oparciu o wyniki badań monitoringowych (Assessment of the chemical status of groundwater based on the results of monitoring tests). AGH Publisher, Kraków (2005)
29. Postawa, A. (ed.): Best practice guide on the sampling and monitoring of metals in drinking water. In: Metals and Related Substances in Drinking Water Series. IWA Publishing, London (2012)
30. Kmiecik, E.: Analytical procedures for ion quantification supporting water treatment processes. In: Bundschuh, J., Tomaszewska, B. (eds.) Geothermal Water Management. CRC Press, Taylor & Francis Group (in press)
31. Rostron, P., Ramsey, M.H.: Cost effective, robust estimation of measurement uncertainty from sampling using unbalanced ANOVA. Accredit. Qual. Assur. **17**(1), 7–14 (2012)
32. Kmiecik, E.: Metodyczne aspekty oceny stanu chemicznego wód podziemnych (Methodological aspects of assessing the chemical status of groundwater). AGH Publisher, Kraków (2011)
33. Drzymała, M.: Chlorowce w wodach kopalnianych GZW i ich wpływ na wody powierzchniowe oraz osady denne zlewni górnej Odry (Chlorides in Upper Silesian Coal Basin mine waters and their impact on surface water and bottom sediments of Upper Odra basin). Doctoral dissertation. AGH-UST (2013)
34. Dwornik, M., Kmiecik, E., Bebek, M.: Wpływ metod analiz na niepewność oznaczeń chlorowców w wodach powierzchniowych zlewni górnej Odry i w wodach kopalnianych do niej odprowadzanych (Influence of analyses method on uncertainty associated with halogens determination in Upper Odra River Basin surface water and coalmine water inducted to it). Przegl. Geol. **63**(10/1), 705–709 (2015)
35. Kmiecik, E., Podgórni, K.: Ocena wpływu zmiany próbobiorcy na niepewność związaną z opróbowaniem w monitoringu wód podziemnych (Estimation of sampler influence on uncertainty associated with sampling in groundwater monitoring). Geol. Bull. PGI **436**(9/1), 253–260 (2009)
36. Postawa, A., Kmiecik, E., Wątor, K.: Rola osoby próbobiorcy w monitoringu jakości wód przeznaczonych do spożycia (The role of a sampler in drinking water quality monitoring). In: Sozański, M.M. (ed.) Water supply and water quality – present issues PZITS 874 (2010)
37. Wątor, K., Kmiecik, E., Postawa, A.: Wybrane problemy kontroli jakości w monitoringu wód przeznaczonych do spożycia przez ludzi (Selected aspects of quality control in monitoring of water intended for human consumption). In: Problemy wykorzystania wód podziemnych w gospodarce komunalnej: materiały na XVIII sympozjum naukowo-techniczne pt. Dokumentowanie i eksploatacja małych i średnich ujęć wód podziemnych. Częstochowa, PZIiTS (2010)
38. Korzec, K., Kmiecik, E., Mika, A., Tomaszewska, B., Wątor, K.: Metodyka opróbowania ujęć wód termalnych – aspekty techniczne (Metodology of thermal water sampling – technical aspects). TPG. Geotermia, Zrównoważony Rozwój **1**, 75–87 (2016)

Evaluating Expert Performance

Human Performance Variability in Task Execution Times Under Generic Human-System Integration Conditions in Naval Operations

Tareq Ahram[1,2(✉)], Waldemar Karwowski[1,2], and Kevin Muhs[2]

[1] Institute for Advanced Systems Engineering, University of Central Florida,
Orlando, FL, USA
tahram@ucf.edu
[2] Department of Industrial Engineering and Management Systems,
University of Central Florida, Orlando, FL, USA

Abstract. This main objective of this paper was to assess and model variability of task performance relevant to human-systems integration efforts in naval ship operations, including the estimation of task execution times under generic task conditions. A method was developed for quantifying and modeling human operator execution times in selected naval ship operations. The proposed method is based on time contributions for each task component with due consideration of three core task performance elements: skills, knowledge and task requirements. The experimental analysis utilized a hybrid approach based on linear regression, weighted scoring method and artificial neural networks. The proposed modeling approach demonstrates promising results for developing a realistic solution for assessment of task performance times relevant to training requirements and total ownership cost for competing technology upgrades, with emphasis on maintaining manpower readiness and mitigating possible performance degradation due to unforeseen mission conditions.

Keywords: Human performance · Variability · Human-system integration

1 Introduction

Human operators show significant variability in performance when operating complex systems during normal Navy ships operations. Little is known about human performance variability in naval ship operations, while various approaches to human-systems integration rely on technology interventions, training, or selecting personnel who operates the system. The technology acquisition communities identified targeted and possible solutions to improve total system performance by quantifying total system error as system error, operator error and a random error. For example, to increase human-system compatibility and to design better systems, various technological solutions need to understand the cause of the variability to determine whether future investments should be directed toward technology or training resources.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_17

From a human performance point of view, both internal and external factors affect the ability and effectiveness of an individual to perform an identified task [1, 2]. Understanding the temporal consequences of these factors in dynamic, time-constrained environments is an essential element for integrating the human component into complex systems.

Variability in performance response times is the ubiquitous outcome of these human temporal consequences, as they form the critical factor to understand the broad range of human-system interactions from simplistic to complex setup. As a result, accurate prediction of the factors affecting human performance variability within the context of individual task performance, both cognitive and physical, as well as the development, refinement, and use of reliable tools in assessing this variability has been a major focus in research for well over 50 years. The broad coverage in literature of factors affecting human performance produced significant gains in understanding the individual elements as well as the organizational factors that impact performance as well as temporal variability. For example, research conducted by Hick [3]; Hyman [4] and Fitts [5] have been generalized and extremely far reaching in the field of human factors and ergonomics (HF/E), whereas others [6] have been exceedingly limited in their scope and application. This variance is not unexpected given the broad desire to create both generalized and adaptive rules to human response as well as the recognition that context, specificity of the task, also plays a significant role. The variance is also indicative of partially unpredicted human behavior nature under stressful situation and the shift in human sciences from prescriptive models to descriptive models in terms of a rational performance standard towards modeling the “actual behavior” as described by Rasmussen [7].

2 Mechanism of Human Task Performance

Successful human task performance requires the cumulative effort of both the cognitive abilities as well as the physical and sensory abilities of the individual completing the task. In this context, cognitive abilities refer to the ability to process information and make task related decisions based on environmental perceptions, knowledge, and memory. Physical abilities describe the coordinated muscle action needed to complete sustained effortful muscular work. In socio-technical systems, the human is central to the system and will respond, enabled by their attention to surrounding stimuli, based on how they perceive their environment through visual and auditory perception [8]. This response, dependent on the type of action required, which will require individuals to leverage their cognitive and physical abilities in completing the task. These actions occur principally in series with one another, although some components of the response may be conducted in parallel. The task environment, on the other hand, consists of all the elements, both internal and external to the individual, that impact the human response [9]. A generalized view of the human task performance relationship is provided in Fig. 1.

Figure 1 identifies possible interdependencies between the defined performance environment and the human component, in order to measure the system impact, both the physical and cognitive performance characteristics on the individual. In addition, the dotted line recognizes existing relationships and likely interplay between the physical and cognitive abilities of the individual in completing the task. An outcome of

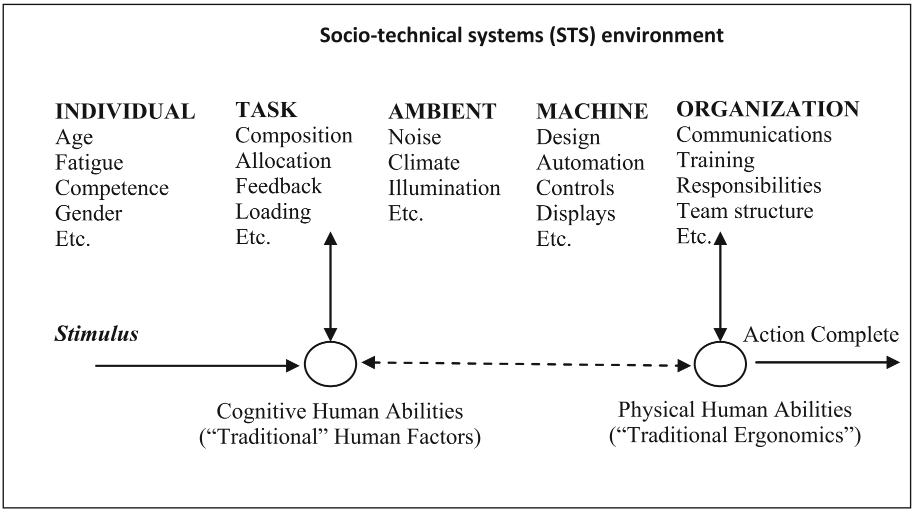


Fig. 1. Relationships and interdependencies of physical and cognitive human capacities within the socio-technical systems (STS) environment (adapted from Marras & Hancock, [9]).

these relationships is the cumulative nature of the time required for both the cognitive and physical functions in the completion of a task, as this will be explained in the taxonomy of human task abilities in the next section.

3 Research Method

This study utilized a dataset of (5860) unique human-system integration tasks ranging from simple, intermediate to complex activities, each of which describing a unique but virtually all possible Navy ship operations. The experimental analysis aims to classify those tasks into groups and categories based on each unique task description, narrative, complexity, importance and performance time. Each entry is associated with task text description, workload, criticality, frequency of task execution, category and minimum, average and maximum task performance times. Each task description, regardless of category or type, was considered as composed of a combination of unique set of task building blocks or components which are knowledge, skills and abilities. For example, some tasks can be described in two components, while those complex (performance demanding) tasks can be constructed with 15 or more components.

Therefore, it was recognized that in order to investigate the basic building blocks for each task, and to better understand the reason of performance variability and/or task complexity, tasks were grouped based on knowledge, skills or abilities composition. It is important to notice that the unique order and sequence of the task components define the task structure, performance time and reason of performance variability during standard ship operation.

Linguistic task description and clustering was conducted by the application of text analytics (text mining) using natural language processing and N-gram approach in MATLAB R12, results identified a taxonomy of (343) possible task group, based on their description word structure which are: 46 Skills, 109 Knowledge Requirements and 52 Abilities, all of which can describe virtually all possible human performance task or operation in a naval ship. Each task (T_c) in the dataset is made up of at least (2) and up to (15) unique human performance component $(2 \leq Tc \leq 15)$, this means that there are 125 possible combinations of tasks based on skills, knowledge and ability requirements as follows:

$$\boldsymbol{Tc} = \left(\sum_{i=1}^{5} \boldsymbol{KnowledgeRequirements} + \sum_{i=1}^{5} \boldsymbol{SkillRequirements} + \sum_{i=1}^{5} \boldsymbol{AbilityRequriements} \right)$$

The combination of various task components describe virtually any possible task within the ship environment. Each of which reflects unique task requirements with respect to *Knowledge (K), Skill (S) and Ability (A)* task components, as shown below in Table 1 for the task performance taxonomy where:

$$\begin{aligned} 0 \leq KnowledgeRequirement \leq 5 \\ 0 \leq SkillRequirement \leq 5 \\ 0 \leq AbilityRequirement \leq 5 \end{aligned}$$

In order to calculate performance time for each task component, the following soft computing methods were applied: weighted scoring (WS), linear regression (LR), and artificial neural networks (ANN's). The weighted scoring method is similar to multi-criteria analysis. This approach involves identification of all the descriptive factors (here performance components) which are relevant to the task; the allocation of weights or importance for each of them reflects their relative importance; for example, performance time for physical tasks reflects physical skill or ability score more than knowledge in relation to physical task performance, while tasks that involves cognitive processing relies on knowledge or for example problem solving abilities more than anything else. The result is a weighted score for each task component, which may be used to indicate and compare the overall performance of the component in numeric terms.

The dataset studies includes the combination of knowledge, skill and abilities requirements for each task, where the normal distribution is used in order to calculate the time distribution for each task component. In order to calculate the time distribution for each task component, a matrix was constructed of size $n \times m = [5860 \times 15]$, where each row corresponds to the unique task components, keep in mind that each task is unique with respect to task requirements composition, therefore a "0" in one cell means that for there is no value for specific task element in that row, since some but not all tasks require certain but not all skill, knowledge or abilities. To calculate the task performance time the following matrix captures this task performance data:

$$\begin{bmatrix} TaskID \\ 1 \\ \vdots \\ \vdots \\ 120027 \end{bmatrix} = [K1\, K2\, K3\, K4\, K5] + [S1\, S2\, S3\, S4\, S5] + [A1\, A2\, A3\, A4\, A5]$$

For example:

$$\begin{bmatrix} Task \\ ID \\ 52231 \end{bmatrix} = [35\,50\,50\,35\,35] + [1\,2\,7\,21\,41] + [15\,9\,7\,20\,24]$$

$$= [\mathrm{Min} = 10,\ \mathrm{Average} = 12.5,\ \mathrm{Max} = 15]$$

Table 1 illustrates the example task breakdown into knowledge, skill and ability requirements and their respective performance times. For example, reading comprehension in task #52231 takes on average 1 min to complete, minimum 0.8 min and maximum 1.2 min. Table 1 shows results from linear assessment of task performance components without applying ANN or MLRP modeling. Keep in mind that in order to create universal task performance components, researchers analyzed all tasks regardless of task category to derive task performance time variability with respect to time demand needed for each knowledge, skill and ability components.

The following is another example illustrating the approach implemented for coding the human performance task components and the calculation method for one example combat/assessment task description:

"Watchstander receives the order to execute the sensor support plan for the UAV"

Table 1. Example of results for simulated task performance time breakdown with respect to task components of knowledge, skill and ability requirements for task #52231.

Task_Num	52231
Workload_Cat	Operational manning
Task_Cat	Combat
Task_Type	Assess
Components	11
Knowledge Weights	3
Skill Weights	4
Ability Weights	4
Knowledge Time	3.41
Skill Time	4.55
Ability Time	4.55
Task_Name	TAO review response error alert for situational awareness
Task_Desc	An ongoing engagement with an air track has failed. The TAO reviews the failure alert and decides whether or not to re-engage using the same response as the one that failed. Or if time allows, to select another response option.
Min_Time	10
Mean_Time	12.5
Max_Time	15

Table 2. Example task breakdown.

Task_Num	Task_Name	Workload_Cat	Task_Cat	Task_Type
6432	Receive execute sensor support Plan-UAV	Operational manning	Combat	Detect

The above example was coded as shown in Table 2 to summarize the workload category, task category and task type, along with the detailed task breakdown into the human performance components of knowledge, skill and abilities required to accomplish this task as shown below. Tasks categorized based on their complexity and performance demands into six task performance-related regions as shown in Fig. 2, notice that the Figure shows original concept developed by the researchers based on theory by de Waard [10]. Successful task performance is based on knowledge pre-requisites, ability assessments and response execution with proper skills.

The various regions in Fig. 2 can be described as Optimal Performance (A2): where the operator is able to cope with workload, and performance remains at levels approaching optimal state. Increased Workload Demand (A3-B-C): operator performance not typically show any noticeable decline in performance despite increased task

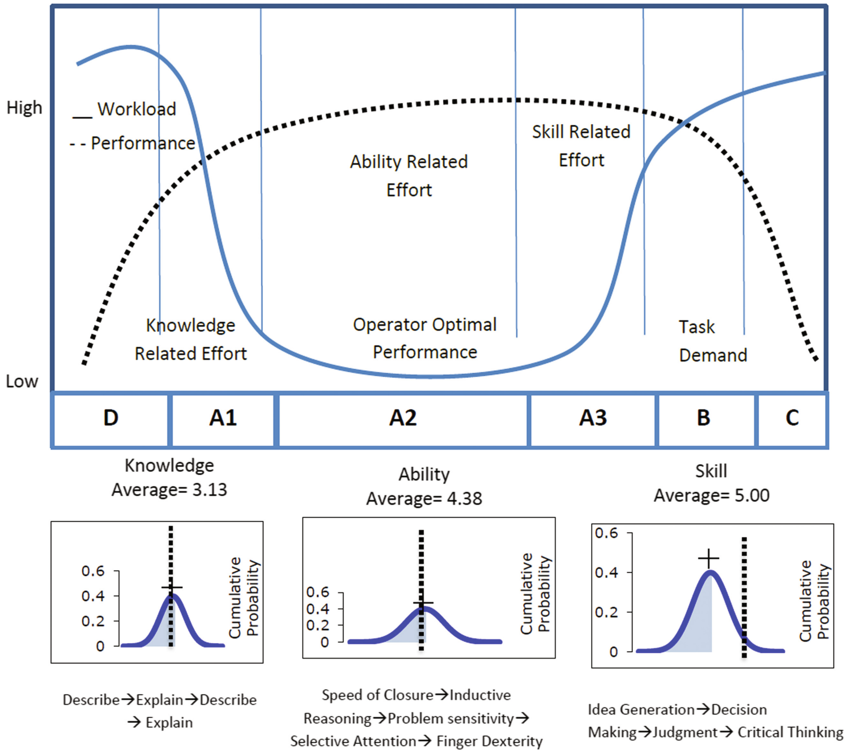


Fig. 2. Standard task performance and workload as a function of task demand

loading until entering A3 region. Decreased Task Demands (A1-D): reduced task demands from the A2 region transition into region A1 in which maintenance of performance relies heavily on increased operator vigilance. "It should be noted that, contrary to popular belief, vigilance is a cognitive activity, and implies increased, rather than decreased, cognitive workload levels. Acceptable levels of vigilance can only be maintained in the A2 region by increased cognitive effort", although in this case, the effort needs to be directed towards the maintenance of a vigilant state by operators keeping the cognitive system in a state-of-readiness for response, rather than directing attention toward incoming task demands as in the case of high workload situations [11].

4 Discussion and Conclusion

Consideration of human performance variability elements is essential for assuring task performance and training effectiveness. This study is based on a concept of the ideal human observer, a novel approach introduced by Krebs [12] for exploring naval manpower performance variability with respect to potential human error and reliability estimates for any functional, tactical or operational naval task. The overall goal of this study was to develop a useful method for quantifying human performance variability, and estimate human performance task execution times and its variability. The variability of human performance execution times were estimated based on the normality assumption for operator population, with the task specific information and task components.

The applied methodology utilized a hybrid approach based on linear regression, weighted scoring method and artificial neural networks. Such a methodology should help better optimize human resource allocations, develop effective training strategies, and optimize manpower total ownership costs in naval operations.

The data analysis performed in this paper should help establish a body of knowledge and enrich the research on task performance measurement with the focus on naval ship operations. The proposed approach should be useful to government acquisition programs and decision makers in assessing cost/error/variability, and to mission critical industries such as in oil and gas exploration or power generation, to ensure proper assignment of tasks, training preparedness, and addressing challenging problems of human performance assessment in the context of continuous technology upgrades, and evolving mission goal. Ultimately, the proposed approach should help in the development of useful solutions to model and/or assess human performance and explain causes of performance variability and reasons for errors or safety risks.

Acknowledgements. This research was sponsored by the Office of Naval Research Contract No. N00014-14-1-0777. Authors would like to acknowledge the support of Dr. William Krebs and Program Management.

References

1. Chen, J.Y., Joyner, C.T.: Individual differences in concurrent performance of Gunner's and robotic operator's tasks. In: Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, pp. 1759–1763. Human Factors and Ergonomics Society, San Francisco (2006)
2. Stanton, N.A., Baber, C.: Modelling of human alarm handling response times: a case study of the Ladbroke Grove rail accident in the UK. Ergonomics **51**(4), 423–440 (2008). doi:10.1080/00140130701695419
3. Hick, W.E.: On the rate of information gain. Q. J. Exp. Psychol. **4**, 11–26 (1952)
4. Hyman, R.: Stimulus information as a determinant of reaction time. J. Exp. Psychol. **45**(3), 188–196 (1953)
5. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol. **47**, 381–391 (1954)
6. Chan, A.H., Shum, V.W., Law, H.W., Hui, I.K.: Precise effects of control position, indicator type, and scale side on human performance. Int. J. Adv. Manufact. **22**, 380–386 (2003). doi:10.1007/s00170-002-1491-z
7. Rasmussen, J.: Risk management in a dynamic society: a modeling problem. Saf. Sci. **27**, 183–213 (1997)
8. Salvendy, G. (ed.): Handbook of Human Factors and Ergonomics. Wiley, New York (1997)
9. Marras, W.S., Hancock, P.A.: Putting mind and body back together: a human systems approach to the integration of the physical and cognitive dimensions of task design and operations. Appl. Ergon. **45**(1), 55–60 (2014)
10. De Waard, D.: The Measurement of Drivers' Mental Workload. Ph.D. Thesis, The Traffic Research Centre VSC, University of Groningen, The Netherlands (1996)
11. Warm, J.S., Dember, W.N., Hancock, P.A.: Vigilance and workload in automated systems. In: Parasuraman, R., Mouloua, M. (eds.) Automation and Human Performance: Theory and Applications. Erlbaum, Mahwah (1996)
12. Krebs, W.: ONR Presentation for Ideal Observer (2013)

Are We Flooding Pilots with Data? – Effects of Situational Awareness Automation Support Concepts on Decision-Making in Modern Military Air Operations

Christopher Roos(✉)

Aerospace Operations, Netherlands Aerospace Centre NLR,
Anthony Fokkerweg 2, 1059 CM Amsterdam, The Netherlands
`Christopher.Roos@nlr.nl`

Abstract. Within highly dynamic situations, the amount of relevant information that a pilot needs to process to make an informed decision can be substantial. With an ever increasing amount of data available to the pilot there is a real risk that not all relevant data can be taken into consideration, or that this process exceeds the time available. To maintain the operational effectiveness of pilots and other military aviation operators in the face of these developments, a prototype system was developed that aims to support the operator in the development of Situational Awareness (SA) by deriving essential information from the large body of data and presenting this to the user in an efficient Human Machine Interface (HMI). Extrapolation of data allows the user to preview the effects of changes and adapt accordingly. Results of a human-in-the-loop simulation exercise are presented in this paper, along with potential concept issues.

Keywords: Situational awareness · Prototype system · Air operations · Adaptation · Resilience · Human factors · Decision-making

1 Introduction

Recent armed conflicts demonstrate that opponents are becoming far better and faster in responding to Western tactics. Increasingly, the situation calls upon the pilot's ability to adapt [1–3]. Within highly dynamic situations such as can be encountered in military operations, the amount of relevant information that a pilot needs to process to make an informed decision can be substantial. With an ever increasing amount of data available to the pilot there is a real risk that not all relevant data can be taken into consideration, or that this process exceeds the time available. To maintain the operational effectiveness of pilots and other military aviation operators in the face of these developments, a study was conducted by the Netherlands Aerospace Centre – NLR for the Royal Netherlands Air Force (RNLAF) with the aim of enhancing the development of situational awareness, and subsequently the decision-making process, by deriving relevant information from large amounts of contextual data.

Unforeseen dynamic events during mission execution can require adaptations: from small changes in the route or changes in tactics to complete abandonment of mission

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_18

goals. The decision making process that underlies these adaptations requires the operator to have a good Situational Awareness (SA). Situational awareness is defined in several levels [4]:

Level 1:	The perception of the elements in the environment within a volume of time and space,
Level 2:	The comprehension of their meaning and
Level 3:	The projection of their status in the near future

The build-up of SA requires extracting relevant information from the environment and integrating this information with existing knowledge into a mental image [5]. As a product of their complexity, military air operations often require operators to poses a high level of SA. When changes occur or adaptations are necessary, only perceiving and comprehending the different elements is insufficient for determining the necessary adaptations the operator has to make. Determining the future status of different elements, and the direct or indirect effect this could have on the operator/team/mission is crucial in the military decision-making process.

However, the process of developing a solid level 3 SA of a situation is challenging and can therefore take considerable time. With recent technological advances, large amounts of information are becoming available to the operator. Extracting the relevant information from this ‘lake of data’ is time consuming. Together, this can greatly impact the effectiveness of SA build-up and ultimately the speed and/or the effectiveness of the decision-making process. Smart use of (automated) data processing methods and efficient interfaces have the potential to help the operator effectively process large amounts of relevant information, thereby making the decision-making process more efficient.

To test this, NLR developed a prototype system (see Fig. 1) that aims to support the operator in the development of SA by deriving essential information from a large body of data and presenting this to the user in an efficient Human Machine Interface (HMI). Through analysis and extrapolation of existing data, a prediction of a future state can be made. This allows the operator to visualise the future effect of an action by ‘previewing’ them, potentially increasing operator resilience by rapidly developing a level 3 SA and choose the most effective course of action.

2 Deriving Relevant Information – A Prototype SA Support System

For rapid development of SA, it is necessary to derive the relevant information from the ‘lake’ of available data. Central to the design philosophy of the system prototype is the construct of “effect on the operator”. Military flight crewmembers determine the relevance of information on the basis of the effect it could have on them/their crew and can be categorised from highly relevant to irrelevant. Information that directly affects the operator has more relevance than information that indirectly affects the operator or does not affect the operator. There are two main questions underlying the features of the prototype:

Fig. 1. System prototype evaluation within NLR Helicopter Pilot Station (HPS)

(1) Can/will it have an effect on the operator?
(2) What is the effect on the operator?

Air operations are often complex and multi-facetted, stretching over longer periods of time. Changes occurring in one stage of the operation can have profound effects later on in the mission. Therefore, understanding the effect of a change on the remainder of the mission is a crucial aspect. By deriving relevant information from large sets of available data, the amount of information presented to the operator is kept to a minimum. A reduction in the number of elements that needs to be processed facilitates effective perception (Level 1 SA) and comprehension (Level 2 SA) of these elements and increases overall processing speed. Projecting comprehension of a situation forward in time and understanding the associated consequences is considered level 3 SA. The system prototype is designed to support the operator in quickly developing level 3 SA by real time calculation of the effect of changes. The system achieves this by using advanced algorithms to analyse several different data streams in real time and extrapolate a logical future state. This technology also allows the operator to 'fast forward in time' to not only see the impact of a change, but also how the operators' (contingency) actions might play out.

By showing what will be the effect on the operator, the operator is allowed to 'preview' the effect of a decision, projecting the information forward in time and effectively providing a real time level 3 SA. The operator is then able to choose the most effective measure, and instantaneously develop understanding of the effect this measure could have. The 'previewing' of the effect of a decision also allows the operator to rapidly reconstruct a higher level SA after adapting to a dynamic event.

This allows the operator to remain flexible and adaptive to further dynamic events that might take place. The example below illustrates how the system achieves this.

Example: while en-route, a previously located and identified enemy vehicle suddenly starts moving towards the Landing Zone of a transport helicopter. Whereas detecting (level 1, “it starts moving”) and understanding (level 2, “it poses a danger”) the threat is relatively straight forward, projecting this information forward in time (level 3, “it can/will affect me”) to determine the effect is not. Whether this vehicle will have an impact on the operator depends on several factors which makes determining whether attention should be given to the vehicle or not very difficult.

Through a path planning algorithm the prototype calculates the most likely route of the vehicle. The system combines this information with general characteristics of the vehicle such as speed, its Weapon Engagement Zone (WEZ) and the helicopters (real time) flight plan data to determine whether the vehicle is a potential threat, and the timeframe in which the vehicle is able to intercept the helicopter. If the vehicle is able to intercept the helicopter, an indication is presented to the operator on the place, time and severity of the threat (see Fig. 2). Indicated by the potential danger, the operator can perform a real time adjustment of the parameters of the flight plan (change route, fly faster or slower, etc.) and immediately see the projected results. By performing real time recalculations of the plan, the operator is able adjust the flight plan and avoid the danger.

Figure 2 (below) shows what the operator is shown in the Graphical User Interface (GUI). The threat is indicated on the map by a red icon, on the flight path (red curtaining) and on the timeline below.

To test the innovative new features described above, a completely new systems architecture had to be build. Within an operationally ready product, these new features would normally be supplemented by a plethora of other common system features such as navigation tools, grids, weather threats, etc. Replicating all these features within the prototype system was not feasible. Within the prototype system only the most essential features were included.

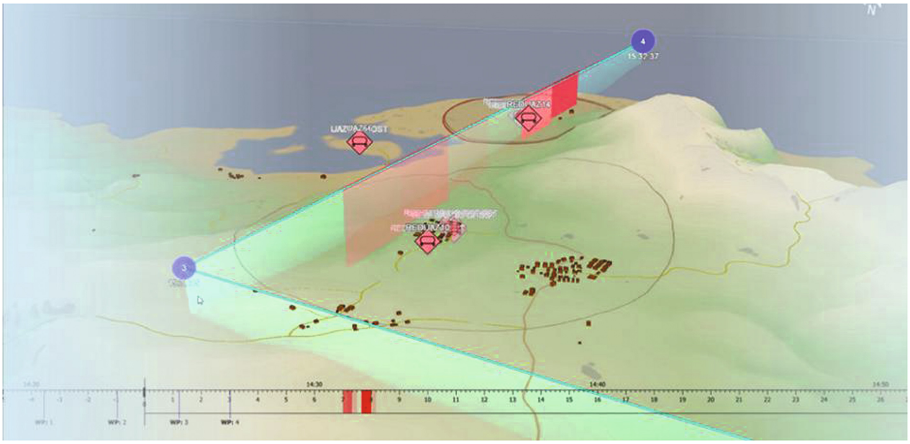


Fig. 2. System prototype GUI providing information on potential future threats

3 Concept Validation

To determine the potential of automated systems in supporting the operator in processing large amounts of data, a Concept Development and Experimentation (CD&E) process was followed. Within this CD&E process, the prototype system was developed and subsequently evaluated with the potential end-users of such a system (military helicopter crews). To develop this system, user requirements were first determined based on in-house knowledge and supplemented with Subject Matter Expert (SME) input. These user requirements were subsequently transformed into functional and technical requirements for the prototype. Through design workshops, a concept design was developed that detailed the functionality of the prototype, and Human-Machine Interaction (HMI). The design was subsequently developed into a working prototype. Finally, the effectiveness of the prototype was evaluated with end-users in a human-in-the-loop simulation exercise at the NLR (see also Fig. 1).

Several simulated missions, characteristic of those encountered in real life, we run. The participants were subsequently asked to fill in a standardised questionnaire. In this standardised evaluation, every participant was asked to determine how (to his/her opinion) the system prototype could work in real life military operations. Specifically, the participants were asked how the system prototype could affect the SA build up, adaptability (resilience) of the operator and what would be the effect of using such a prototype on overall mission effectiveness. With each rating, rationales were asked. Specifically, the participants were also asked if, and how the system prototype could cause problems in operational use.

4 Results

The participating SMEs were asked to rate (on a scale from 1 to 5, 1 being *"totally disagree"* and 5 being *"totally agree"*) whether the developed prototype system could be beneficial in supporting the operator in several aspects of his/her decision-making. The following summarized results were found:

Personal SA:	$M = 4.6$, $SD = 0.5$
Shared SA:	$M = 4.1$, $SD = 0.4$
Adaptation/resilience:	$M = 4.3$, $SD = 0.8$
Mission effectiveness:	$M = 4.6$, $SD = 0.5$

Further inquiries were made into the underlying aspects of these effects. The results of the evaluation indicate the development of personal and shared Situational Awareness can be supported by the prototype system. A number of aspects contribute to the potential of the system to increase personal SA. Mainly, the system provides the user with an up-to-date overview of both friend and foe (positional data) and indicates changes within this. For the user, this relates to knowing where he and his team members are in relation to the enemy, in which direction they are moving and a prediction of their potential future locations. In addition, the ability to create and share (new) plans allows the users to (re-)establish a 'picture' of the mission, target and route.

This allows the user to set (updated) operational information against the (updated) mission plans.

Concerning shared SA, the main contributing factors were the ability of the system to provide a clear overview of the situation, specifically the position of both friendly and enemy forces. Furthermore, the system shows the flight paths and positions of the 'ownship', and other team members and allows the user to extrapolate his (and his team members) location at a future point in time. By showing expected hazards when 'retasking' a flight plan, the system eases the game plan Command and Control (C2) while in-flight.

Participants indicated the improvements in mission effectiveness would be a result of a reduction in risk by providing a better overview of hazards and saving time [to process the information]. Through an improved SA, the user is able to more effectively avoid threats. Furthermore, all players have similar, updated information available. This allows the team to adapt faster and more effectively to changes.

5 Discussion

Although the results of the simulation exercises indicate the system could be beneficial in the development of SA, it does not escape some of the more fundamental issues regarding SA and decision-making and some of the system features also produced (new) negative side effects.

In the process of data extrapolation, data is analysed and processed within the system, resulting in a filtering of the relevant information. Operational data is – in practice - however far from 100% reliable; it can be unverified, ambiguous, out-dated, incomplete or sometimes even false. This not only requires a high level of flexibility from the perspective of system architecture, but also requires specific strategies for dealing with partly- or un-reliable information.

How reliable the data (or processed information) could or should be, depends on the means and time available to validate the information and the information risk tolerance of the operator. Figure 3 below illustrates the information validation relevance function as a product of the information risk tolerance and the time to validate.

Contrary to civil operations, within the constraints of a military operation, the required validity of information is not a set standard. The validity of information that is required by the operator before using it for decision-making depends on the situation the operator is in (risk/priority) and the time available to validate information (cost). As the amount of time available increases, the operator is better able to validate the data and information that is presented to him. As the amount of available time increases, the tolerance to invalidated information decreases. There are limitations to this function, for example when the information becomes too unreliable to use or when the time taken to validate becomes larger than the refresh rate of the information. Depending on the context, the amount of required validity can however differ; under 'normal conditions' the amount or required validity might run linear to the time available (a) whereas in a high priority (e.g. potential friendly fire) scenario, the operator might have a lower tolerate for risk of invalidated information (b). Conversely, if the operator is under imminent threat, the tolerance for invalidated information might be higher than normal.

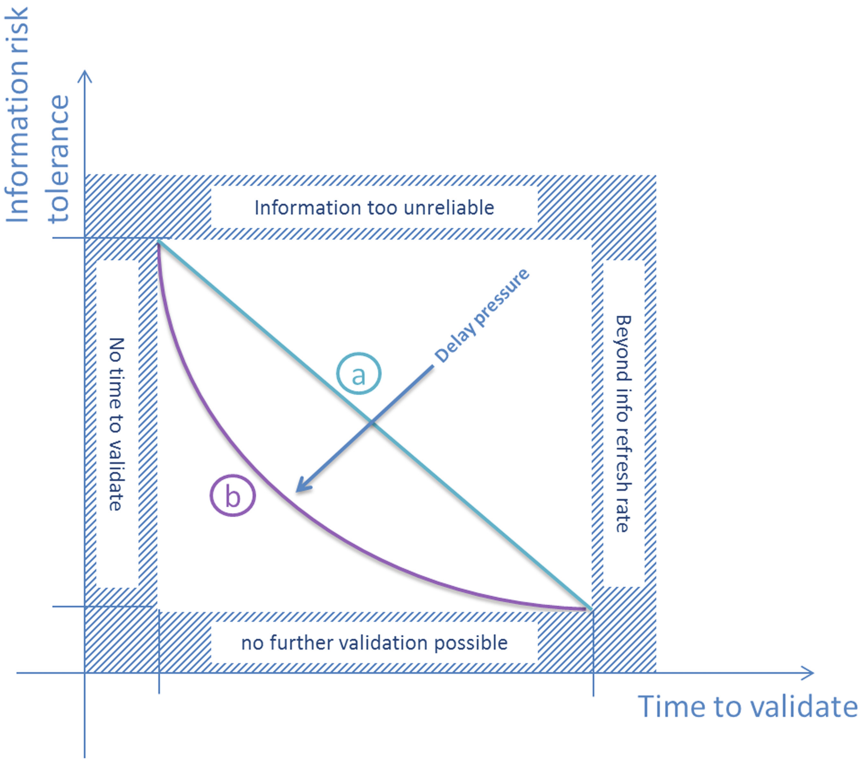


Fig. 3. Information validation relevance function (information risk tolerance x time to validate)

Furthermore, as data in the system is analysed, processed and subsequently presented to the operator as information, the imperfections inherent in the data become masked. By masking these imperfections, the operator is unable to ascertain the true value of data/information, increasing the likelihood he might perceive the information presented as absolute truth. There are two distinct problems that might develop from this: (1) The operator might develop a false sense of security. As information is presented as absolute and truth the operator will not be motivated to validate the information further, even if time permits it. (2) Situational Awareness is developed on the basis of the information that is present and processed by the operator. Information that is incorrectly perceived as absolute truth thus can cause incorrect SA development. This incorrect SA needs to be deconstructed and subsequently reconstructed, most likely in a situation with smaller time margins and decision-making options.

Lastly, within the system, the user is able to see near-real-time information of friendly and enemy forces. Predictions or prognoses can be made on future states of your 'ownship', team members and enemy forces. The SMEs that participated in the validation of the system indicated that without a good indication of the reliability of the information, this could lead to target fixation. Target fixation can reduce the scope of the person operating the system. Furthermore, the SMEs cautioned that the prediction of future states can cause erroneous construction of SA, taking prediction as fact. When later presented with a (slightly) different reality, it will be especially difficult to detect this and reconstruct the SA.

6 Conclusions

The aim of the research program was to test innovative (technological) concepts that could support the military flight crew in developing and maintaining situational awareness. Through a CD&E process, a SA enhancing system was developed and subsequently tested in a semi-realistic (simulation) environment with the intended end-users (military helicopter pilots/crew). The results of this CD&E process indicate the developed system can have a substantial positive effect on the personal and shared Situational Awareness of the military operator.

Subsequently these SA improvements can have a positive effect on the adaptation to dynamic mission events and ultimately result in a positive effect on mission effectiveness. SMEs participating in the simulation exercises indicated the system could reduce risk through an enhanced, commonly shared SA, reduced necessity to coordinate and more efficient decision-making processes.

However, the system could not overcome inherent, fundamental aspects and several negative side effects of system use were noted. The concept utilises advanced algorithms to extrapolate and ‘predict’ red and blue force future states. As military operations are highly complex and have a high level of inherent variability and freedom of movement, it is likely the predictions divert from reality in various degrees. When predictions are used to build up SA, an incorrect ‘picture’ might develop, preventing the operator from effective decision-making or requiring him to rebuild his SA on location. The extrapolation of information into future states through ‘previewing’ the effect of a decision is therefore – until truly reliable systems emerge – only a supporting function. This function allows the operator to get an indication that the intended actions will have the desired effect. In addition, the function could be preferential in situations where the operators’ SA is too low for effective decision-making, e.g. due to extremely dynamic circumstances. The operator should, when circumstances permit, subsequently allocate available resources to re-establish a high level SA through normal processes. Additionally, the course of action chosen by the system should always be validated to determine if it (still) best matches the mission goals.

References

1. Osinga, F.P.B.: Science, Strategy and War: The Strategic Theory of John Boyd. Routledge, Abingdon (2007)
2. HCSS: VLEUGELS VOOR NEDERLAND - Naar nieuwe partnerschappen voor de militaire lucht- en ruimtevaart, 3 August 2016. http://www.hcss.nl/reports/nieuwe-hcss-studie-voor-de-koninklijke-luchtmacht-vleugels-voor-nederland-naar-nieuwe-partnerschappen-voor-militaire-lucht-en-ruimtevaart/206/
3. Anrig, C.F.: The Quest for Relevant Air Power. Air University Press, Montgomery (2011)

4. Endsley, M.R.: Design and evaluation for Situational Awareness enhancement. In: Proceedings of the Human Factors Society 32nd Annual Meeting, Santa Monica, pp. 97–101. Human Factors Society (1988)
5. Dominguez, C., Vidulich, M., Vogel, E., McMillan, G.: Situation awareness: papers and annotated bibliography. Armstrong Laboratory, Human System Center, Ref. AL/CF-TR-1994-0085 (1994)

Comparison of Mutual Awareness in Analog Vs. Digital Control Rooms

Kateryna Savchenko(✉), Heather Medema, Ronald Boring,
and Thomas Ulrich

Idaho National Laboratory, P.O. Box 1625, Idaho Falls, ID 83415, USA
{kateryna.savchenko,heather.medema,ronald.boring,
thomas.ulrich}@inl.gov

Abstract. Control rooms in nuclear power plants are complex, collaborative working environments rife with potential for human error. As control rooms evolve from analog to digital interfaces, crew communication strategies must change as well. With the increase in automation and the use of digital HMIs, operators no longer rely on large annunciator panels, but have instead moved to personal sitting workstations. The technology shift causes operators to focus only on their screens, reducing interaction between crewmembers. Therefore, the collaboration and coordination of task demands requires greater attention to communication, vigilance and mutual awareness, or collective knowledge of the situation. This paper will investigate, through literature review and expert interviews the impact of the technology shift and identify significant and critical improvements that can be implemented in the main control rooms to increase safety and reliability.

Keywords: Nuclear power plants · Main control rooms · Human-Computer interaction · Situation awareness · Mutual awareness · Teamwork

1 Introduction

Nuclear power plant (NPP) main control rooms (MCRs) are complex, dynamic environments. As such, operators must monitor and control complex systems while under stressful conditions with the potential for severe consequences if performed incorrectly [1]. Therefore, teamwork culture serves a vital role in enabling this complex monitoring and control in a safe manner. Work environments, such as NPP MCRs, organize teams hierarchically and assign specific roles and functions to support successful integration, synthesis and sharing of information [2].

The majority of work within the MCR centers on procedures carried out by two reactor operators (ROs) overseen by one senior reactor operator (SRO). ROs primarily are tasked with maintaining safe and correct plant operations, optimizing parameters, coordinating the functioning of the reactor and its systems, and detecting and reacting to plant deviations from the normal conditions and states. SROs are responsible for the safe manipulation of the controls of a nuclear reactor and give direction of ROs to manipulate these controls. Furthermore, they are required to effectively plan, maintain,

© Springer International Publishing AG (outside the USA) 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_19

and supervise efficient plant operations in the MCR, as well as direct and implement emergency operating procedures and event reporting [2, 3].

SROs and ROs often experience high levels of workload in such a complex environment where they have to cope with multiple information sources, performance pressure, and changing scenarios. Furthermore, this high workload strains the operators' ability to perform multiple activities in parallel, which can lead to various problems such as inaccurate communication.

Management of the NPP during normal and abnormal situations requires effective team communication and coordination. Standardization of communication and support technology is critical for a number of reasons (1) to reduce the potential impact of human errors attributed to the increasing complexity of new digital main control rooms, (2) to balance the crew demands and available resources to keep mutual awareness between SROs and ROs at a high level and, (3) to prevent the immediate and severe consequences of poor team performance [4].

This paper will investigate, through a literature review and expert interviews, the impacts resulting from the increasing complexity of new digital main control rooms. Comparison of the new digital interfaces with existing analog interfaces in NPPs has the potential to shed light on the lack of mutual awareness and communication breakdowns between the SROs and ROs in the MCRs. This paper describes the concept of teamwork and teams in terms of situation awareness and mutual awareness within crews in MCRs. Additionally, we will compare mutual awareness in analog and digital control rooms and discuss ways to mitigate the potential decrease in mutual awareness suffered in crews operating digital control rooms.

While this paper highlights potential breakdowns in communication and mutual awareness, it is important to note that a simple failure of this sort does not automatically equate to an unsafe condition at an NPP. Plants entail multiple redundant safety systems, processes, and second-checking staff. Where breakdowns occur, for example, in threeway communication, these issues are typically identified on the spot by fellow crew members and remedied. However, many of the safeguard processes to prevent mishaps may need to be revisited in light of new technologies in the MCR. Where there is the potential for established and trusted safeguard processes to fail in new contexts, these must be identified and mitigated.

2 Situation Awareness

Situation awareness is central to achieving optimal performance in the MCR. Endsley [5] defined Situation Awareness (SA) as "the primary basis for subsequent decision making and performance in the operation of complex, dynamic systems…" At the lowest level of SA, the operator takes in relevant information from the environment, the system, and self. Within the mid-level stage, information is integrated relative to task goals. Finally, in the highest levels of SA, the operator uses information gathered to predict future events as well as system states. In a dynamic environment such as an NPP MCR, seeing a big picture could reduce risks by identifying potentially problematic situations. Situation awareness operates at both the individual and at the team level. Operators must first build SA individually from the information they have immediately available to them. Then the team can form situation awareness at a higher level by communicating aspects from each individuals SA into a collective team SA.

Team SA is shared SA about plant conditions. In contrast, mutual awareness is the crews' awareness of each other. The next section will discuss in more detail the concept of mutual awareness within MCR operating crews.

3 Mutual Awareness

For a team to operate effectively it is important that the whereabouts, actions, and intentions of each individual can be communicated to other team members so that mutual awareness can be created. Mutual awareness (MA) can be defined as "an awareness of the state of a cooperative effort [6]". It is the knowledge of what other team members are doing, how they are doing it, as well as how and when they can affect each other's activities. Mutual awareness is important for individuals to be informed about each other's actions throughout the task. Team members try to fine-tune their own activities to provide their colleagues with cues about their intentions and other different kinds of information relevant to their activities. This awareness is established through various actions of individuals and is affected by the type of action, the number of actions, and when the actions occurred. Mutual awareness can be maintained through the oral exchange of the information, gestures, body language, and artifacts [7]. Furthermore, the crew can access mutual awareness through vision, sound, odors, vibrations, touch, and movement present in the shared work environment [6].

4 Changes in Mutual Awareness Due to Technology Shift

Interviews with five subject matter experts on control rooms highlighted a number of concerns regarding mutual awareness in the face of changing technology. As MCR technology shifts from analog displays to digital interfaces, significant changes to MA become apparent due to the differences between the two information display formats. In the traditional MCR, the large analog control panels span large areas primarily due to the sheer size of the instruments themselves. These larger panels naturally afford working as a team, simply because the information is distributed in many places. Operators each check individual values and relay these values back to the SRO to build the SA. This distribution of information assists the operators in understanding how their small aspect of the system relates to the larger operation of the plant and makes them aware of each other's activities and actions [7].

Control panels are associated with specific plant functions and components, i.e., the reactor, main steam and steam generators, and are each housed on their own control panel. These analog boards also contain arrays of indicators, buttons, and controls with which ROs and SROs monitor and control to support the functions for each plant system [2]. Each panel has a larger prominent array of annunciators that provide contextual cues to help operators quickly orient themselves to the functions occurring within each control panel. This analog representation also enables multiple operators to simultaneously see all the information at-a-glance and mentally integrate information across the boards to build a comprehensive high-level representation of the plant. Analog interfaces offer the advantage of a tactile feedback from pushing the buttons and using the controls only accessible on specific boards. Furthermore, instant immediate tactile, audio, and visual feedback is available, including click sounds and visual changes in positioning of the physical control. Thus, the operators' mental model and expectations of how equipment and controls are supposed to operate are supported with physical mapping of plant components to controls and indicators arranged along the control panels.

Mutual awareness is enhanced through the salient and visible cues of the operators themselves physically positioned around the analog MCR [2]. Therefore, ROs and SRO are able to coordinate and align their own activities with each other. For example, seeing ROs at the certain control board provides cues to the SRO about plans and procedures they are engaging in at the moment, thus confirming or not the correct actions of the ROs as well as expectations of the SRO [8].

New digital control rooms introduce a variety of new technologies, increased automation, and employment of graphical interfaces [2]. Operators typically sit at separate workstations with a digital control system where they have different small displays which enable them to navigate and operate the plant [7]. Digital interfaces are more compact, flexible, and configurable to particular tasking of the operator including useful trending displays and overall consistency with the design of indicators and controls. Furthermore, information about the whole system can be available at any location of the control room. Operators focus on their own screens and can perform their activities autonomously [9]. System functions are increasingly allocated to the automated computer controller, which moves the operator to the position of supervisors [10]. MA is expected to be supported by human-system interfaces (HSIs) with integrated information and a common overview display. However, separate workstations and inability to share information between ROs as well as the SRO creates an opportunity for error due to the lack of communication [11]. The SRO may lose the ability to supervise the activities of ROs and becomes a passive observer with higher workload due to higher responsibility [12]. Thus MA is maintained through the oral exchange of information that requires more increased communication frequency [7, 11, 13].

5 Communication Issues Due to Technology Shift

Despite the many benefits of automation, there are noted ramifications as well. A number of critical HSI factors must be considered as MCR technology shifts from conventional, analog displays to new digital instrumentation and control (I&C) systems. There are a number of human factors concepts that arise with the transition and potentially affect SA and MA.

In new digital control rooms, operators remain in shared space, yet the means by which operators obtain and share critical plant information has critically changed. Rather than visualizing plant system data and status via large wall hung annunciator panels and other plant control panels, information is now available via digital interfaces. These visual display units (VDU) bring benefits for safety and flexibility by combining information from various sources [14]. A great breadth of information may be concentrated within one interface and may be difficult to sort through for the desired piece of information or tool. This phenomenon, coined the Keyhole Effect [15, 16], was named to describe the idea of "peeping through a keyhole to find the relevant piece of data in a room full of possibilities behind the closed door [17]." A keyhole effect in an MCR creates difficulty in fully understanding the system in question, because the operator is only afforded partial system information at any given time and must search for additional information within the display.

Another human factors concept is operator-out-of-the-loop performance problems. The term *operator out of the loop* (OOTL) refers to both system performance and human performance problems that arise as the MCR operator struggles to maintain situation awareness [18]. Operator reliance on the automation leads to a number of potential, critical missteps including lack or loss of adequate skills, a shift from an active to passive state in information processing, a change in the feedback provided to the operator, and finally to vigilance and complacency obstacles for accurate situation awareness. These ultimately lead to an inability to take back the reins from the automation and operate the plant in manual mode should automation failure occur [19].

Table 1. Summary of experts' interviews.

	Issue	Solution
Teamwork	Failures of team coordination cause disruptions in plant operations	Proper training, peer checking and briefings to improve effective communication is needed
Mutual awareness	Team members are not aware of each other's activities and intentions	Increased frequency of communication and exchange of information is needed
Communication	Lack of MA causes breakdowns in three-way protocol	Operators can increase the use of briefings and peer checking as well as enlist in additional training to effectively communicate
Control room environment	Technology shift causes operators to perform plant operations in isolation, experiencing overload of information	Regular peer checking, briefings, proper training and use of mimic screen can help to exchange information

6 Conclusions

In this paper, we've explored the field of research concerning the potential negative impact of changes in MA between SROs and ROs with technology shifts in the MCR of NPPs. Table 1 summarizes several of the issues found in interviews with subject matter experts. Team collaboration and coordination play an important role in digital operating systems as well as in causing incidents. With the use of fixed analog control boards, ROs and SROs could see the big picture of the plant, but digital systems may place the operators at spread out workstations. ROs may be less able to share information with each other and maintain high MA. Below are the ways we propose to combat the lack of MA.

- *Training*. The procedural and strategic knowledge of the system benefits operators in providing more time and attention resources to communicate with each other as well as understanding how the system works [7, 8]. The knowledge can be gained through the proper training. Causes of human error in the nuclear industry can be a lack of proper training and understanding of operational procedures [20–22]. The nuclear industry needs to train operators to reduce the potential knowledge and experience gap during the emerging technology shift in the MCR. Every operator must be able to work productively and safely around other operators to share the critical information and increase MA. Training on plant's procedures, capabilities, vulnerabilities, limitations, and hierarchy of priorities is necessary to achieve that goal.
- *Staff briefings and peer checking*. Briefings of operators can support high levels of MA as well as peer checking (PC). Briefings are good to get everyone back on the same page, and to promote and ensure the coordinated effort in accomplishing the goal of safe operations of the NPP [23]. During the PC, two people self-check in parallel and agree on the correct action to be carried out on the appropriate component. Action is the primary focus of PC. This also brings a fresh set of eyes to the issue. The peer who has knowledge and familiarity with the activity and intended results might notice the hazard that the performer cannot see [24].
- *Mimic screen.* Redesign of the display interfaces can improve MA and support communication between operators under task conditions as well. Adding a mutual awareness tool (MAT) or "mimic screen" to the digital interfaces of a NPP could improve crew's MA [7, 25]. The tool would "mimic" what ROs are doing, which procedure and step in the procedure they are using, and systems each team member is on. The SRO would have the means to monitor and verify the parallel activities of ROs. Less time would be spent on the discussion and exchange of information promoting operators working in isolation.

Any design flaws have an opportunity to bury critical information and create overload of information for the ROs and SRO. This is the first of a series of papers exploring MA and investigating means for effectively reducing issues linked to transitions of technology in the MCR. In terms of future research, a thorough development and evaluation of the new technologies in terms of potential negative and positive effects is needed to ensure the appropriate application in the control room to support team performance, MA, and plant safety [2]. Further, the researchers will expand the knowledge base by gathering and evaluating input from actual operator users.

Acknowledgements. This work of authorship was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Idaho National Laboratory is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DEAC07-05ID14517.

References

1. Paris, C.R., Salas, E., Cannon-Bowers, J.A.: Teamwork in multi-person systems: A review and analysis. Ergonomics **43**, 1052–1075 (2000)
2. Wright, M., Kaber, D.: Effects of automation of information-processing functions on teamwork. Hum. Factors **47**(1), 50–66 (2005)
3. Sage, A., Rouse, W.B.: Handbook of Systems Engineering and Management. Wiley, New York (1999)
4. Salas, E.: On teams, teamwork, and team performance: discoveries and de-velopments. Hum. Factors **50**(3), 540–547 (2008)
5. Endsley, M.R.: Measurement of situation awareness in dynamic systems. Hum. Factors **37**, 65–84 (1995)
6. Schmidt, K.: Some Notes on Mutual Awareness. COTCOS Awareness Sig Work-shop, Paris (1998)
7. Yenn, T.-C., Lin, C.J., Jou, Y.-T., Hsieh, T.-L., Yang, C.-W.: The effect of human information processing and workload on staffing level in the main control room. Paper presented at the 6th Nuclear Plant Instrumentation, Controls, and Human Machine Interface Technology (NPIC&HMIT), Knoxville, April 2009
8. Juhász, M., Soós, J.K.: Human aspects of NPP operator teamwork, nuclear power - control, reliability and human factors. In: Tsvetkov, P. (ed.) InTech (2011). doi:10.5772/17046
9. Lin, C.J., Yenn, T.-C., Yang, C.-W.: Automation design in advanced control rooms of the modernized nuclear power plants. Saf. Sci. **48**(1), 63–71 (2010)
10. Nachreiner, F., Nickel, P., Meyer, I.: Human factors in process control systems: The design of human-machine interfaces. Saf. Sci. **44**, 5–26 (2006)
11. Min, D.H., Chung, Y.H., Yoon, W.C.: Comparative analysis of communication at main control rooms of nuclear power plants. In: Proceedings of the 9th IFAC/IFIP/ IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human–Machine Systems, Atlanta (2004)
12. Liu, P., Li, Z.: Comparison between conventional and digital nuclear power plant main control rooms: a task complexity perspective, Part II: detailed results and analysis. Int. J. Ind. Ergon. **51**, 10–20 (2016)
13. Yuan, X., She, M., Li, Z., Zhang, Y., Wu, X.: Mutual awareness: enhanced by interface design and improving team performance in incident diagnosis under computerized working environment. Int. J. Ind. Ergon. **54**, 65–72 (2016)
14. Wahlström, B.: Differences between analog and digital I&C. In: Proceedings of the 9th International Conference on Nuclear Plant Instrumentation, Control & Human–Machine Interface Technologies (NPIC & HMIT 2015) (2015)
15. Woods, D.D.: Visual momentum: a concept to improve the cognitive coupling of person and compter. Int. J. Man Mach. Stud. **21**, 229–244 (1984)
16. Woods, D.D.: Toward a theoretical base for representation design in the computer medium: ecological perception and aiding human cognition. In: J. Flach (1995)
17. Dekker, S.: Patient Safety: A Human Factors Approach. CRC Press, Taylor and Francis Group, Boca Raton (2016)
18. Kaber, D., Endsley, M.: Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. Process Saf. Prog. **16**(3), 126–131 (1997)
19. Endsley, M.R., Kiris, E.O.: The out-of-the-loop performance problem and level of control in automation. Hum. Factors J. Hum. Factors Ergon. Soc. **37**(2), 381–394 (1995)

20. Flin, R., O'Connor, P., Mearns, K.: Crew resource management: improving teamwork in high reliability industries. Team Perform. Manag. **8**(3/4), 68–78 (2002)
21. Harrington, D., Kello, J.: Systematic evaluation of nuclear operator team skills training paper presented at the American Nuclear Society, San Francisco, November 1993
22. Rich, V.: Training of operators needed for nuclear power station safety. Nature **329**(6140), 572 (1987)
23. Tehrani, M., Molesworth, B.R.: Pre-flight safety briefings, mood and information retention. Appl. Ergon. **51**, 377–382 (2015)
24. Department of Energy.: Human Performance Improvement Handbook, vol. 1 of 2. DOE-HDBK-1028 (2009)
25. Nova, N., Wehrle, T., Goslin, J., Bourquin, Y., Dillenbourg, P.: Collaboration in a multi-user game: impacts of an awareness tool on mutual modeling. Multimedia Tools and Appl. **32**(2), 161–183 (2007)
26. Yang, C.-W., Yenn, T.-C., Lin, C.J.: Team error analysis in a digitalized control room by using a novel crew reliability model. Hum. Factors Man. **21**, 429–439 (2011). doi:10.1002/hfm.20318

Characteristics Analysis for the Effect of Mental Fatigue in Monitoring Task

Yiqi Ai, Qianxiang Zhou(✉), Yun Tian, and Zhongqi Liu

Key Laboratory for Biomechanics and Mechanobiology
of the Ministry of Education, School of Biology and Medical Engineering,
Beihang University, Beijing, China
aiyiqijiuxiao@163.com, zqxg@buaa.edu.cn

Abstract. In general, mental fatigue refers to the phenomenon of declining ability of human perception, memory, attention and other cognitive caused by insufficient sleep, biological rhythm disorders or long time focusing. The main research work of this paper is as follows: Make a research on the reasonable way to induce mental fatigue. The volunteers in fatigue group continuously operate 140 min number "2-back" task, while the volunteers in normal group only operate "2-back" task at the first and last 10 min. Subjects' task performance and ERP component are compared between the first 10 min and the last 10 min. The dynamic change of task performance and ERP component was studied. The results confirmed that 140 min "2-back" task successfully induces mental fatigue.

Keywords: Mental fatigue · 2-Back task · Monitoring operation · ERP

1 Introduction

Human fatigue can be divided into two physical fatigue and mental fatigue. Mental fatigue refers to brain dysfunction leading by nerve cell fatigue [1]. Mental fatigue is a psychological - physiological discomfort state people feeling subjectively and objectively after high-intensity or long-term mental activity [2, 3]. Fatigue is a gradual accumulation process that is common in modern production and daily life. Mental fatigue people usually is unable to complete self-motivation and internal driving force of mental work [4].

There are many factors that cause mental fatigue, the mainly two factors are objective factors and subjective factors. Bad work environment, difficult task or few rest time [5], lacking of sleep and other objective factors all can destroy the physiological and psychological ability and have fatigue effect. With the emergence of mental fatigue, usually people will lack interest of a job or task, have negative emotions and fell tired. Negative subjective state, in turn, has a negative impact on mental state, resulting in cognitive task performance decreased [6, 7].

Common methods of inducing mental fatigue include sleep deprivation, continuous cognitive manipulation and other methods. Here we are discussing the continuous cognitive manipulation. The N-back paradigm was first proposed by Kirchner to analyze age differences in short-term memory [8]. The most common paradigm for the

© Springer International Publishing AG 2018

R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_20

study of neural mechanisms in working memory is the N-back paradigm, since the N-back task requires the subjects to monitor and update the working memory contently in real time [9]. In recent years, N-back paradigm has been widely used in the study of mental workload and mental fatigue. When N = 2, the current stimulus is compared with the last but one stimulus. Due to much occupation of the memory capacity, 2-back task often used for moderate load experiments [10].

The formation of mental fatigue is multiple, and its manifestations are diverse, such as drowsiness, fatigue, workload, mood changes, etc. The main tools for the evaluation and prediction of mental fatigue are as follows: (1) subjective evaluation method (2) based on the performance of the task evaluation method (3) electrophysiological evaluation method (4) eye movement evaluation method. In this paper, we used the performance data and EEG data to evaluate mental fatigue [11].

Studies have shown that the “2-back” task can induce mental fatigue, but there is no clear definition for upper and lower limits of task duration. Discussion of the mental state changes during “2-back” operation process are insufficiency. Therefore, this paper discusses the relationship between “2-back” task and mental fatigue by performance data and EEG data.

2 Material and Method

2.1 Subjects

A total of 32 males voluntarily participated in this experiment. Their ages ranged from 20 to 30 years old, bachelor Degree or above, right handed. They all have normal or corrected to normal visual acuity and have no sensitivity difference in perception of color red and green. Volunteers’ basic cognitive ability is normal and have good sleep habit. They don’t smoke or addict to alcohol. 4 h before the experiment began, volunteers cannot intake of alcohol or caffeine. None of them have had any experience with this type of experiment.

2.2 Mental Fatigue Induced Method

The digital “2-back” task program is wrote by E-Prime2.0 (Fig. 1) [12–14]. Number 0 to 9 shows in the central of the screen randomly. The background is black and the number is white. Each stimuli display 500 ms. Time interval between two stimuli changes within the range of 1500 ms to 3000 ms randomly. The ratio between target stimuli to distractive stimuli is 1:2, presented in pseudo-random order [15]. The subjects judge whether the current number was the same as the last second number. If the two numbers is the same, press f key by left index finger. If not, then press j key by right index finger [16].

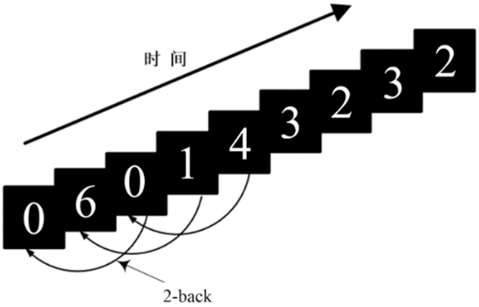


Fig. 1. “2-back” task

3 Conclusions of Mental Fatigue Induced Method

3.1 Performance Data

The paired samples T test was performed to test the performance data of the “2-back” pretest and posttest tasks in the normal group. The reaction time difference between pretest and posttest tasks has no statistically significant meaning ($t = 1.710$, $p = 0.098$). The difference between the stability of the pretest and posttest tasks has no statistically significant meaning ($t = 1.193$, $p = 0.242$). There is no statistically significant meaning of the difference between the error rate of pretest and posttest tasks ($t = 0.170$, $p = 0.866$). There is no statistically significant meaning of the difference between the omission rate of pretest and posttest tasks ($t = -0.269$, $p = 0.789$). The performance data proves that the 20 min “2-back” task doesn’t result in mental fatigue. After mental fatigue induction period, the subjects in normal group have normal mental status.

Divided operation of fatigue group into 14 segments averagely (pre, block1 ~ 12, post, each segments last 10 min). Figure 2 shows the change of each segment’s performance data with time passing by (Time-On-Task).

The overall task error rate shows an upward trend. The last 20 min (block12 ~ post) in the task’s reaction time decreased, error rate decreased, leakage rate decreased. It is because that the subjects become excited at the end of the “2-back” task, which is causing by subjective adjustment. The overall trend of “2-back” mission stability is similar to the overall trend of response time. From the performance point of view, the 140 min “2-back” task induced mental fatigue successfully.

Figure 3 shows the post-error reaction time, post-correct reaction time and the slowing of each fatigue group segment. According to Fig. 3, subjects in fatigue group’s reaction time slowed down after an error occurs in all segments. The slowing value of “2-back” posttest task is significantly less than the slowing value of the pretest task (paired sample t test, $t = 2.852$, $p = 0.012$).

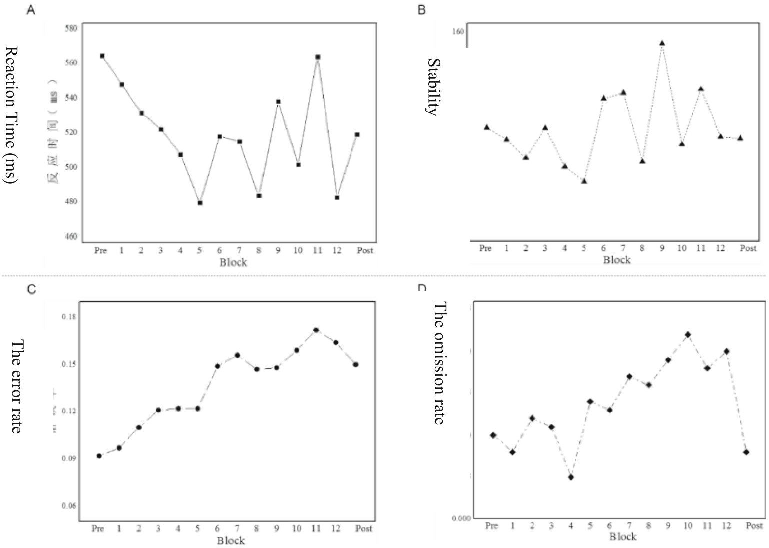


Fig. 2. Performance data of each "2-back" task segment in the fatigue group

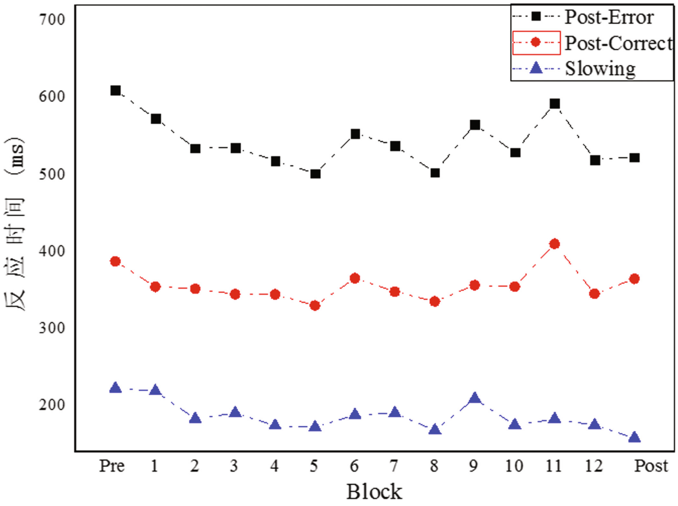


Fig. 3. The post-error slowing effect of the "2-back" task in the fatigue group

3.2 EEG Data

Descriptive statistics and repeated measures ANOVA were used to analyze the P300 amplitude, latency of normal group subjects of "2-back" pretest and posttest tasks. The main effect of normal group subjects' P300 amplitude is not significant in the pretest and posttest (F = 0.055, p = 0.818). The main effect is not significant on the three leads

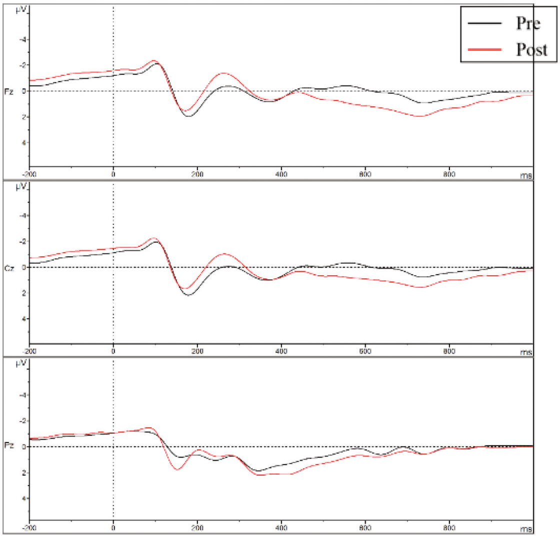


Fig. 4. P300 waveform of pretest and posttest task in normal group

Fz, Cz, Pz (F = 2.611, p = 0.123). And there is no significant interaction between time and lead (F = 0.132, p = 0.732).

Before and after mental fatigue induction task, there is no significant change in P300 amplitude and latency in normal subjects ($p > 0.05$). EEG data can prove that 20 min 2-back task does not cause mental fatigue. Brain wave pattern (Fig. 4) can directly reflect the above analysis' results of P300 amplitude and latency of the normal group subjects (Fig. 5).

Using descriptive statistics and repeated measures ANOVA to analyze the P300 amplitude, latency of fatigue group subjects of "2-back" pretest and posttest tasks. The main effect of normal group subjects' P300 amplitude is not significant in the pretest and posttest (F = 0.999, p = 0.333). The main effect is not significant on the three leads Fz, Cz, Pz (F = 0.902, p = 0.359). Interaction effect of time and lead is remarkable (F = 8.132, p = 0.009). Simple effect analysis shows the effect of time factor on Fz, Cz and Pz lead is not significant (Fz: F = 1.770, p = 0.203; Cz: F = 1.630, p = 0.221; Pz: F = 0.000, p = 0.966). The result shows that time factor is not affected by lead factor. On the pretest the effect of lead is not significant (F = 1.410, p = 0.259) while on the posttest the effect is significant (F = 3.430, p = 0.046). This result prove that lead factor is influenced by time factor. P300 latency's main effect is significant in the pretest and posttest (F = 14.042, p = 0.002) while is not significant in the Fz, Cz, Pz lead (F = 0.262, p = 0.617), and there is no significant interaction between time and lead (F = 2.469, p = 0.137).

With "2-back" task time extended, subjects in fatigue group's P300 amplitude had no significant change ($p > 0.05$) and the incubation period extended significantly ($p < 0.05$). EEG data proved that 140 min "2-back" task induced mental fatigue

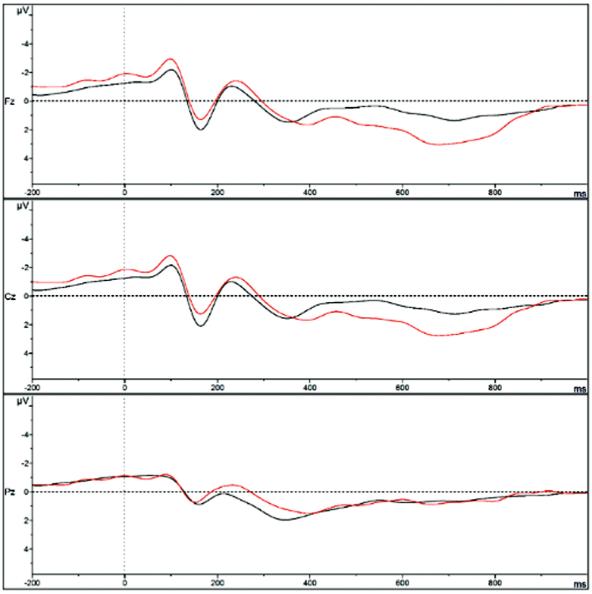


Fig. 5. P300 waveform of pretest and posttest task in fatigue group

successfully. Brain wave pattern (Fig. 6) directly reflect the above analysis' results of P300 amplitude and latency of the fatigue group subjects.

Divided operation of fatigue group into 14 segments averagely (pre, block1 ~ 12, post, each segments last 10 min). Figures 6 and 7 display each segment's performance data's change with time (Time-On-Task).

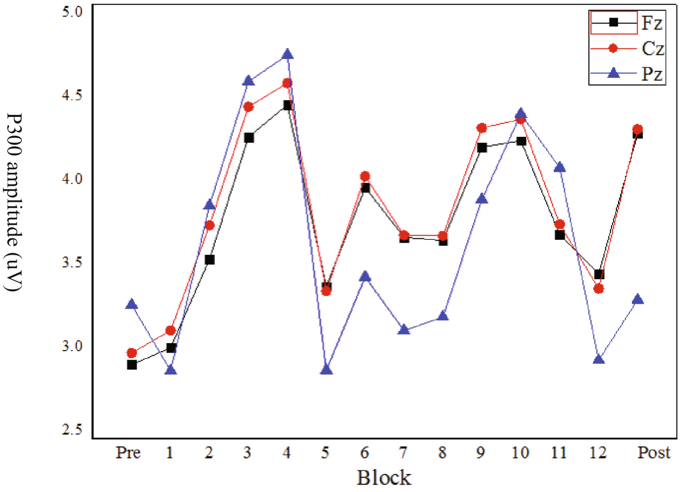


Fig. 6. "2-back" task fatigue group's P300 amplitude in each segment.

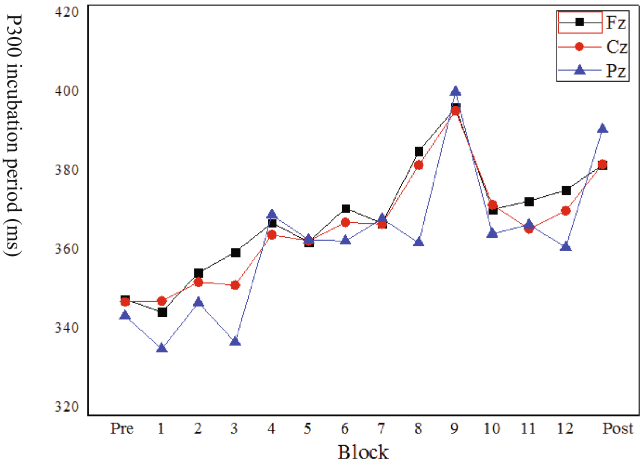


Fig. 7. “2-back” task fatigue group’s incubation period in each segment.

During the first 50 min (pre ~ block4) of “2-back” task, P300 amplitude and P300 incubation period’s rising trend indicated that subjects’ fatigue degree was increasing. During block5, P300 amplitude declined suddenly, which means operation of the subjects was sloppy. In the last 80 min (block6 ~ post), P300 amplitude fluctuated greatly which is due to subjects’ self-motivation and adjustment. The length of P300 incubation period was increasing, which can indicative subjects were still mentally fatigued during this time. Above all, EEG further proves that 140 min “2-back” task successfully induced mental fatigue.

Figure 8 displays scatter plot and trend line of ERN amplitude of each segments in fatigue group (Fz lead). With task time increasing, absolute value of ERN amplitude decreases. That shows that the behavior monitoring ability and error cognitive ability of the subjects decreased. The fatigue group subjects’ ERN amplitude in the posttest task was significantly less than that in the pretest task (Fz: t = −8.014, p = 0.000) (Fig. 9).

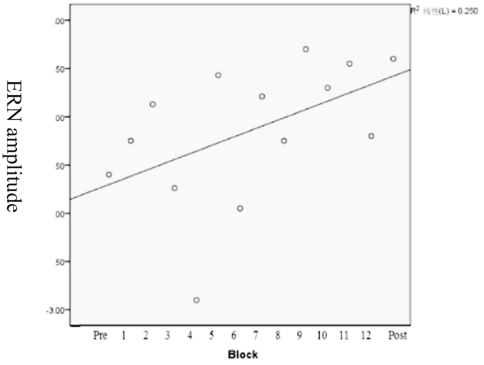


Fig. 8. Scatter plot and trend line of ERN amplitude of each segments in fatigue group (Fz lead)

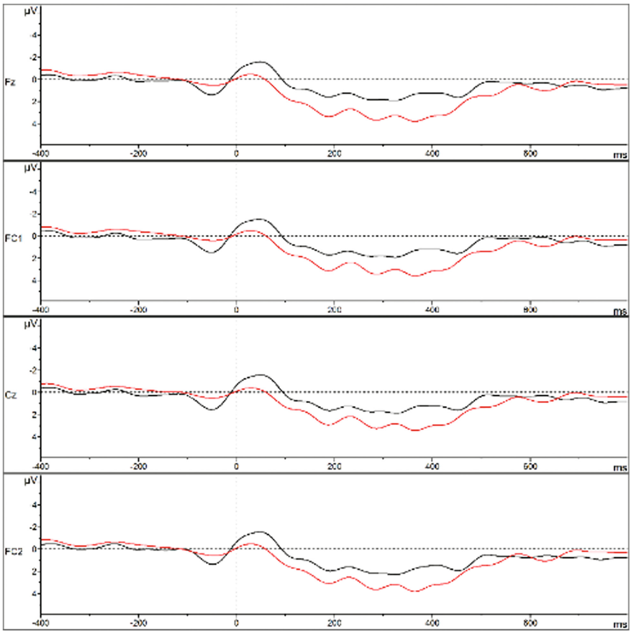


Fig. 9. The ERN waveform of the "2-back" pretest and posttest tasks in the fatigue group

4 Conclusion

140 min "2-back" task successfully induced mental fatigue. From the analysis of performance data and P300 data of the 140 min "2-back" task of the fatigue group, till the 60 min subjects reached the mental fatigue state, with the task time pass by, the deepening of mental fatigue degree. In the last 20 min of "2-back" task, due to the subjective adjustment of the subjects, the performance data has a good trend, but the whole is still in mental fatigue state. The dynamic change of task performance and ERP component was studied. The results confirmed that 140 min "2-back" task successfully induces mental fatigue.

Acknowledgements. This research was funded by Electronic information equipment system research of Key laboratory of basic research projects of national defense technology (DXZT-JC-ZZ-2015-016).

References

1. Mizuno, K., Tanaka, M., Yamaguti, K., et al.: Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity. J. Behav. Brain Functions **7**(15), 1–7 (2001)
2. Chaudhuri, A., Behan, P.O.: Fatigue in neurological disorders. J. Lancet **363**(9413), 978–988 (2004)

3. Hansson, E., Rönnbäck, L.: Altered neuronal–glial signaling in glutamatergic transmission as a unifying mechanism in chronic pain and mental fatigue. J. Neurochem. Res. **29**(5), 989–996 (2004)
4. Staud, R.: Peripheral and central mechanisms of fatigue in inflammatory and noninflammatory rheumatic diseases. J. Curr. Rheumatol. Rep. **14**(6), 539–548 (2012)
5. Stasi, L.L.D., Mccamy, M.B., Catena, A., et al.: Microsaccade and drift dynamics reflect mental fatigue. J. Eur. Neurosci. **38**(3), 2389–2398 (2013)
6. Lal, S.K.L., Henderson, R.J., Carter, N., et al.: Effect of feedback signal and psychological characteristics on blood pressure self-manipulation capability. J. Psychophysiol. **35**(4), 405–412 (1998)
7. Heller, W., Nitschke, J.B., Etienne, M.A., et al.: Patterns of regional brain activity differentiate types of anxiety. J. Abnorm. Psychol. **106**(3), 376–385 (1997)
8. Kirchner, W.K.: Age differences in short-term retention of rapidly changing information. J. Exp. Psychol. **55**(4), 352–358 (1958)
9. Jansma, J.M., Ramsey, N.F., Coppola, R., et al.: Specific versus nonspecific brain activity in a parametric N-Back task. J. Neuroimage **12**(6), 688–697 (2000)
10. Crevits, L., Simons, B., Wildenbeest, J.: Effect of sleep deprivation on saccades and eyelid blinking. J. Eur. Neurol. **50**(3), 176–180 (2003)
11. Rabbitt, D.P.: How old and young subjects monitor and control responses for accuracy and speed. J. Br. Psychol. **70**(2), 305–311 (2011)
12. Botvinick, M.M., Braver, T.S., Barch, D.M., et al.: Conflict monitoring and cognitive control. J. Psychol. Rev. **108**(3), 624–652 (2001)
13. Dutilh, G., Ravenzwaaij, D.V., Nieuwenhuis, S., et al.: How to measure post-error slowing: a confound and a simple solution. J. Math. Psychol. **56**(3), 208–216 (2012)
14. Coles, M.G.H., Scheffers, M.K., Holroyd, C.B.: Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. J. Biol. Psychol. **56**(3), 173–189 (2001)
15. Endrass, T., Franke, C., Kathmann, N.: Error awareness in a saccade countermanding task. J. Psychophysiol. **19**(4), 275–280 (2005)
16. Mathewson, K.J., Dywan, J., Segalowitz, S.J.: Brain bases of error-related ERPs as influenced by age and task. J. Biol. Psychol. **70**(2), 88–104 (2005)

Evaluation of Decision Making Processes in Critical Situations

Lisa Fromm(✉), Gundula Zerbes, Christian Plegge,
and Thomas Alexander

Fraunhofer Institute for Communication, Information Processing
and Ergonomics FKIE, Zanderstr. 5, 53177 Bonn, Germany
`{lisa.fromm,gundula.zerbes,christian.plegge,`
`thomas.alexander}@fkie.fraunhofer.de`

Abstract. Emergency services, policemen and members of the armed forces act in high workload environments where fast and correct decisions are essential. Therefore, decision criterions and automated processes play an important role in drill and training. The evaluation of the trainees' performance is usually based on subjective ratings by the instructors, which depend on the availability of many instructors to provide individual feedback. The goal of our ongoing research work is to develop evaluation techniques that enable objective ratings for the trainees' performance. As gaze behavior is seen as key element of trainees' performance, tracing methods are evaluated and head-mounted eye tracking is ascertained as promising. Furthermore, we give an overview about ongoing work, including software (laboratory experiment) and hardware development (field experiment) of an objective rating tool.

Keywords: Critical situations · Training · Decision making · Gaze behavior · Eye tracking

1 Introduction

In their highly demanding work environment, emergency services are frequently exposed to critical situations with high cognitive and physical workload. In these environments, fast and correct decisions are essential. Because of the limited capacity and time to reflect on the situation and judge the possible consequences, rule-based decision making processes and underlying decision criterions play an important role in drill and training. In this context, the following questions are of specific interest: Which are the critical characteristics of a situation and how can they be perceived and processed in an accurate manner? In training sessions, instructors tend to teach processes, motion schemata and gaze behavior until they are automated. The evaluation of such training sessions still remains difficult. In most cases, subjective ratings and judgements of the instructors are the only way to evaluate training effectiveness and efficiency. But for a more constructive and detailed feedback, instructors have to be appropriately equipped with measurement systems and supported by subsequent evaluation techniques that enable objective ratings of their trainees' performance.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_21

Different methods are available to trace human decision making. An overview of existing methods will be given in this article. These methods are reviewed in terms of their practical applicability for the evaluation of decision making processes in critical situations.

One of the most frequently used methods to track decision making processes, e.g. in critical driver's behavior research [1–3], is eye tracking. Because of the importance of gaze behavior in dynamic, critical situations and the potential for its application during training sessions, we primarily address this method. We will refer to ergonomic criteria, possibilities and technical limitations to analyze the applicability of this method to track decision making processes in safety-critical situations. Among other parameters, we tested technical opportunities and limitations, such as marker detection, reach of data recording, user-friendly integration in the forces' equipment, or the operational effort and benefit of the system. Based on these findings, we will present a concept of a further study to evaluate our results and recommendations.

2 Tracing Decision Making Processes

The term "decision making" is used in different ways: In the narrower sense, decision making describes the process of selecting the best option out of different alternatives. In deployments of emergency forces, the cognitive process called judgement is most important. Judgement means the classification of an object [4]. Hastie and Dawes [5] define judgement as "the human ability to infer, estimate and predict the character of unknown events". Judgments can be categorized in three ways; statements can be either comparative ("X is more agreeable than Y"), quantitative ("X is very agreeable"), or categorical ("X is agreeable"). Relative judgements, estimation, and classification have been used as response formats to investigate judgement strategies [6]. Regarding emergency forces, all three types of judgements are relevant, e.g. the decision between different options of entering a building, the estimation of the potential risk of smoldering fire, or the classification whether an object, a person, or a situation is threatening or not.

Three questions have been most important in research on human judgment. First, how are objective magnitudes (e.g., loudness of a tone or the brightness of a light) mapped onto subjective, internal scales? Researchers in this field search for regularities which characterize this mapping. Second, how accurate are judgements and how can they be improved? And third, which are the underlying cognitive processes of judgement? Concerning the conception of the mind as an information processor, researchers are interested in modeling how different judgmental cues are searched, processed, and combined [6].

Several methods and techniques exist for tracing decision making, such as verbal protocols, information search displays, and eye movement monitoring. These methods are used to identify and track psychological events that occur prior to the response, such as cognitive states, stages, or processes. Several advantages of eye movement recordings in contrast to verbal protocols or information search displays were identified: They capture a broader range of information sampling acts (executed with or without conscious awareness); they can be used supplementary when having concurrent verbal protocols or serve as powerful cues for retrospective protocols; gaze behavior can serve as an indirect measure of the distribution of visual attention; and – in contrast to the other methods – they do not produce artificial demands which might alter the decision process (such as the execution of a manual act in the information search display paradigm) [7]. As eye movement monitoring is potentially applicable in dynamic situations, e.g. during training sessions of emergency forces, studies that used eye tracking for decision making research were reviewed in terms of their methodological approach and its possible transfer to dynamic contexts.

3 Evaluation of Eye Tracking for Critical Situations

The area of the human foveal vision is somehow constrained. To perceive objects or persons clearly, humans need to point their gaze at it [8]. While the focus of attention is not necessarily consistent with visual fixation (e.g., an auditory stimulus catches our attention while continuing the conversation with a colleague), a fixated point initially catches attention [9, 10]. Admittedly, people are able to shift their attention covertly to areas of the visual field away from their point of gaze [11], but during natural vision, direction of gaze is considered to be tightly coupled to the orientation of attention [12–14]. Therefore, eye tracking is a useful tool to visualize the shifting attention of the brain. The basis of assessing cognitive processes by use of eye tracking is the differentiation between saccades (little, fast movements of the eye) and fixations (short periods of stagnation which allow for perceiving the relevant information). For a review of eye movement measures, see [15, 16].

A number of eye tracking systems is available on the market [17]. Usually, they have been specifically developed for application in different research fields, such as driver's behavior [18], human–computer interaction and usability research [19], market research [20], etc. Stationary systems, such as remote eye tracker that are integrated into displays, are ineligible for application in dynamic training environments. Therefore, we focus on head-mounted systems. Because of its availability and former integration of the system into safety glasses, we exemplarily examined the Dikablis System by Ergoneers (Fig. 1) in terms of its applicability for highly dynamic training environments. The head section is fixed at forehead, nose bridge and back of the head, while the weight is mainly worn by nose and forehead. The infrared cameras are positioned below the eyes (Fig. 1, *left*). Other wearable parts of the system are carried in a backpack, which include storage battery, tablet for data recording and connector box (Fig. 1, *right*).

In a first study, technical potentials and limitations of the system were examined. Criterions included marker detection (required to define areas of interest for automatic analysis), range of wireless connection for data recording while simultaneously tracing the video data on a laptop, costs and benefits when learning how to run the system, as well as user-friendly integration into safety gear [21]. Marker detection and range of wireless connection for data recording are relevant for adapting the eye tracking software for usage in training sites and for adequate positioning of markers in the scenery. The maximal distance between marker and field camera of the eye tracking system is depicted in Table 1. The detection algorithms are provided by the software.

Fig. 1. Head-mounted eye tracking system (*left*) and required equipment for data recording (*right*).

Table 1. Maximal distance for marker detection.

Size of marker [cm]	50 × 50	20 × 20	15 × 15	10 × 10	5 × 5
Normal detection	11 m	8 m	7 m	5 m	2,5 m
Fast detection algorithm	11 m	12 m	10 m	7 m	5 m
Normal detection algorithm	23 m	18 m	13 m	12 m	7 m
Slow detection algorithm	40 m	20 m	18 m	13 m	7 m

The so-called slow detection algorithm unsurprisingly needs the most time to run through the video data.

In Table 2 the maximum angle of the marker axis relative to the camera is depicted. As expected, the largest angles between field camera and marker are possible when using bigger markers at higher distances. The maximum angle for functioning marker detection is 25°.

Table 2. Maximum angle of the marker relative to the camera.

Size of marker [cm]	15 × 15	10 × 10	5 × 5
Distance of 67 cm	10°	15°	20°
Distance of 250 cm	25°	25°	25°

The range of wireless connection for data recording was tested by varying the distance between router and eye tracking system. The maximal linear distance amounted to at least 55 m. Barriers, such as walls, reduced the maximal distance to 20 m, depending on the texture and build of the walls.

Results also indicated that learning the basic functions of the software is quite simple, even for novices. After a short software introduction by an expert, novices were able to calibrate the eye tracking system and record a video within eight minutes. But for being able to perform analysis without guidance of a software expert, training courses are necessary.

Moreover, the integration of the eye tracking system into the protective gear, with a focus on helmet and safety glasses, is of special interest. Further optimization is recommended regarding the comfort of the head section, especially in the area of forehead and nose. Using eye tracking in training sessions demands sufficient comfort for several hours. A possible approach to improve comfort issues might be the integration of the eye tracking components into safety glasses. This was already realized for jet aircraft pilots [22], but the combination with ballistic protection of safety glasses in use by security forces is challenging and requires further research and development. Further examination is also needed in terms of the tracking of dynamic sceneries, especially under variable lightning conditions. Some researchers seem to be on a good way to robust eye tracking under variable lightning conditions [23].

4 Future Work: Eye Tracking for Training Critical Situations

In future work, we will focus on the development of an eye tracking device to support training of emergency forces on specific training sites. A crucial requirement for applicability is the integration of such a device into common procedures in live simulation training. Usually training sequences last about 30–60 min and are followed by debriefings and preparation of the training site for the next training sequence. A promising approach for successful integration of eye tracking is to use the device during training sequences and deliver relevant information fast, so that it supports debriefing.

Further research and development on the way to an eye tracking system for training of decision making processes in critical situations is necessary in two ways: First, which kind of eye tracking data is of interest in training of critical situations and how can this data easily be provided for instructors? These questions concern the software part of further development. Second, the hardware of the eye tracking system needs to be adapted to the circumstances of dynamic training environments.

Regarding the software requirements, results of recent studies provide the basis for further work in this area. Expert interviews revealed insights about organizational conditions and benefits of eye tracking for educational purposes. A reasonable period for analyzing eye tracking data during training sessions was defined as 10 to 15 s within an exercise. The time between exercise and debriefing should not exceed 15 min. Benefits of eye tracking for educational purposes were seen in the tracking of particular sequences of an exercise and the use of this data for debriefing. Especially the detection of endangerments, including persons, objects or situational characteristics, is of high interest for the instructors [21]. Although the perception of endangerments is the basis for fast reactions, an objective evaluation tool is currently not available. To provide a technical solution for objective evaluations, we will examine gaze behavior in dynamic, critical decision making processes and evaluate eye tracking as a method to visualize gaze behavior relevant for the decision process.

4.1 Laboratory Experiment to Identify Relevant Performance Data

The first part of future work will be a laboratory experiment that indicates whether gaze behavior is a proper parameter for analyzing trainees' performance during decision making processes in critical situations. The planned laboratory experiment is based on psychological research and models of relevant cognitive processes. Presumably, the detection of endangerments requires two sequential cognitive processes: The primary process is visual search. In the context of emergency forces, this means searching for features of a situation that indicate endangerment. If a feature of a situation is already perceived, the second process called judgement follows, as defined earlier [5]. In this case, the person needs to decide whether the target is threatening or not and which kind of reaction is needed. Finally, a correct or adequate motor reaction can be initiated. The paradigm of visual search is used to study the deployment of attention. In this paradigm, the subject looks for a target item among a set of distractor items, including sceneries where a target item is missing. The analysis of reaction time reveals insights about attentional processes [24]. According to a leading model of visual search [25], visual search is divided into two subsequent processes: A first automatic, fast processing allows to perceive and examine all stimuli of the visual field superficially. Thereby, the resemblance of the stimuli with the target stimuli is tested. On the basis of resemblance, the stimuli are serially processed in a second, slower and controlled way in order to detect stimuli with target features.

The second basis for our laboratory experiment is the Diffusion Drift Model [26, 27]. It intends to explain simple and fast decision processes with two alternatives (e.g., reaction necessary or not). According to this model, three cognitive processes are necessary to reach a decision: encoding, decision and answer. Information is gathered and accumulated towards one of the possible alternatives. If the accumulated information reaches a threshold, the alternative is chosen. An integrated model, linking the visual search paradigm and the Diffusion Drift Model, could help to explain cognitive processes in decision making processes of emergency forces. First, the surroundings are scanned in terms of potential endangerments. Second, stimuli containing features of potential endangerments are further processed and a decision is made regarding the necessary reaction, e.g., to shoot or not in case of confronting an assailant. The idea of an integrated model of decision making is questioned in our laboratory experiment. The challenge is to find parameters (e.g., eye tracking data) that reflect cognitive processes during decision making processes in critical situations.

To visualize and analyze gaze behavior during visual search tasks and judgement separately, it seems reasonable to make use of a paradigm called subtraction method. Therefore, two tasks are compared with each other: The first task requires searching for a target and a fast reaction (e.g., pressing a button) if the target is detected. The second task includes searching for a target, but additionally, the subject is asked to judge the target (e.g., a categorization like dangerous vs. not dangerous). If we compare reaction times, errors, eye tracking data, etc. of both test conditions, we can separate specific processes and/or gaze behavior for judgement. Possible measurements include mean fixation duration, single fixation duration, first fixation duration, and gaze duration [15]. If eye tracking data reveals insights of judgement processes, this data could be used for evaluation of judgement processes under real-life conditions.

As mechanisms of attentional deployment differ considerably in the lab compared to the real-world [10, 28], a further step of our future research will be to examine gaze behavior of emergency forces in a dynamic, more context-specific (field) experiment. The revealed data, including evaluation regarding ergonomic criteria, can be used for the development of a software tool which helps to automatically analyze relevant eye tracking data and assists with data interpretation.

4.2 Field Experiment to Evaluate Hardware Development

As expounded above, eye tracking is a useful evaluation tool for training of emergency forces, because of the opportunity to trace the trainees' visual focus of attention [29]. Therefore, the second part of future work is to develop a hardware system for eye tracking during training sessions of emergency forces. The aim is to test the developed eye tracking system in a real training scenario including evaluations of trainee's gaze behavior by the instructors. The usage of eye tracking for training of emergency forces might have additional advantages apart from analyzing gaze behavior. A promising measurement technique called Index of Cognitive Activity (ICA) uses eye tracking data, particularly pupil dilation, to measure cognitive workload [30]. The application of this parameter to measure cognitive workload during training sessions of emergency forces is to be evaluated. Because of dwindling cognitive efforts when actions are performed more automatically, cognitive workload could be a promising additional parameter to capture learning progress.

5 Conclusion

In this article, we described ongoing work aiming at the development of an objective rating tool for training of decision making processes in critical situations. We defined gaze behavior as a useful parameter to track trainees' visual focus of attention, the basis for target-aimed decision making processes. We presented a concept for a laboratory experiment to examine cognitive processes during decision making. The development of a software tool will be based on these findings according to supply relevant data for tracking decision making processes. The second step will be the development of a hardware system for usage during dynamic training sessions that can be integrated into safety gear. The training tool, consisting of software and hardware, will then be evaluated in a field experiment under real circumstances.

References

1. Fakhrhosseini, S.M., Jeon, M., Lautala, P., Nelson, D.: An investigation on driver behaviors and eye-movement patterns at grade crossings using a driving simulator. In: 2015 Joint Rail Conference, San Jose, California, USA, Monday 23 March 2015, pp. V001T06A012. doi:10.1115/JRC2015-5731

2. Pradhan, A.K., Hammel, K.R., DeRamus, R., Pollatsek, A., Noyce, D.A., Fisher, D.L.: Using eye movements to evaluate effects of driver age on risk perception in a driving simulator. Hum. Fact. **47**(4), 840–852 (2005)
3. Salvucci, D.D., Liu, A.: The time course of a lane change. driver control and eye-movement behavior. Transp. Res. Part F: Traffic Psychol. Behav. **5**, 123–132 (2002). doi:10.1016/S1369-8478(02)00011-6
4. Pärnamets, P., Johansson, R., Gidlöf, K., Wallin, A.: How information availability interacts with visual attention during judgment and decision tasks. J. Behav. Dec. Making **29**, 218–231 (2016). doi:10.1002/bdm.1902
5. Hastie, R., Dawes, R.M.: Rational Choice in an Uncertain World. The Psychology of Judgment and Decision Making, 2nd edn. Sage Publications, Los Angeles (2010)
6. Pachur, T., Broder, A.: Judgment: a Cognitive Processing Perspective. Wiley interdisciplinary reviews. Cognitive science (2013). doi:10.1002/wcs.1259
7. Glaholt, M.G., Reingold, E.M.: Eye movement monitoring as a process tracing methodology in decision making research. J. Neurosci. Psychol. Economics **4**, 125 (2011). doi:10.1037/a0020692
8. Goldstein, E.B. (ed.): Wahrnehmungspsychologie. Der Grundkurs, 7th edn. Spektrum Akad. Verl., Berlin (2011)
9. Zelinsky, G.J., Rao, R.P.N., Hayhoe, M.M., Ballard, D.H.: Eye movements reveal the spatiotemporal dynamics of visual search. Psychol. Sci. **8**, 448–453 (1997). doi:10.1111/j.1467-9280.1997.tb00459.x
10. Hayhoe, M., Ballard, D.: Eye movements in natural behavior. Trends Cogn. Sci. **9**, 188–194 (2005). doi:10.1016/j.tics.2005.02.009
11. Posner, M.I., Snyder, C.R., Davidson, B.J.: Attention and the detection of signals. J. Exp. Psychol. **109**(2), 160–174 (1980)
12. Hoffman, J.E., Subramaniam, B.: The role of visual attention in saccadic eye movements. Percept. Psychophys. **57**, 787–795 (1995). doi:10.3758/BF03206794
13. Deubel, H., Schneider, W.X.: Saccade target selection and object recognition. evidence for a common attentional mechanism. Vis. Res. **36**, 1827–1837 (1996). doi:10.1016/0042-6989(95)00294-4
14. Hoffmann, J.E.: Visual attention and eye movements. In: Pashler, H.E. (ed.) Attention. Studies in Cognition Series, pp. 119–150. Psychology Press, Hove (1998)
15. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychol. Bull. **124**(3), 372–422 (1998)
16. Rayner, K.: Eye movements and attention in reading, scene perception, and visual search. Q. J. Exp. Psychol. **62**, 1457–1506 (2009). doi:10.1080/17470210902816461
17. Holmqvist, K.: Eye Tracking. A Comprehensive Guide to Methods and Measures, 1st edn. Oxford University Press, Oxford (2011)
18. Lange, C., Wohlfarter, M., Bubb, H.: Automated analysis of eye-tracking data for the evaluation of driver information systems according to ISO/TS 15007-2:2001. In: Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M.Y., Weikum, G., Jacko, J.A. (eds.) Human-Computer Interaction. New Trends, vol. 5610. LNCS, pp. 105–110. Springer, Heidelberg (2009)
19. Jacob, R.J.K., Karn, K.S.: Eye tracking in human–computer interaction and usability research: ready to deliver the promises. In: Hyönä, J., Radach, R., Deubel, H. (eds.) The Mind's Eye. Cognitive and Applied Aspects of Eye Movement Research, 1st edn., pp. 573–605. Elsevier, Amsterdam u.a (2003)

20. dos Santos, R.D.O.J., de Oliveira, J.H.C., Rocha, J.B., Giraldi, J.D.M.E.: Eye tracking in neuromarketing: a research agenda for marketing studies. Int. J. Psychol. Stud. **7**(1), 32 (2015). doi:10.5539/ijps.v7n1p32
21. Fromm, L., Zerbes, G., Plegge, C., Alexander, T.: Evaluation der Methoden zur Erfassung und Analyse der Entscheidungsfindung in kritischen Situationen am Beispiel Eyetracking. In: Gesellschaft für Arbeitswissenschaft e.V. (GfA) (ed.) 63. GfA Frühjahrskongress, Brugg-Windisch, 15–17 Feb 2017. GfA-Press, Dortmund, Zürich (in press)
22. Ergoneers: official website (2016). http://www.ergoneers.com/. Accessed 16 Nov 2016
23. Zhu, Z., Ji, Q.: Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. Comput. Vis. Image Underst. **98**, 124–154 (2005). doi:10.1016/j.cviu.2004.07.012
24. Wolfe, J.M.: Guided search 2.0 a revised model of visual search. Psychon. Bull. Rev. **1**, 202–238 (1994). doi:10.3758/BF03200774
25. Wolfe, J.M., Cave, K.R., Franzel, S.L.: Guided search: an alternative to the feature integration model for visual search. J. Exp. Psychol. Hum. Percept. Perform. **15**(3), 419–433 (1989)
26. Ratcliff, R.: A theory of memory retrieval. Psychol. Rev. **85**, 59–108 (1978). doi:10.1037/0033-295X.85.2.59
27. Ratcliff, R., McKoon, G.: The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. **20**, 873–922 (2008). doi:10.1162/neco.2008.12-06-420
28. Kingstone, A., Smilek, D., Eastwood, J.D.: Cognitive ethology: a new approach for studying human cognition. Br. J. Psychol. (London, England: 1953) (2008). doi:10.1348/000712607X251243
29. Westhoven, M., Plegge, C., Henrich, T., Alexander, T.: Posture based recognition of the visual focus of attention for adaptive mobile information systems. In: Schmorrow, D.D., Fidopiastis, C.M. (eds.) Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience. LNCS, pp. 416–427. Springer
30. Marshall, S.P.: The index of cognitive activity: measuring cognitive workload. In: Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants New Century, New Trends, Scottsdale, AZ, USA, 15–19 September 2002, pp. 7-5–7-9. doi:10.1109/HFPP.2002.1042860

Advanced Analysis Techniques for Human Error

Detection of Typical Progress Patterns of Industrial Incidents by Text Mining Technique

Toru Nakata(✉) and Mohammad Sohrab

Artificial Intelligence Research Center,
National Institute of Advanced Industrial Science and Technology,
Aomi 2-4-7, Koto-ku, Tokyo, Japan
`toru-nakata@aist.go.jp`

Abstract. To prevent accidents, it is very important to learn why and how past accidents occurred and escalated. The information of accidents is mostly recorded in natural language texts, which is not convenient to analyze the flow of events in the accidents. This paper proposes a method to recognize typical flow of events in a large set of text reports. By focusing two adjacent sentences, our system succeeded to detect typical pairs of predecessor word and successor word. Then we can recognize the typical flows of accidents.

Keywords: Incident Report · Text Mining · Human Factors · Safety Engineering

1 Introduction

Even though general level of technology is advancing, our society is still suffering serious industrial accidents. In Japan, for example, we have experienced several fatal accidents at chemical plant complexes run by major companies almost every year. The most famous cases are Toso accident 2011 (1 fatality), Mitsui Chemical accident 2012 (1 fatality), Nihon Shokubai accident 2012 (1 fatality), and Mitsubishi Material accident 2014 (5 fatalities). Some parts of the accident progress resemble each other. For instance, excessive attention on particular meters made the worker fail to watch crucial information. We can say the workers could have prevent newer accidents if they knew precedents accidents.

Learning from past accidents is very important to prevent accident in the future. That is why many societies concerning industrial safety are collecting records of past accidents [1].

Like ordinary traffic accidents, industrial accidents are easily repeated. Although some accidents can occur as unprecedented and unexpected pattern, most cases of them have parts similar to typical pattern of past accidents. Even unheard-of accidents may have partial similarity to common accidents. Those who know well about past accidents can notice risks for future accidents.

To learn past accidents, we usually rely on text information, such as newswires or official reports edited by expert investigation committees. The best data shapes to

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_22

understand accidents are raw and vivid data like videos or evidential objects of the accident scene. Such data are not convenient to preserve and to collect. So we usually write text reports with summarizing the accidents, and we abandon the raw data.

For certain purposes, text is a convenient shape of information. By reading accident reports, we can understand the story of the accident deeply. But text is costly shape for statistical processing. For instance, it requires careful consideration to find similarities and differences between different accidents by reading their text information. We have to spend long time for it.

There is huge quantity of text information reporting industrial accidents in the world, and the amount is increasing. Human beings cannot read all of them any longer, so the most of texts are left unread and unused.

Natural language processing (NLP) technology can be utilized for the task of reading and understanding of such huge text information. We can use NLP to detect similarities and differences among accident and to clarify causality of events in the accidents. Such analysis will help us to prevent future accidents.

This paper proposes a NLP method that can process huge amount of incident reports to understand typical patterns of progress of incidents.

2 What We Should Extract from Accident Reports

We use the term of ‘incident’ or ‘accident’ for harmful and unpleasant event. An incident is a series of events, which have causality among them and end with bad result.

In general, human can find various types of information from reports, if we have enough time to read and to consider on them. When we try automatic processing of reports, we cannot make the system extract information as much as humans can. We must choose a certain kind of information that can be extracted by the system and is useful for accident prevention.

2.1 Causality is the Most Important Part of Information

Finding causality is crucial to prevent similar incidents, because we cannot hinder occurrence of the bad event without knowing its cause events.

Also, analysis of causality is important to find similarities or differences among incidents. Different incidents may have partially common flows of events. We should detect common patterns of incidents without being deluded by minor differences.

Safety engineering employs graph methods to analyze causality. Traditional methods represented by Event Tree Analysis (ETA), Fault Tree Analysis (FTA), and HAZOP are used to generate graphs of causality among events in an incident [2].

It usually requires deep consideration of experts to composing event causality graphs, so we cannot generate the graphs easily. Automation of the graph composition is strongly required.

Even though automatic detection of events and causal relationship described in texts is one of most active topics in natural language processing, its particular difficulties are becoming clear.

First of all, the definition of causal relationship is difficult [3]. While we usually regard causal relationship such as “A person will die if he eats potassium cyanide,” we do not accept awkward (but logically true) relationships like “Churchill was not born if Cleopatra was not born.”

Moreover, there is a problem of false correlation like “Mr. A’s age will go up by two if age of Mr. B goes up by two.” Fake correlations cannot be eliminated by observing only their correlation coefficient. Deep consideration on meaning of texts with common sense is required for detection of causal relationship.

In this paper, we avoid trying precise detection of causality. Instead, our system merely analyzes statistical correlations among words. That may not reflect strict causality of events, but we may find clues about how the accidents progressed.

2.2 Limitation of Simple Text Classification Method

We assume that there are representative and common patterns of incidents. Knowing such patterns, we can prevent similar incident in the future. So detection of such patterns is one of purpose for text mining of the report.

One of ordinary methods to detect typical patterns of documents is text classification. Under this procedure, we regard each report as a set of contained words (so-called ‘bag of words’ or BoW). Each report is transformed into a set of numbers (or we can call it a vector), which are appearance frequencies of words. We ignore neither grammatical structure nor order of the sentences to keep the analysis simple. We can observe similarity among the reports by comparing BoW vector.

Figure 1 shows an analysis result of similarity among actual accident reports of aviation. (We will explain the detail of the reports at Sect. 4.) Unfortunately, the distribution of report similarity was vague, so we could not extract particular information from that. This is because this BoW analysis was too simple.

3 Proposed Method for Incident Report Analysis

We observe appearance order of words in the incident reports. As shown above, treating each report as one BoW was not enough for analysis. We therefore should pay more attention on order of sentences and order of words.

We assume there are typical patterns of word appearance order, which reflect particular common patterns of event progress of incidents.

We explain the procedure over an example of an incident report below.

> *I told her that it was 11 minutes until scheduled departure. I had also informed the gate agents twice during boarding that oversized carry-on baggage was entering the aircraft. The agent Supervisor stated that the cargo bin had already closed.*

First, we apply normalization: we eliminate ‘stop words’ (common terms appear mainly for grammatical control and does not reflect the context well) and transform left terms to their original form. In addition, we put the markers to indicate the head and the end of a report.

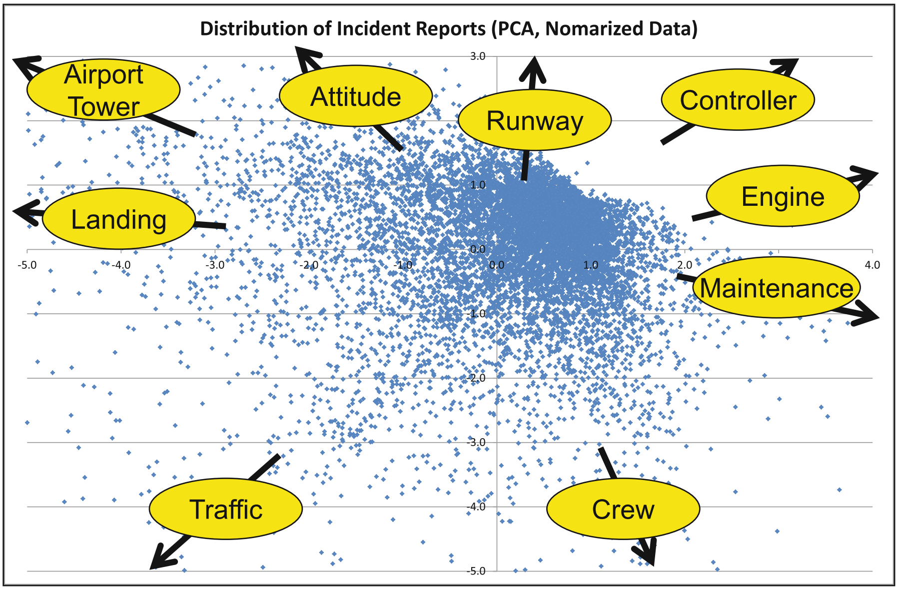


Fig. 1. Conventional Principal Component Analysis result of the aviation incident reports. Each dot represents a particular report. The distances among the dots reflect similarity of words contained in reports. The arrows indicate words typically appearing in the reports of their directions. The distribution is heavily concentrated in the center, so it is hard to find differences among the reports around the center.

- *{'StoryStart'}*
- *{tell, minute, schedule (verb), departure.}*
- *{inform, gate, agent, twice, board (verb), oversize, carry-on, baggage, enter, aircraft}*
- *{agent, supervisor, state, cargo, bin, close}*
- *{'StoryEnd'}*

Second, we count frequency of word pairs. In this example, the first sentence contains term 'tell', and the second sentence has term 'inform'. So we add one to frequency record of the pair of (tell → inform). Likewise, we count pairs of words between two neighboring sentences: we count the occurrence of the pair of (oversize → agent), which appears in second and third sentences. We do not count the pair of (depature → bin), since those words does not appear in neighboring sentences but first and third sentences.

Counting occurrences of such word orders, we get a transition matrix, which indicates statistical tendency about which word appears in next sentence after appearance of which term.

If there are strong patterns of word appearance flow, the transition matrix contains their component. We can extract such component by using method of matrix factorization and so on.

Finally we can detect typical and common pattern of incidents among a large set of reports.

4 Experiment

In general, firm tendencies of word order in stories are not exactly equal to of causal relationship, even though we hope the order may be used as an indicator of causality. We test the effectiveness of the proposed method through an experiment with real incident reports.

4.1 Dataset of Aviation Incident Report

We used 4,469 reports of NASA's Aviation Safety Reporting System (ASRS). The data consist of all reports of 2013. We chose the particular part named 'narrative', which is made up of recollected story of event written by people who concerned the incidents. Figure 1 is a part of a real ASRS narrative report.
Characteristics of the data are as the following:

- Language: English
- Author: Persons who concerned the incidents (i.e. pilots, ground crew, cabin crew, etc.)
- Amount of reports: 4 469. (All data of 2013.)
- Amount of words: 1 365 260.
- Amount of kinds of words: 28 615.
- Amount of sentences: 110 963.
- In average, a report may contain 305 words in 25 sentences.

4.2 Visualization Method

In this paper, we visualize the result with the following procedure. We will get the transition matrix T consisting of t_{ij}, which is occurrence frequency of ordinal word pair $(word_i \rightarrow word_j)$.

We can rate tendency appearance order of each word in reports by calculating the following indexes.

$$a_i = \sum_j t_{ij} - \sum_j t_{ji} \tag{1}$$

$$b_i = \sum_{tij > 0} a_j - \sum_{tij < 0} a_j \tag{2}$$

We call b_i as Word Order Index of word i. (Index of a_i may express some characteristic about word order, but it is less stable than b_i.)

Second, we measure commonality of context among the words. We assign a 2-dimensional position vector for each word. Assume $\vec{v}_i = (x_i, y_i)$ is the vector for word i. At the beginning, the vectors are set with random numbers.

We adjust the vectors in iteration of (3) to (5) to make the vector express commonality of context.

$$c_{ij} = t_{ij} + t_{ji} \tag{3}$$

$$\vec{v}_i := \sum_j c_{ij} \cdot \vec{v}_j / \sum_j c_{ij} \tag{4}$$

$$\vec{v}_i := \left(1 + \frac{K}{|\vec{v}_i|^2}\right)\vec{v}_i \tag{5}$$

K is a constant coefficient.

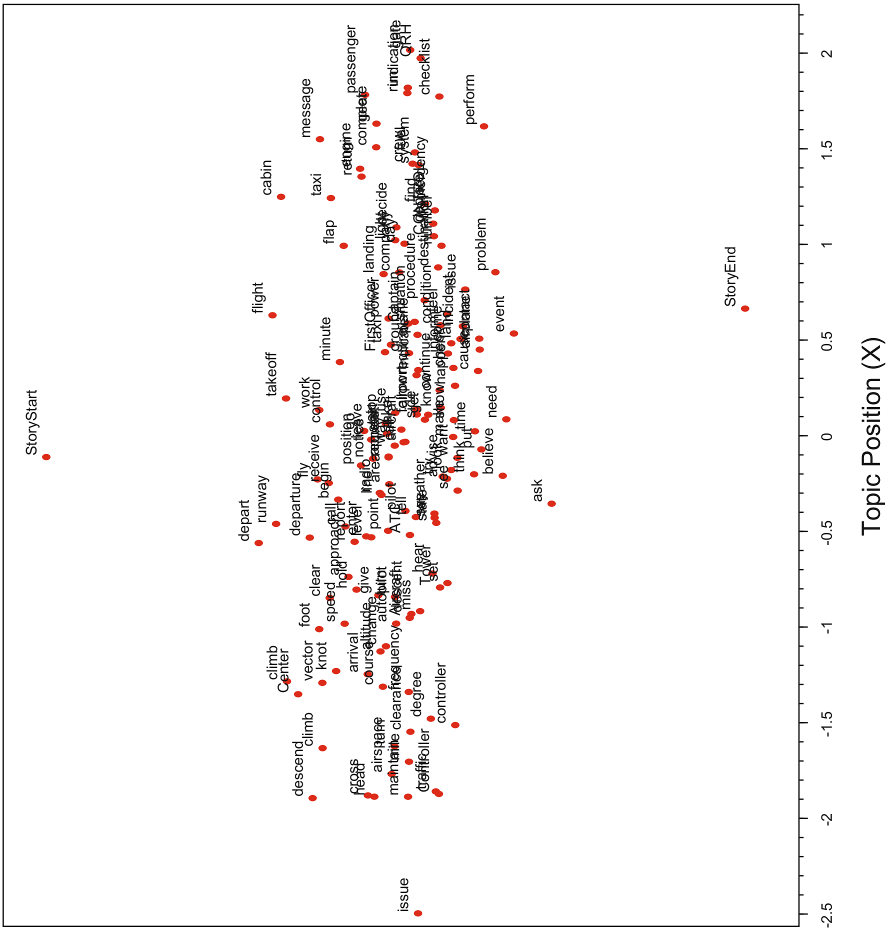


Fig. 2. A visualization of the word transition result. Distribution of all words with the markers of beginning and end of report.

By processing this, groups of words that tend to appear in same sequences of sentences together will have certain position vectors similar to each other.

4.3 Results and Discussion

Figure 2 shows the result about Word Order Index. In typical cases, reports starts with the point of ‘StoryStart’ located at top of the plot, then the report story moves toward the point of ‘StoryEnd’ at the bottom. During the movement, words located on the route tend to appear in the story (Fig. 3).

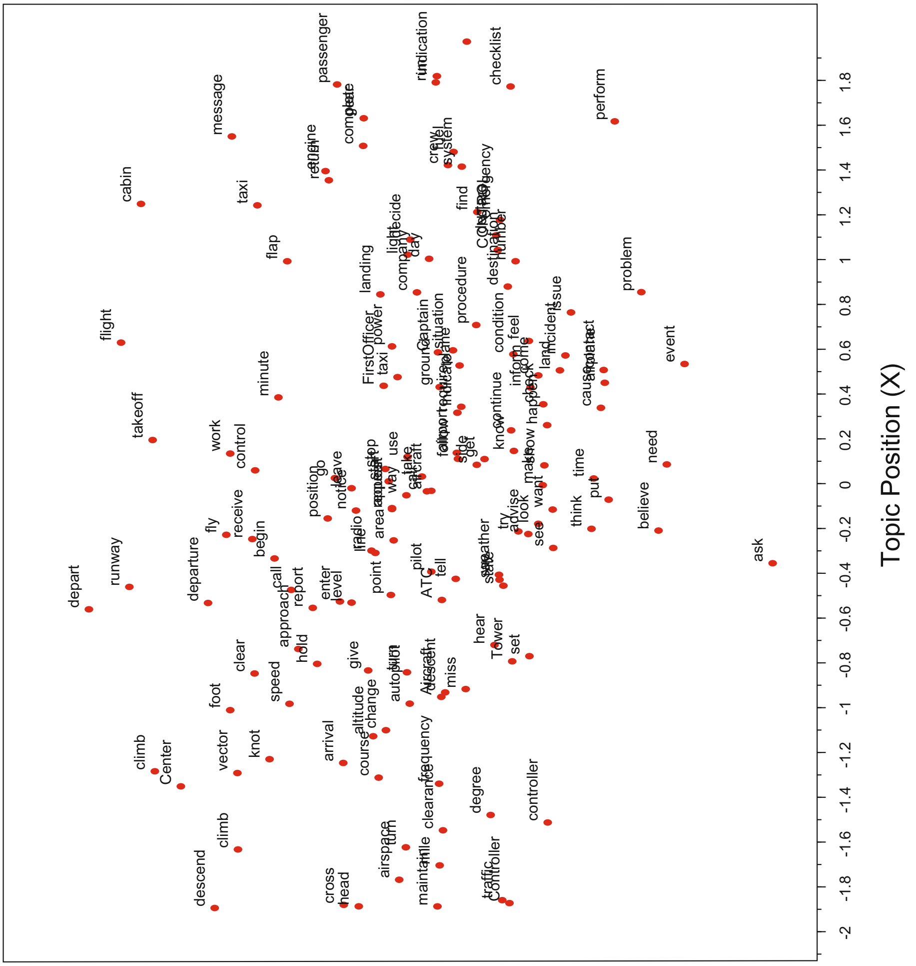


Fig. 3. Zoom-up of the center area of Fig. 2.

Figure 4, which is the result of word position vectors, gives a view for commonality of context.

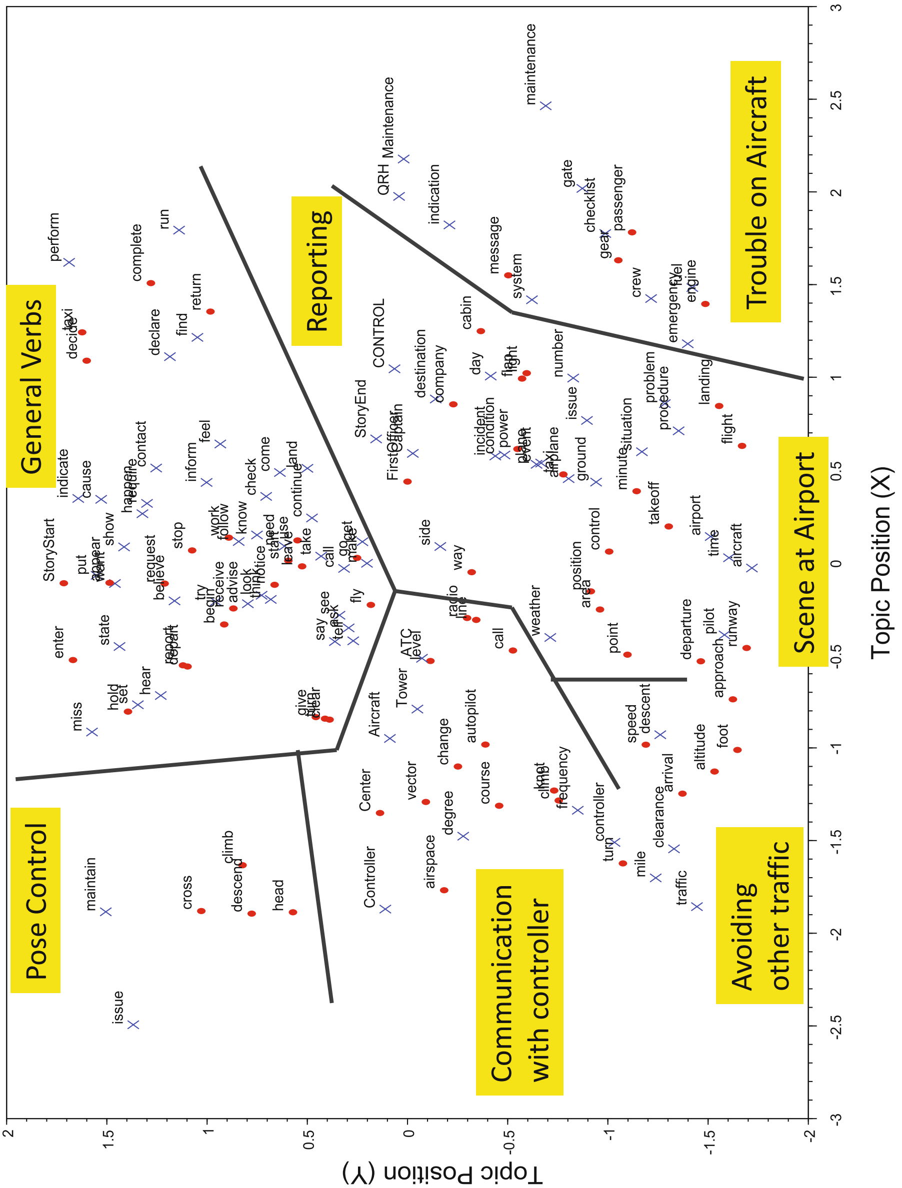


Fig. 4. Multi-dimensional scaling result of words familiarity with interpretations of meaning of clusters. A red dot indicates that word order index of its word is plus (i.e. the word tends to appear earlier in reports.). A blue cross stands for negative word order index.

We found some word clusters in the distribution of Fig. 4, so we subjectively classified them and gave labels of their meanings: as the following:

- Cluster 1: general verbs
- Cluster 2: words about pose control: *'descend', 'climb', 'head (verb)',* etc.
- Cluster 3: words about communication with the aviation controller: *'Controller', 'Tower', 'call', 'radio',* etc.
- Cluster 4: words about avoiding near-miss with other airplanes: *'traffic', 'clearance', 'turn',* etc.
- Cluster 5: words for events at a airport: *'runway', 'takeoff', 'departure', 'landing',* etc.
- Cluster 6: words for mechanical troubles on aircraft: *'engine', 'gear', 'system', 'maintenance',* etc.
- Cluster 7: words for reporting incidents: *'incident', 'issue', 'day',* etc.

In general, stories tend to move within same clusters or their neighbors. We understand that the clusters represent typical patterns of whole stories or particular scene of actual incidents. For instance, troubles like Cluster 4 and 6 are very common incident patterns.

Thanks to this text mining technique, we extract common incident patterns without reading all of huge set of the reports.

5 Conclusion

This paper proposed a text-mining method to analyze texts of accident reports automatically. The method extracts the flows of events of accidents by observing transition of words in neighboring two sentences. We applied the method on large dataset of 4,468 reports about real aviation incidents and found the typical flows.

This work is still at very beginning state, and we ignored information about synonyms and hypernyms. In future work, we will employ ontology to extract events in report texts more precisely.

Acknowledgements. This paper is based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

1. Center for Chemical Process Safety: Guidelines for Implementing Process Safety Management, 2nd edn. (2016)
2. International Electrotechnical Commission: IEC 61508-7:2000 Functional Safety of Electrical Electronic/Programmable Electronic Safety-Related Systems — Part 7, Overview of Techniques and Measures (2000)
3. Pearl, J.: Causality: Models, Reasoning and Inference. Cambridge University Press, Cambridge (2000)
4. Nakata, T.: Text-mining on incident reports to find knowledge on industrial safety. In: IEEE Annual Reliability and Maintainability Symposium (2017)

Use of a Big Data Mining Technique to Extract Relative Importance of Performance Shaping Factors from Event Investigation Reports

Jinkyun Park(✉), Yochan Kim, and Wondea Jung

Risk and Environmental Safety Research Division,
Korea Atomic Energy Research Institute, 989-111 Daedeok-Daero,
Yuseong-Gu, Daejeon 34057, Republic of Korea
{kshpjk, yochankim, wdjung}@kaeri.re.kr

Abstract. In this study, the relative importance of significant performance shaping factors (PSFs), which is critical for estimating the human error probability (HEP) of a given task environment is extracted from event investigation reports of domestic nuclear power plants (NPPs). Each event was caused by one or more human performance related problems (i.e., human errors), and its investigation report includes detailed information describing why the corresponding event has occurred. Based on 10 event reports, 47,220 data records were identified, which represent the task environment of 11 human errors in terms of significant PSFs. After that, the relative importance of the associated PSFs was analyzed by using a CART (Classification and Regression Tree) method that is one of the representative techniques to scrutinize the characteristics of big data.

Keywords: Human reliability analysis · Performance shaping factors · Event investigation report · Classification and regression tree · Nuclear power plant

1 Introduction

The operation of complicated systems such as railroad networks, commercial airplanes, and NPPs (Nuclear Power Plants) historically showed that even a small event can cause a catastrophic failures resulting in enormous casualties. The Fukushima accident on March 11, 2011 clearly demonstrated this experience. In this regard, it is very important to emphasize that a human performance related problem (e.g., human error) is one of the significant factors affecting the safety of the complicated systems. Subsequently, in order to enhance their safety through minimizing the likelihood of human error, a lot of practical approaches have been deployed for several decades across many industrial sectors.

From this standpoint, one of the most practical solutions would be the management of safety critical tasks based on a systematic framework and/or technique, such as an HRA (Human Reliability Analysis). In other words, if HRA practitioners are able to identify plausible error forcing factors (e.g., PSFs; Performance Shaping Factors) for a given task context, effective countermeasures that are helpful for reducing the possibility of human error can be specified by thinking of how to eliminate the associated

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_23

PSFs. In this light, it is very important to clarify the catalog of PSFs with their relative importance on the possibility of human error in a systematic manner.

For this reason, the collection of HRA data from event reports has been carried out for several decades in many industrial sectors, which reflect the response of human operators who were faced with various kinds of error-prone situations. At the same time, the catalog of significant PSFs was also proposed based on the HRA data. Unfortunately, in terms of determining the relative importance of significant PSFs, most of existing methods are insufficient from two perspectives: (1) a lack of technical underpinnings for analyzing HRA data, and (2) a lack of an ability to manipulate huge amount of HRA data [1, 2].

In order to address the abovementioned issues, in this study, the relative importance of significant PSFs was estimated based on a big data mining technique. To this end, event reports attributable to human errors were collected. After that, the context information of each human error was investigated based on the description of an event report, which is critical for representing the task environment of each human error. Finally, the relative importance of significant PSFs was calculated based on a CART (Classification and Regression Tree) method.

2 Collecting Event Reports

When a safety significant event (such as an event resulting in the initiation of engineered safety features and unexpected reactor trip) has occurred in the NPPs of the Republic of Korea, it is mandatory to report detailed information to the public. In this regard, the role of the KINS (Korea Institution of Nuclear Safety), which is the nuclear regulatory body of the Republic of Korea, is to dispatch a special force to investigate including (1) what went wrong, (2) the progression of a safety significant event, and (3) countermeasures to prevent from the recurrence of similar events. Once the investigation of the safety significant event is finished, the KINS uploads all kinds of information to the public via the Internet since 2002 [3]. For this reason, total 193 investigation reports that cover from January 2002 to December 2013 were reviewed in

Table 1. Human errors identified from the analysis of event reports

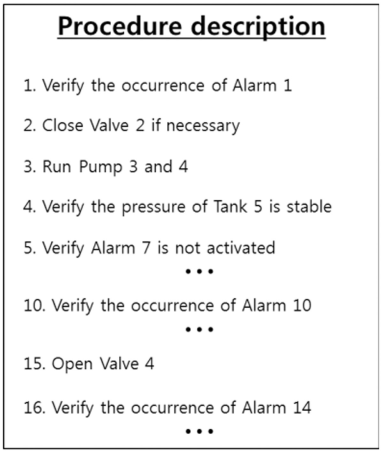
Event ID	Task type	Error mode	Remark
1	Manipulating simple (discrete) control	EOC	Normal condition
2	Manipulating simple (discrete) control	EOO	Normal condition
3	Manipulating simple (discrete) control	EOC	Off-normal condition
4	Manipulating simple (discrete) control	EOC	Normal condition
5	Manipulating dynamically	EOC	Off-normal condition
6	Transferring step in procedure	EOO	Off-normal condition
7	Manipulating simple (discrete) control	EOC	Off-normal condition
8	Manipulating dynamically	EOC	Off-normal condition
9	Manipulating simple (discrete) control	EOC	Normal condition
10	Manipulating simple (discrete) control	EOO	Off-normal condition
11	Manipulating dynamically	EOC	Off-normal condition

detail. As a result, 11 kinds of human errors are successfully identified from event investigation reports. Table 1 epitomizes human errors with the associated task types.

For example, an EOC (Error of Commission) was observed from the first event report, which happened during the performance of simple control task in a normal operation condition. Similarly, the EOC of the same task was identified from the seventh event report, which has occurred during an off-normal condition. However, in the case of the fifth event report, an EOC was recognized when a human operator was trying to control a certain component along with a dynamically varying situation. More detailed explanation on the task type of Table 1 will be given in the next section.

3 Identifying Task Types

Let us assume that, during the off-normal condition of an NPP, human operators have to conduct a hypothetical procedure depicted in Fig. 1, which is indispensable for coping with it (i.e., restoring the status of the NPP to a normal condition by removing a root cause resulting in the off-normal condition).



Procedure description

1. Verify the occurrence of Alarm 1
2. Close Valve 2 if necessary
3. Run Pump 3 and 4
4. Verify the pressure of Tank 5 is stable
5. Verify Alarm 7 is not activated

• • •

10. Verify the occurrence of Alarm 10

• • •

15. Open Valve 4
16. Verify the occurrence of Alarm 14

• • •

Fig. 1. Description of a hypothetical procedure, reproduced from Ref. [4].

As can be seen from Fig. 1, in order to properly deal with an off-normal condition at hand, human operators have to accomplish a set of required tasks (or actions) along with the predefined sequence of a procedure. Here, it is very interesting to point out that the contents of tasks can be regrouped into several categories based on their nature. For example, it is possible to say that the task type of four tasks (i.e., the first, fifth, tenth and sixteenth task) can be represented by ‘Verifying alarm occurrence’ because their nature is to verify whether or not a specific alarm is generated. This strongly implies that a firm taxonomy that can be used to clarify the types of tasks included in a procedure. For this reason, the catalog of task types and the associated error modes with respect to the representative cognitive activities of human operators (e.g.,

Table 2. Catalog of task types with the associated error modes used in this study, reproduced from Ref. [5].

ID	Task type	Error mode
1	Verifying alarm occurrence	EOO (Error of Omission), EOC
2	Verifying state of indicator	EOO, EOC
3	Synthetically verifying information	EOO, EOC
4	Reading simple value	EOO, EOC
5	Comparing parameter	EOO, EOC
6	Comparing in graph constraint	EOO, EOC
7	Comparing for abnormality	EOO, EOC
8	Evaluating trend	EOO, EOC
9	Entering step in procedure	EOO
10	Transferring procedure	EOO, EOC
11	Transferring step in procedure	EOO, EOC
12	Directing information gathering	EOO, EOC
13	Directing manipulation	EOO, EOC
14	Directing notification/request	EOO, EOC
15	Diagnosing	EOO, EOC
16	Identifying overall status	EOO, EOC
17	Predicting	EOO, EOC
18	Manipulating simple (discrete) control	EOO, EOC
19	Manipulating simple (continuous) control	EOO, EOC
20	Manipulating dynamically	EOO, EOC
21	Notifying/requesting to the outside of a control room	EOO, EOC

information gathering and situation interpreting) is adopted in this study. Table 2 summarizes task types and the associated error modes.

From Table 2, most of task types are self-explainable. For example, 'Verifying state of indicator' denotes a task type of human operators who have to read the status of an indicator, while that of 'Comparing parameter' designates the comparison of the values of two or more process parameters, such as pressurizer pressure and main steam pressure. However, there are several task types that have remarkable features.

First, the task type of 'Synthetically verifying information' means the determination of a component state using two or more information sources. Although its nature is very similar to that of 'Identifying overall status,' the definition of the former should be distinguished from that of the latter because there are times when human operators need to decide the status of a component by integrating additional information. For example, the operability of a pump can be confirmed by the integration of supplementary information such as the readiness of lubrication pumps and the associated valve alignments.

Second, the task type of 'Comparing for abnormality' should be considered when human operators need to check the abnormality of a component and/or system through the comparison of several process parameters that can be varied with respect to the status of an NPP. For example, let us assume an arbitrary task such as 'Determine if pressurizer pressure is less than 96 kg/cm^2 (105 kg/cm^2 for ADVERSE condition).' In order to accomplish this task, human operators need to determine whether or not the status of a containment is in 'ADVERSE' condition, of which the satisfaction can be confirmed when either containment pressure is greater than 0.35 kg/cm^2 or the radiation dose of the containment is greater than 105 R/hr. This means that the nature of 'Comparing for abnormality' task is definitely different from that of 'Comparing parameter' task. Similarly, since the task of 'Comparing in graph constraint' requires human operators to clarify the satisfaction of a specific condition by using a predefined parameter curve (e.g., determining an appropriate flow rate from a certain reservoir based on its pressure curve), it is indispensable to distinguish this task type from 'Comparing parameter' task.

Fourth, it is necessary to discriminate the types of control tasks into threefold along with their nature: (1) simple (discrete) control, (2) simple (continuous) control, and (3) dynamic control. Here, the definition of each control task type can be summarized as follows.

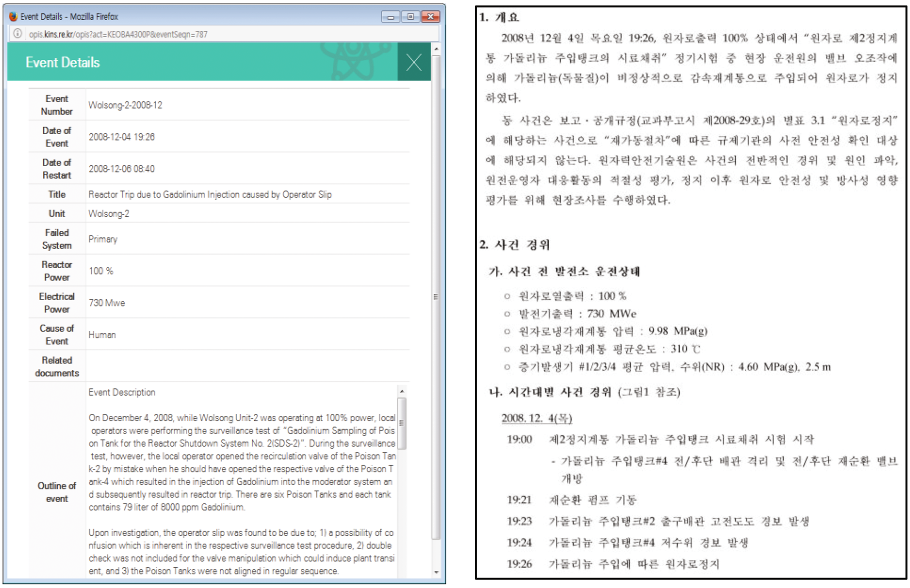
- Simple (discrete) control: a control task that can be accomplished by a dichotomous control (e.g., on-off, open-close, or start-stop);
- Simple (continuous) control: a control task that can be accomplished by selecting a set of discrete states or a specific point within a continuous range;
- Dynamic control: a control task that requires the monitoring and/or regulating two or more dedicated components (e.g., adjusting a flow rate in accordance with the water level of a certain tank).

4 Analyzing Event Reports

As can be seen from Table 2, except the task type of 'Entering step in procedure,' two kinds of human error modes (i.e., EOO and EOC) are commonly considered in all the task types. In general, the EOO implies that human operators did not carry out a required task while the EOC did the required task in a wrong way. For example, the EOC of 'Directing information gathering' task type denotes that a human operator instructed another human operator to read a wrong indicator. It is to be noted that, in terms of 'Entering step in procedure' task, it is not reasonable to count its EOC because of the EOC of 'Transferring step in procedure' task. In other words, the EOC of the former directly corresponds to that of the latter.

Based on the taxonomy of task types and the associated error modes, it is possible to analyze the contents of a procedure. In this regard, let us consider an unexpected reactor trip occurred in one of the domestic NPPs in 2008. Figure 2a shows a snapshot of the event report provided from the website of the OPIS (Operational Performance Information System for nuclear power plant), which is an official database managed by the KINS [3]. In addition, Fig. 2b shows a part of the event report published by the KINS (written in Korean).

According to the investigation report, this event has occurred in December 4, 2008 at one of the domestic NPPs. At that time, human operators were performing a maintenance procedure (supposed to be conducted in every seven days), which



Event Details

Event Number	Wolsong-2-2008-12
Date of Event	2008-12-04 19:26
Date of Restart	2008-12-06 08:40
Title	Reactor Trip due to Gadolinium Injection caused by Operator Slip
Unit	Wolsong-2
Failed System	Primary
Reactor Power	100 %
Electrical Power	730 Mwe
Cause of Event	Human
Related documents	
Outline of event	Event Description On December 4, 2008, while Wolsong Unit-2 was operating at 100% power, local operators were performing the surveillance test of "Gadolinium Sampling of Poison Tank for the Reactor Shutdown System No. 2(SDS-2)". During the surveillance test, however, the local operator opened the recirculation valve of the Poison Tank-2 by mistake when he should have opened the respective valve of the Poison Tank-4 which resulted in the injection of Gadolinium into the moderator system and subsequently resulted in reactor trip. There are six Poison Tanks and each tank contains 79 liter of 8000 ppm Gadolinium. Upon investigation, the operator slip was found to be due to; 1) a possibility of confusion which is inherent in the respective surveillance test procedure, 2) double check was not included for the valve manipulation which could induce plant transient, and 3) the Poison Tanks were not aligned in regular sequence.

(a)

1. 개요

2008년 12월 4일 목요일 19:26, 원자로출력 100% 상태에서 "원자로 제2정지계통 가돌리늄 주입탱크의 시료채취" 정기시험 중 현장 운전원의 밸브 오조작에 의해 가돌리늄(독물질)이 비정상적으로 감속재계통으로 주입되어 원자로가 정지하였다.

동 사건은 보고 · 공개규정(교과부고시 제2008-29호)의 별표 3.1 "원자로정지"에 해당하는 사건으로 "재가동절차"에 따른 규제기관의 사전 안전성 확인 대상에 해당되지 않는다. 원자력안전기술원은 사건의 전반적인 경위 및 원인 파악, 원전운영자 대응활동의 적절성 평가, 정지 이후 원자로 안전성 및 방사성 영향 평가를 위해 현장조사를 수행하였다.

2. 사건 경위

가. 사건 전 발전소 운전상태

- 원자로열출력 : 100 %
- 발전기출력 : 730 MWe
- 원자로냉각재계통 압력 : 9.98 MPa(g)
- 원자로냉각재계통 평균온도 : 310 ℃
- 증기발생기 #1/2/3/4 평균 압력, 수위(NR) : 4.60 MPa(g), 2.5 m

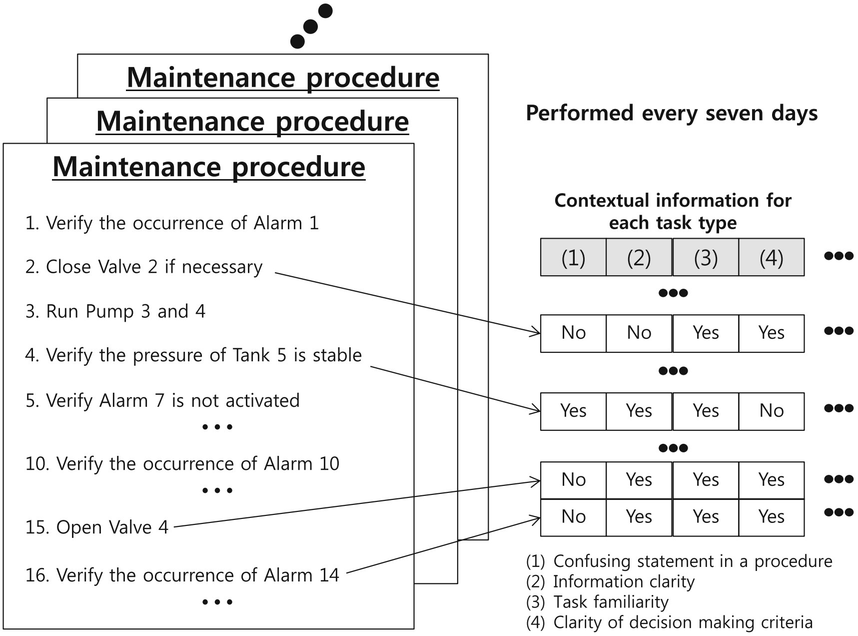
나. 시간대별 사건 경위 (그림1 참조)

2008. 12. 4(목)

19:00 제2정지계통 가돌리늄 주입탱크 시료채취 시험 시작
  - 가돌리늄 주입탱크#4 전/후단 배관 격리 및 전/후단 재순환 밸브 개방
19:21 재순환 펌프 기동
19:23 가돌리늄 주입탱크#2 출구배관 고전도도 경보 발생
19:24 가돌리늄 주입탱크#4 저수위 경보 발생
19:26 가돌리늄 주입에 따른 원자로정지

(b)

Fig. 2. Summary of an unexpected reactor trip occurred in a domestic NPP.



(1)	(2)	(3)	(4)
No	No	Yes	Yes
Yes	Yes	Yes	No
No	Yes	Yes	Yes
No	Yes	Yes	Yes

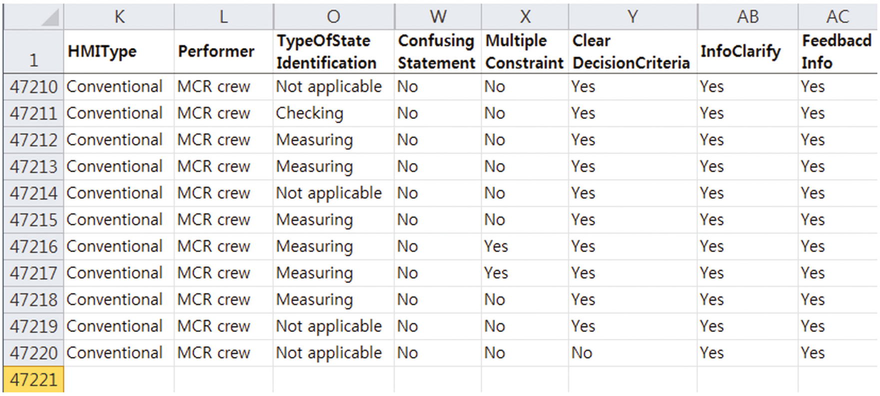
(1) Confusing statement in a procedure
(2) Information clarity
(3) Task familiarity
(4) Clarity of decision making criteria

Fig. 3. Extracting contextual information from the description of an event report.

describes how to sample the concentration of gadolinium nitrate solution for one of six storage tanks. Unfortunately, since human operators open a wrong valve during the performance of the maintenance procedure, a large amount of gadolinium nitrate solution was injected to a reactor resulting in an unexpected reactor trip. In this light, it is possible to distinguish various kinds of contextual information related to the performance of the maintenance procedure. Figure 3 depicts an underlying idea about how to extract the contextual information.

For example, Park et al. proposed the catalog of data items with the associated instances that are essential for identifying contextual information related to when and why human error has occurred [5]. Typical data items and the associated instances are: (1) HMI (Human Machine Interface) type, (2) task familiarity (Y/N), (3) time pressure (Y/N), (4) confusing statement (Y/N), (5) clarity of decision making criteria (Y/N), (6) information clarity (Y/N), and (7) feedback information (Y/N). With these data items, their instances can be properly marked based on the contents of an event report. That is, if there are any explanations in the event report indicating that human operators did not have sufficient experience on the performance of the maintenance procedure, the instance of the task familiarity should be assigned as ‘Y.’ Similarly, if the contents of the maintenance procedure were described with many conditional statements (e.g., IF-THEN-ELSE), the instance of the confusing statement should be designated as ‘Y.’ More interesting point is that most of maintenance procedures are conducted with a fixed time interval. This indicates that the amount of contextual information could be drastically increased. That is, as far as human operators have to use a same procedure, the instances of several data items (e.g., time pressure, confusing statement, and clarify of decision making criteria) would be also identical.

For this reason, 11 event reports dealing with human errors shown in Table 1 are analyzed in detail. As a result, a total of 47,220 records were secured from the event reports, which specify the contextual information of each task type summarized in Table 2. Figure 4 shows a part of contextual information extracted from the analysis of event reports.



	K	L	O	W	X	Y	AB	AC
1	HMIType	Performer	TypeOfState Identification	Confusing Statement	Multiple Constraint	Clear DecisionCriteria	InfoClarify	Feedbacd Info
47210	Conventional	MCR crew	Not applicable	No	No	Yes	Yes	Yes
47211	Conventional	MCR crew	Checking	No	No	Yes	Yes	Yes
47212	Conventional	MCR crew	Measuring	No	No	Yes	Yes	Yes
47213	Conventional	MCR crew	Measuring	No	No	Yes	Yes	Yes
47214	Conventional	MCR crew	Not applicable	No	No	Yes	Yes	Yes
47215	Conventional	MCR crew	Measuring	No	No	Yes	Yes	Yes
47216	Conventional	MCR crew	Measuring	No	Yes	Yes	Yes	Yes
47217	Conventional	MCR crew	Measuring	No	Yes	Yes	Yes	Yes
47218	Conventional	MCR crew	Measuring	No	No	Yes	Yes	Yes
47219	Conventional	MCR crew	Not applicable	No	No	Yes	Yes	Yes
47220	Conventional	MCR crew	Not applicable	No	No	No	Yes	Yes
47221								

Fig. 4. A part of records for the contextual information of each task type.

Based on the records of the contextual information, a CART (Classification and Regression Tree) analysis that is one of the representative techniques for visualizing the characteristics of big data was carried out with respect to each data item [6]. Table 3 highlights important results obtained from the CART analysis.

Table 3. Some results obtained from the CART analysis

Data item	Data instance	Relative importance
HMI type	Conventional	1
	Computer-based	2.6
Time pressure	No	1
	Yes	5.3
Clear decision making criteria	Yes	1
	No	15.1
Feedback information	Yes	1
	No	95.9

For example, in the case of ‘Time pressure,’ it is observed that the ratio of human error rates (i.e., human error rate under the time pressure divided by human error rate without the time pressure) is 5.3. In contrast, in the case of ‘Clear decision making criteria,’ it is revealed that the ratio of human error rates is 15.1. This means that, in terms of the likelihood of human error, the provision of a clear decision making criterion is more important about five times than the existence of the time pressure. In addition, the provision of the feedback information is utmost important for reducing the likelihood of human error because its importance seems to be six times comparing to that of the clear decision making criteria.

5 General Conclusion

One of the key factors affecting the quality of HRA results would be the determining the relative importance of significant PSFs (i.e., PSF multipliers). In this regard, many researchers proposed the list of significant PSFs such as (1) procedure quality, (2) ergonomics/HMI, and (3) available time (e.g., time pressure) [7, 8]. Unfortunately, most of PSF multipliers were determined without appropriate technical underpinnings (e.g., expert’s judgement).

Here, it is very interesting to point out that the relative importance of data items for identifying contextual information can be used to determine PSF multipliers. For example, one of the data items included in Table 3 is ‘Time pressure,’ which is directly comparable with a significant PSF – the available time. In addition, it is possible to say that the data item of ‘Clear decision making criteria’ is a typical characteristic specifying the quality of a procedure. If so, although limited numbers of event reports were analyzed, the result of this study can be used as a technical basis, which allows us to estimate PSF multipliers in an objective way.

References

1. International Atomic Energy Agency (IAEA): Collection and Classification of Human Reliability Data for Use in Probabilistic Safety Assessments. IAEA-TECDOC-1048, Vienna (1998)
2. Nuclear Energy Agency (NEA): HRA Data and Recommended Actions to Support the Collection and Exchange of HRA Data. Report of Working Group on Risk Assessment, NEA/CSNI/R(2008)9 (2008)
3. Operational Performance Information System (OPIS). http://www.opis.kins.re.kr
4. Park, J., Kim, Y., Jung, W.: Application of the HuREX Framework for Calculating Empirical Human Error Probabilities from the Operation Experience of Domestic Nuclear Power Plants – Preliminary Results. KAERI/TR-6474, Korea Atomic Energy Research Institute, Daejeon, Republic of Korea (2016)
5. Park, J., Kim, Y., Jung, W.: A Framework to Estimate HEPs from the Full-scope Simulators of NPPs: Unsafe Act Definition, Identification and Quantification. KAERI/TR-6401, Korea Atomic Energy Research Institute, Daejeon, Republic of Korea (2016)
6. Shmueli, G., Patel, N.R., Bruce, P.C.: Data Mining for Business Intelligence. Wiley, New Jersey (2010)
7. Barriere, M., Bley, D., Cooper, S., Forester, J., Kolaczkowski, A., Luckas, W., Parry, G., Ramey-Smith, A., Thompson, C., Whitehead, D., Wreathall, J.: Technical Basis and Implementation Guidelines for a Technique for Human Event Analysis (ATHEANA). NUREG-1624, Rev. 1, U.S. Nuclear Regulatory Commission, Washington DC (2000)
8. Lois, E., Dang, V.N., Forester, J., Broberg, H., Massaiu, S., Hildebrandt, M., Braarud, P.Ø., Parry, G., Julius, J., Boring, R., Männistö, I., Bye, A.: International HRA Empirical Study – Description of Overall Approach and Pilot Phase Results from Comparing HRA Methods to Simulator Performance Data. NUREG/IA-2016, vol. 1, U.S. Nuclear Regulatory Commission, Washington, DC (2009)

Text Mining for Procedure-Level Primitives in Human Reliability Analysis

Sarah M. Ewing[1(✉)], Ronald L. Boring[1], Martin Rasmussen[2], and Thomas Ulrich[1]

[1] Idaho National Laboratory, PO Box 1625, Idaho Falls, ID 83415, USA
{sarah.ewing, ronald.boring, thomas.ulrich}@inl.gov
[2] NTNU Social Research, Studio Apertura, Dragvoll Allé 38 B,
7491 Trondheim, Norway
martin.rasmussen@svt.ntnu.no

Abstract. The classification of nuclear power plant procedures at the sub-task level can be accomplished via text mining. This method can inform dynamic human reliability calculations without manual coding. Several approaches to text classification are considered with results provided. When a discrete discriminant analysis is applied to the text, this results in clear identification procedure primitive greater than 88% of the time. Other analysis methods considered are Euclidian difference, principal component analysis, and single value decomposition. The text mining approach automatically decomposes procedure steps as Procedure Level Primitives, which are mapped to task level primitives in the Goals, Operation, Methods, and Section Rules (GOMS) human reliability analysis (HRA) method. The GOMS-HRA method is used as the basis for estimating operator timing and error probability. This approach also provides a tool that may be incorporated in dynamic HRA methods such as the Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER) framework.

Keywords: Human reliability analysis · Computation-based human reliability analysis · Human error · GOMS-HRA · Text mining

1 Introduction

The quantification of nuclear power plant (NPP) anomalous events as a probability over time is called dynamic probability risk assessment (PRA). The use of PRA in NPPs has become commonplace in NPPs, with quantification methods implemented throughout the entire U.S. NPP fleet. Examples of PRA methodology implemented by regulators include the Systems Analysis Programs for Hands-on Integrated Reliability Evaluations (SAPHIRE) and the Standardized Plant Analysis Risk (SPAR) models. However, the human component in each NPP is difficult to quantify due to commission and omission errors. Closer inspection of the NPP operation manuals that are implemented can give real-time quantitative information on human behavior, with insights into the specific human actions that need to take place in order for an NPP to operate.

The classification of NPP procedures can be accomplished via the use of text mining capabilities. This method can then be used to inform dynamic human reliability

© Springer International Publishing AG (outside the USA) 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*, Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_24

calculations without the need for tedious manual coding or analysis. This approach includes an objective assessment of the procedures based on previous expert assessments conducted by analysts. Providing an initial objective assessment allows experts to identify anomalies in the algorithms and contribute to an objective result that will provide consistent outcomes.

The application of a Goals, Operation, Methods, and Section Rules (GOMS) model as applied to NPP operator actions is detailed in [1]. And the subsequent application to NPP operation manuals and association to timing data to complete steps are detailed in [2]. In this exploration, the procedures were taken from NPP control room manuals, and as such only GOMS that can be associated with main control room actions were considered. A list of the GOMS procedures as detailed in [1, 2] is provided in Table 1. The association of GOMS, automatically, can be created through a text mining framework.

Table 1. A list of GOMS primitives as defined by [1, 2]. GOMS primitives considered are indicated with **.

Primitive	Description	
Ac	Performing required physical actions on the control boards	**
Af	Performing required physical actions in the field	
Cc	Looking for required information on the control boards	**
Cf	Looking for required information in the field	
Rc	Obtaining required information on the control boards	**
Rf	Obtaining required information in the field	
Ip	Producing verbal or written instructions	**
Ir	Receiving verbal or written instructions	**
Sc	Selecting or setting a value on the control boards	**
Sf	Selecting or setting a value in the field	
Dp	Making a decision based on procedures	**
Dw	Making a decision without available procedures	
W	Waiting	

2 Methods

Data mining is the extraction of meaningful patterns and information from large amounts of data. In the same respect, text mining refers to the process of defining intriguing and relevant conclusions from text [3]. The application of text mining was applied to NPP control room procedures so that a better understanding of the 'procedure' performance shaping factor can be achieved. Seven procedural manuals were acquired from a U.S. NPP [4–10]. The text was captured out of portable document format (PDF) files using the suite of Microsoft products, R 3.2.2 and SAS 9.3 [11, 12].

After the text was pulled from the PDF files, it was formatted into four different levels. These levels are defined by expert HRA analysts and will be referred to as a Levels 1 through 4; an example is provided in Table 2. For the purpose of analysis, the procedure manual was analyzed at the fourth level, because this is where most of the

Table 2. An example of the levels of actions defined in the procedural manual for the NPP main turbine.

Procedural manual text	Levels			
	Level 1	Level 2	Level 3	Level 4
Instructions	6			
Main turbine startup	6	6.1		
Prerequisites	6	6.1	6.1.1	
The feedwater system is in service per feedwater and condensate	6	6.1	6.1.1	6.1.1.1
The main turbine lube oil system is in service per main turbine lube oil system	6	6.1	6.1.1	6.1.1.2
The generator seal oil system is in service per generator seal oil system	6	6.1	6.1.1	6.1.1.3
The main generator is filled with hydrogen per generator hydrogen	6	6.1	6.1.1	6.1.1.4
The stator cooling system is in service per stator cooling system	6	6.1	6.1.1	6.1.1.5
The stator cooling water system trips have been reset per stator cooling system	6	6.1	6.1.1	6.1.1.6

control room instructions are clearly defined. Additionally, the fourth level is the level at which GOMS-HRA most naturally translates. The seven different operation manuals contained more than 2,100 fourth-level procedures. Table 2 is an example of the differing levels as defined in the NPP procedural manual regarding the main turbine.

These procedures were then altered into a format that is easier to text mine. This is completed via the removal of stop words; these are typically conjunctive words that have little meaning when content is analyzed (e.g., "and," "but," "or," and "with"). Then each non-conjunctive word in the manual had the suffix removed; this is called stemming. This is completed so that similar words would be counted as the exact same word. For example, "charge," "charging," "charged," and "charges" would all be defined as differed words before stemming is completed; after stemming, they are all "charg-". In addition to implementing stemming and stop word removal, all punctuation and numbers are removed as software identifies RCP's, "RCPs", and RCPs as different from one another when no content difference exists. An example of stop word removal, stemming, and punctuation removal on a procedural step can be seen in Table 3.

Once the text has been prepared, a text matrix is generated that identifies the number of times a word stem is in a subsection. The seven procedural manuals produced more than 2,000 unique word stems. A bag-of-words approach was taken such that the context of each word was ignored, except in special cases. One such case was due to the frequency use of the term charging pump; it was analyzed as "chargingpump." The context of the word stems is integral, because two different words can mean the same thing (synonymy) and the same word can have two or more meanings in different contexts (polysemy). While this realization exits, it is difficult to quantitatively capture this information. An example of a text matrix with five word stems can be seen in Table 4.

Table 3. An example of stemming, stop word removal, and deletion of numbers and punctuation performed on a procedural manual step.

Before	IF BOTH of the following occur at any time during this procedure: • Pressurizer level lowers to 33% • Restoration of charging is NOT Impending THEN trip the reactor **NOTE:** Multiple indications and SM/CRS discretion should be applied to diagnosing Charging Pump gas binding
After	Follow occur any time procedure pressure level lower restore charge impend trip reactor multiple indic smcrs discrete appli diagnose charging pump gas bind

Table 4. A text matrix with the original procedure and formatted procedure, along with a selection of five stem words and their respective counts.

Original procedure with punctuation	Procedure formatted	Text matrix				
		Action	Charge	Chargingpump	chbhs523	Close
IF BOTH of the following occur at any time during this procedure • Pressurizer level lowers to 33% • Restoration of charging is NOT Impending THEN trip the reactor **NOTE**: Multiple indications and SM/CRS discretion should be applied to diagnosing Charging Pump gas binding	Follow occur any time procedure pressure level lower restor charge impend trip reactor multiple indic smcrs discrete appli diagnose chargingpump gas bind	0	1	1	0	0
IF Charging Pump gas binding is indicated by ANY of the following: • Charging header flow fluctuations • Charging header pressure fluctuations • Charging header flow less than expected for running charging pumps • Charging suction source (VCT, RWT) level lost THEN perform Appendix G, Responding to Gas Binding of Charging Pumps	Chargingpump gas bind indic follow charge header flow fluctuate charge header pressure fluctuate charge header flow less expect run chargingpump charg suction source vct rwt level lost perform appendix G	1	4	2	0	0

Additionally, while the context of word stems was not able to be captured, parts of speech were captured (i.e., noun, verb, and adjective). This was conducted with a hybrid of natural language processing algorithms and excerpts of tables from the Professional Procedure Association's manual [13]. The context of the word was briefly considered as an analytical approach but was not retained due to inaccuracies and time constraints. All of the above techniques were applied to the analysis of the seven procedural manuals, which had more than 2,000 unique word stems for more than 2,100 Level 4 procedures. Thus, more than 4,200,000 observations were considered in matrix form. The most frequent word stems in the 2,100 fourth-level procedures are provided in Fig. 1.

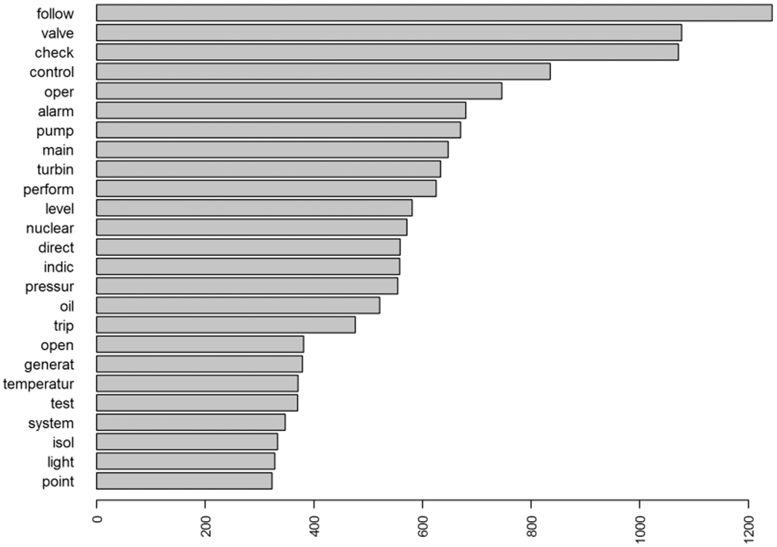


Fig. 1. Bar chart of the top 25 occurring word stems in the seven manuals for fourth-level procedures.

3 Analysis and Results

There are many analysis methods that can be implemented on a text matrix. Due to the large data nature of this analysis, reduction of dimensions or noise is desired. Some methods considered include principal component analysis (PCA), ridge regression, single value decomposition (SVD), and expert judgment [14, 15]. Then analytical methods were further implemented to the text matrix codex to define the GOMS primitives. While all these methods were explored, only the details of PCA, SVD, and expert judgment are detailed herein. To provide meaningful results, a randomly selected subset of 148 of the 2,100 procedures was mapped to GOMS; this created a codex upon which meaningful conclusions can be mapped. The top-occurring word stems are provided in Fig. 2. As such, the analytical methods are applied to the subset

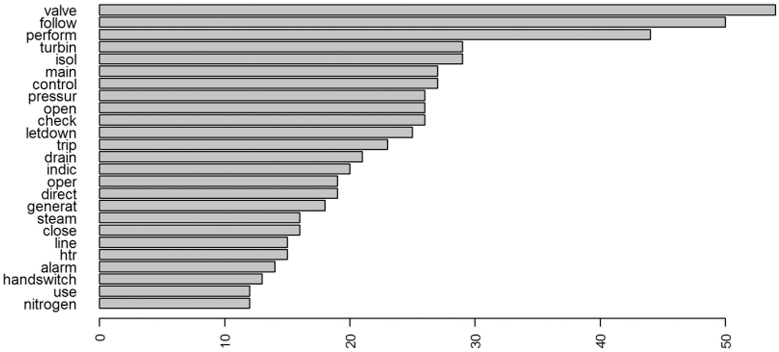


Fig. 2. Bar chart of the top 25 occurring word stems in the GOMS codex of procedures.

of 148 procedures. For the methods to be confirmed as more generalizable, a larger codex needs to be considered.

3.1 Dimension and Noise Reduction

Principal Component Analysis. PCA uses a text matrix of the words to create linear combinations of word stems. These new variables are called Eigen vectors and are orthogonal to one another. The number of Eigen vectors created is equal to the number of variables, or word stems, that are in the initial text matrix. With 33 Eigen vectors, 90% of the variance is explained. A way to visualize the first two Eigen vectors that explain the most variation is provided in a bi-plot in Fig. 3.

The word stems have meaning based on their angle to one another (Fig. 3). The arrows in the figure are at different angles to one another, indicating the level of correlation. When the arrows are at 90°, this indicates orthogonality, or a lack of correlation between word stems. And parallel arrows are considered to be highly correlated. Arrows at 180° from one another are inversely related. Based on this, words like “follow” and “perform” are considered essentially parallel. “Check” and “drain” are 180° from each other, indicating an inverse relationship.

While this method provides informative descriptive statistics and dimension reduction, identifying the stems that are strongly correlated with the GOMS primitives is not straightforward in this form. Thus, other methods were considered for auto calculating GOMS primitives to NPP procedures.

Single Value Decomposition. SVD is a statistical method to reduce the noise of irrelevant variables. SVD describes data by reducing the sum of the difference between the text matrix vectors, the details of which are described in [15]. The positive aspect is that SVD does not overrate the similarity between two words, in contrast to the PCA approach. Unlike other methods, SVD does not automatically toss out highly correlated

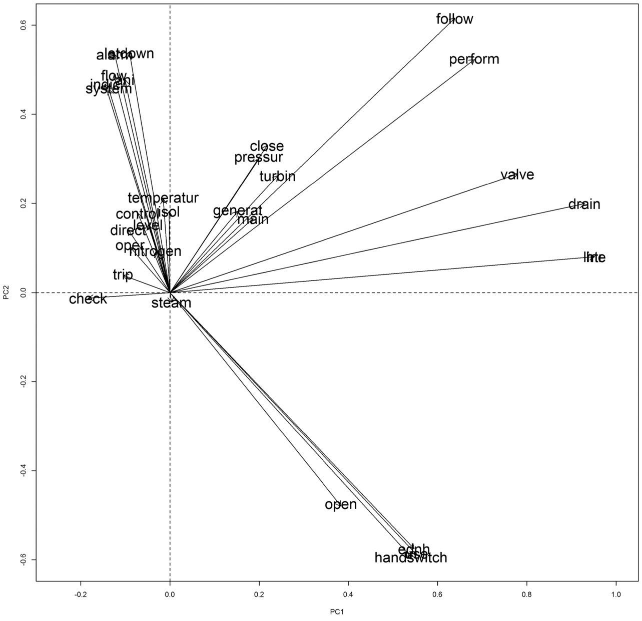


Fig. 3. A PCA bi-plot of the first two Eigen vectors with only the top 30 word stems considered.

word stems. However, the output for SVD is similar to that of PCA and as such was not utilized as a method to reduce dimensions in GOMS.

Expert Opinion. While the previously defined methods for noise reduction in the number of word stems may be more descriptive, their implications are not always straightforward. As such, expert opinion was employed that involved dropping all the word stems that had three or less occurrences. Three was decided upon because it was the median number of occurrences of word stems in the codex. This resulted in 84 word stems remaining in the text matrix. Further dimension and noise reduction was completed using analytical techniques, with unique results being applied to each GOMS primitive type.

3.2 Analysis Methods

The results of the analysis provide word stems that are strongly correlated with GOMS primitives. The methods considered include naive Bayes, random forest, logistic regression, heat map algorithms, Euclidian hierarchical clustering (EHC), correlation networks, and Bayesian discrete discriminant (BDD) analysis. Details from EHC, correlation network, and BDD are provided below.

Euclidian Hierarchical Clustering. The first step to EHC is to calculate the distance matrix by the Euclidian method. The distance matrix provides the distance between two vectors such that it is implemented between the rows of the text matrix [14].

Once the distance between rows is computed, the resulting matrix is considered a matrix of dissimilarities. The distance, or dissimilarity, matrix does not necessarily calculate the literal distance between words and is a way to empirically evaluate the data. The rows of our text matrix are the stem word, so when the dissimilarity matrix is calculated, it is calculating the difference of the procedures based on the frequency of the word stems. This matrix can be represented graphically as a dendrogram, as seen in Fig. 4.

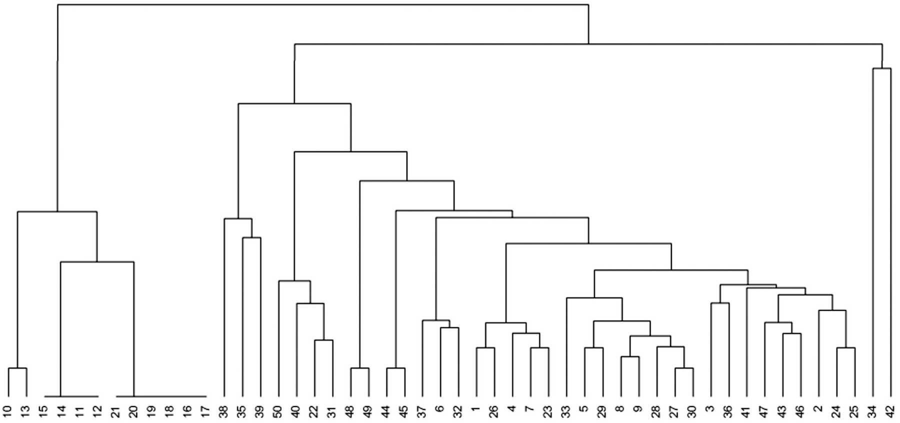


Fig. 4. A Euclidian single cluster dendrogram on the NPP procedures, where the numbers at the bottom are NPP procedures in the codex.

The numbers at the bottom of Fig. 4 are the identification numbers associated with the procedures in the codex. A hierarchical cluster analysis is applied to the dissimilarity matrix for n clusters, where n is defined subjectively by the expert. Based on data configuration, the number of clusters selected is seven, corresponding the number of GOMS that are being investigated. This is then examined against the GOMS groups, which resulted in 11% accuracy. As such, further methods were considered for defining the GOMS type.

Correlation Network. When investigating the dependence between multiple variables, a correlation matrix can be constructed. In this case, the correlation between procedures is being evaluated. The result is a matrix containing the correlation coefficients between each of the procedures. While a matrix contains a lot of information, visualization of that data can be difficult and chaotic. Thus, a network was constructed to better visualize the correlation relationships between the stem words, as in Fig. 5.

The thickness of the lines between the stem word nodes denotes the strength of the correlation. In addition to line thickness, the colors of the lines indicate if the correlation is positive (black) or negative (grey). Oddly enough, there are no strong negative correlations, or thick grey lines, whereas there is a strong positive relationship between

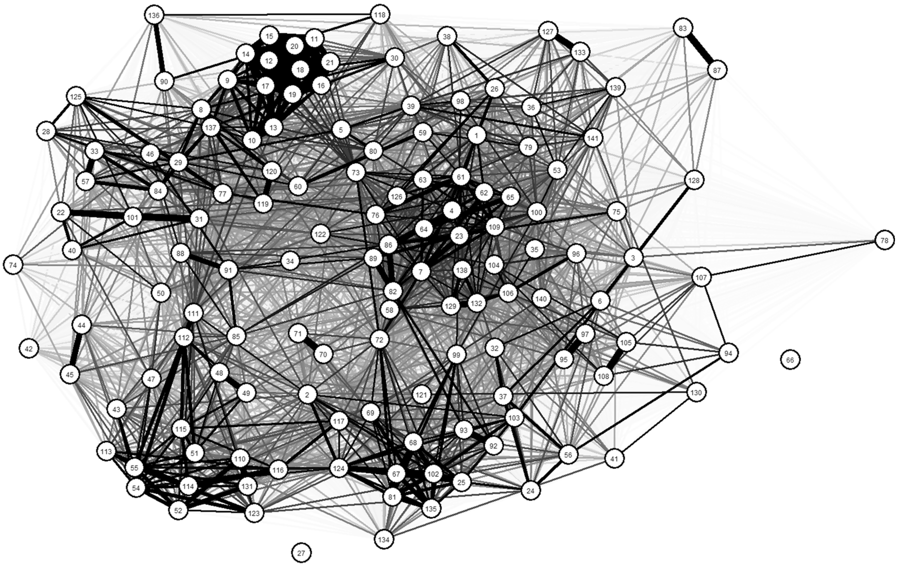


Fig. 5. Network of the word stems based on the correlation matrix. Black indicates a positive correlation, and grey a negative. The nodes, or circles, are the procedures in the codex.

clumps of procedures. These clumps may lend themselves to mapping to the GOMS primitives; however, there only appear to be 4 or 5 clumps at most, while seven GOMS primitives are defined in the codex. As such, another method to define the GOMS primitives was explored.

Discrete Discriminant Analysis. BDD is implemented with the assumption that the frequency of GOMS primitives in the codex is representative of all NPP manuals. Initially, a discrete multinomial distribution of GOMS primitives was assumed; however, this produced low accuracy. Thus, each GOMS primitive was dummy coded and assessed individually, which is in line with expert opinion; the details of discrete discriminant analysis are provided in [16]. Each procedure in an NPP manual may be composed of multiple primitives. A binominal BDD for each GOMS primitive lends itself to identification of multiple GOMS per a procedure.

To further reduce the word stems utilized, stepwise selection based on an Akaike information criterion was applied [17, 18]. Then an algorithm to fit all possible discrete discriminant analysis combinations was executed, with the best performing model defined based on the lowest Akaike information criterion value. The resulting word stems were retained; the accuracy for each GOMS is provided in Table 5. Due to Ir and Sc having such a low frequency in the codex, any results for Ir and Sc are not considered accurate and are not presented.

Table 5. Results from the discrete discriminant analysis for each GOMS primative. The frequency of example procedures are provided along with the analysis accuracy and the word stems that were included in the model.

GOMS primitive	Frequency	Prediction accuracy	Discriminant analysis results
Ac	30	95%	Cool exist manual trbl leak regen ensure high output refer bottle place air test ani level handswitch alarm close trip letdown check control turbine perform valve
Cc	45	88%	Instal low speed gate initi leak run output bottl place action flow system level handswitch close direct trip letdown pressure isol turbine follow valve
Rc	26	94%	Cool cooldown greater instal low gate supply breaker reactor section flow ani steam generate direct drain trip letdown check pressure
Ip	18	95%	Enter smcrs maintain regen auxiliary direct pressure turbine
Ir	5	100%	NOT ACCURATE
Sc	2	94%	NOT ACCURATE
Dp	15	98%	Speed leak lpturbin maintain end loss output rcs refer breaker place section service ani perform follow

4 Results and Conclusions

Text mining, as applied to NPP control room operation manuals, provides a lot of descriptive statistics that can better inform the future development of manuals and error quantification methods. The number of unique word stems, more than 2,000, in NPP control room operations manuals is relatively low compared to other invocations of the English language in everyday life. Experts have suggested that this is because NPP manuals need to be easily understood, even in situations of extreme stress and when English is a second language. Many other interesting findings may still come to light from these documents that will give unique insights to NPP control room interworkings.

Many dimension-reduction methods were employed with the final technique executed, including expert opinion, stepwise selection, and creation of all possible models. Analysis methods for identification of the GOMS primitives to the procedures are accomplished by associating multiple GOMS to a procedure. While the examination only considered the mapping of the one GOMS to procedures, applying a BDD analysis is highly effective with all models, indicating 88% or greater accuracy. To have more accurate results, more examples of GOMS primitive mappings need to be provided so that more generalizable results can be obtained that apply to more than just seven NPP operation manuals.

The highly accurate automation of typing NPP procedures into multiple GOMS primitives is a step toward creating a dynamic framework that can calculate a realistic human error probability in real time. This real-time assessment will be based on the procedures that control-room and field operators implements. Further quantitative research needs to be completed describing the event trees and the other associated performance shaping factors that will have an impact on a control room operator's ability to complete tasks.

Acknowledgments. Every effort has been made to ensure the accuracy of the findings and conclusions in this paper, and any errors reside solely with the authors. This work of authorship was prepared as an account of work sponsored by Idaho National Laboratory, an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Idaho National Laboratory is a multi-program laboratory operated by Battelle Energy Alliance LLC for the United States Department of Energy under Contract DE-AC07-05ID14517.

References

1. Boring, R., Rasmussen, M., Ulrich, T., Ewing, S., Mandelli, D.: Task and procedure level primitives for modeling human error. In: Proceedings of the 8th Applied Human Factors and Ergonomics, Los Angeles (2017, in press)
2. Ulrich, T., Boring, R., Ewing, S., Rasmussen, M., Mandelli, D.: Operator timing of task level primitives for use in computation-based human reliability analysis. In: Proceedings of the 8th Applied Human Factors and Ergonomics, Los Angeles (2017, in press)
3. Gupta, V., Lehal, G.S.: A survey of text mining techniques and applications. J. Emerg. Technol. Web Intell. **1**, 60–76 (2009)
4. U.S. NPP Generating Station, "EXCESSIVE RCS LEAKRATE"
5. U.S. NPP Nuclear Generating Station, "STANDARD POST TRIP ACTIONS"
6. U.S. NPP Nuclear Generating Station, "The Steam Generator Tube Rupture"
7. U.S. NPP Nuclear Generating Station, "Panel 6 Alarm Responses"
8. U.S. NPP Nuclear Generating Station, "Panel 7 Alarm Responses"
9. U.S. NPP Nuclear Generating Station, "Loss of Charging or Letdown"
10. U.S. NPP Nuclear Generating Station, "Main Turbine"
11. R Core Team: R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2016)
12. SAS Institute Inc: Base SAS® 9.3 Procedures Guide. SAS Institute Inc., Cary, NC (2011)
13. Procedure Professionals Association: Procedure Writer's Manual PPA AP-907-005 R2 (2016)
14. Murtagh, F.: Multivariate Data Analysis with Fortran, C and Java Code. Queen's University Belfast, and Astronomical Observatory Strasbourg, Belfast
15. Albright, R.: Taming Text with the SVD. SAS Institute Inc, Cary (2004)
16. Knoke, J.D.: Discriminant analysis with discrete and continuous variables. Biometrics **38**(1), 191–200 (1982)
17. Hocking, R.R.: A biometrics invited paper. the analysis and selection of variables in linear regression. Biometrics **32**(1), 1–49 (1976)
18. Beal, D.J.: Information criteria methods in SAS for multiple linear regression models. In: 15th Annual South East SAS Users Group (SESUG) Proceedings, South Carolina (2007)

The Virtual Human Reliability Analyst

Martin Rasmussen[1](✉), Ronald Boring[2], Thomas Ulrich[3],
and Sarah Ewing[2]

[1] NTNU Social Research, Studio Apertura,
Dragvoll Allé 38 B, 7491 Trondheim, Norway
martin.rasmussen@ntnu.no
[2] Idaho National Laboratory, P.O. Box 1625, Idaho Falls, ID 83415-3818, USA
{ronald.boring, sarah.ewing}@inl.gov
[3] Universitity of Idaho, 709 S Deakin Street, Moscow, ID 83844, USA
thomas.ulrich@inl.gov

Abstract. This paper introduces the virtual human reliability analyst model (VHRAM). The VHRAM is an approach that automates the HRA process to enable HRA elements to be included in simulations in general and simulation based risk analysis in particular. Inspirations from clinical AI and game development are discussed as well as the possibilities for a VHRAM to be used outside of a simulated virtual twin of a nuclear power plant.

Keywords: Human reliability analysis · Computation-Based human reliability analysis · Dynamic human reliability analysis · Virtual analyst · Virtual human reliability analysis model

1 Introduction

Through forty years, and at least that many methods, human reliability analysis (HRA) has been used to analyze, explain and predict the human element of complex systems that hold a potential for major accidents. HRA originated in the weapon assembly industry [1], but the nuclear power industry has been the front-runner in both method development [2–4] and application through most of HRA history [1, 5, 6]. Indeed, other industries have been urged to look towards the nuclear domain for guidance on how they have used HRA to analyze the human aspect of major accident risk (i.e., the petroleum industry after the 2011 Macondo accident, [7]). When other domains have adapted HRA methodology to their needs, the starting point has often been nuclear application intended methods [8–10]. This paper discusses another form of HRA adaptation—not the adaptation from one industry to another, but rather from static paper-based HRA traditionally conducted by an analyst completing worksheets concerning a physical system to computer-based HRA in a virtual simulation of a physical system, i.e., a “virtual twin.” A virtual twin (also known as digital twin, [11]) is a virtual representation created to function similarly as the physical system it is modelled after, and as the name implies it strives to be very similar to the original system. A virtual twin is used for tasks such as design changes, simulation, monitoring or optimization. Depending on what is being evaluated, the virtual twin can include

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_25

aspects such as the physical measurements and placement of all the components of a system, interactions between the components, process simulation and physics engines. The possibilities of how virtual twins can be used increases as improved computational power is continuously enabling new possibilities to accurately and realistically simulate complex systems (using tools like RAVEN [12] and RELAP-7 [13]). These simulations can among other things be used to increase the understanding of the risks at a nuclear power plant (NPP) through simulating expected plant states over thousands of years or simulating a complex scenario thousands of times. The human element has, however, not been a key element in these simulations despite the fact that control room operators have an important role in both normal operations and particularity in accident scenarios. This paper presents the idea of taking some of the lessons learned from traditional static HRA and using them to capture the human element in computation-based simulations of the same types of complex systems where HRA has been used thus far.

2 Differences between Traditional HRA and Computer-Based HRA: Opportunities and Challenges

There are many differences between a traditional HRA conducted on a physical installation and the proposed use on a virtual twin of the installation. The main difference is the presence or absence of a human reliability analyst. It is not practical to introduce a person to make manual decisions at each iteration of a simulation. This would be a very resource-demanding task, especially if the simulation is set to analyze the same scenario thousands of times, or simulate a plant state over many years. To avoid this issue, the tasks of the human reliability analyst must be automated, or in other words, the creation of a virtual human reliability analyst model (VHRAM).

While this will enable the coupling with a plant simulation, it will also create some challenges in the use of existing HRA methods, as most of the existing HRA methods have relied heavily on the subjective evaluations of the human reliability analyst [14–16]. However, inter-analyst variability—whether caused by subjective biases or a poor fit of methods to events—serves as a major limitation in conventional HRA. Even if a static HRA method is dynamicized, it is possible to create a VHRAM that uses the method consistently. Subjective assessments can be minimized and replaced by consistent and replicable virtual analyses. For example, an HRA method that models task complexity based on an analyst's subjective assessment of the level of task complexity can instead be made to autocalculate the level of task complexity based on available parameters of the plant, task, and situation [14–16].

There are several additional advantages of the VHRAM approach. A classic HRA problem has been that the HRA efforts have been performed after most of the risk analysis is already conducted, including how the scenario develops. This leaves less possibility to investigate how human actions would influence the evolution of a scenario rather than following the predefined path outlined during the risk analysis efforts. An automated model on the other hand could feed back into the plant simulation influencing how the scenario develops. Choices made by a human in the NPP can have an extensive effect on how a scenario unfolds, and this should also be the case in a simulation. To support the examination of the human actions, the simulation must be capable of supporting a dynamic scenario in which operator actions can alter the course of the scenario as it develops. An additional advantage of the VHRAM approach is reducing the subjective element from the analyst. The VHRAM approach supports a more standardized method for inputting human error probability (HEP) quantifications for operator actions.

HRA traditionally uses rather simple mathematical formulas to calculate an HEP, often using a version of nominal HEP multiplied with performance shaping factors (PSFs) that degrade or improve performance [3, 4, 8]. The simple formulas are suitable for the worksheet analysis conducted by hand and provide a high level of traceability, as it is easy to see where and why a human action is predicted to fail. The automation of the human reliability analyst will reduce the need for a simple formula, enabling the possibility to include aspects such as interactions between PSFs, dependencies between tasks, and continuous PSF levels, all of which have been included in few of the traditional HRA methods. This is not to say that the quantification approach used in traditional HRA should or will be discarded, but computerized HRA will have the possibility to refine the quantification approach if it can be shown that it improves the method and reduces epistemic uncertainty in the analysis.

3 Intelligent Agents

The absolute simplest version of automating the HRA process conducted by a human reliability analyst would be to use a nominal HEP or non-informed HEP distribution for all human tasks. This is an approach that has been used in technical quantitative analyses without HRA and in analyses where the goal is to determine if the human has a critical role in the scenario by setting the HEP to 1.0 (or close) to find the likelihood of an accident if human actions fail. However, the introduction of a static HEP for all human tasks, while simple, does not seem to be on par with the high fidelity modeling of the other systems in a plant simulation [17].

On the other extreme, we have a simulated system including a model of human cognition with the ability to perform similarly to a real human operator with all the variability and creativity a human can express. However, artificial intelligence (AI) technology has not yet come to the point where this is entirely feasible.

In the VHRAM we attempt to find a suitable middle ground between these two extremes. Instead of attempting to model the entire cognition of an operator we are instead trying to create a model that evaluates the situation the operator would find himself or herself in by including the most important PSFs. This will create a representation much like the one a human reliability analyst would create using worksheets to evaluate the situation of the operator, instead of attempting to model his or her cognition. To emphasize this focus we have chosen the term VHRAM and not virtual operator model, where it might have seemed that we were attempting to model the full cognition of an operator.

3.1 AI and the Concept of the Intelligent Agent

The idea of an intelligent agent performing human-like actions within a system, for different purposes and with different levels of sophistication is not new and has been explored in several fields. Throughout history many myths and philosophers describe the idea of an inanimate object obtaining a mechanical version of human intelligence. A more direct AI reference is found in the works of Alan Turing including the test of the sophistication of an intelligent agent in the famous Turing Test [18]. AI research generally represents the upper range of this sophistication in the creation of systems that either think rationally and/or emulate human thought [19], but in fact most of computer programming does in some way fit within this sophistication scale through a version of rule based commands executed via provided inputs.

A practical application is seen in several fields where the goal is for the intelligent agent to assist, or even replace, the human performing the task. Several examples of this type of intelligent system are seen in our everyday lives. Google attempts to understand our search phrases and deliver the results we desire, Netflix attempts to anticipate what we want to watch, and online advertisements are personalized in an attempt to increase the chance of users viewing and clicking them. We also find examples of intelligent agents in fields that are traditionally unassociated with academia, such as computer games [20], and the methods used to create these intelligent agents, that aren't in themselves full-blown AI, should not be discounted.

Academic AI. AI is naturally an area of interest for computer science and several engineering disciplines as these are areas where advancements in AI primarily occur. However, it has also created interesting academic discussions in the fields of neurology, psychology and philosophy. Our knowledge of the brain and human cognition expands daily, but we are far away from a complete understanding, of what some, intriguingly and somewhat paradoxically using their human cognition, have described as the most complex system known to the human race. While we have many different models of memory, cognition, intelligence and consciousness, they are all simply models—a simplification of how we understand an abstract and complicated concept. Interestingly, thus far the field of AI—the field that strives to build intelligent entities [19]—are faced with the challenge of creating something which we do not yet fully understand. Some of these questions are outside of the scope of this paper, but it is interesting to note how many different fields are involved in the AI topic and the potential that AI research holds to contribute to all of these topics as the field develops.

Clinical AI. Within medicine, both the desire to increase accuracy of diagnosis and the work towards lowering medical costs has led to many different versions of intelligent agents through wide range of different methods (e.g. [21–23]). A version of a non-disease-specific clinical AI was created using a combination of a Markov decision process and dynamic decision networks [21]. The AI used a combination of existing clinical data and simulation of sequential decisions paths to develop plans that would both reduce patient costs and increase patient outcomes.

While there are naturally many differences between the tasks of a medical doctor and a human reliability analyst, there are similarities in how an intelligent agent could be structured. The way a medical doctor considers symptoms in the diagnosis of a patient is similar to how a human reliability analyst considers PSFs to diagnose how a control room operator is expected to perform.

A clinical AI will generally be created to diagnose as flawlessly as possible without realistically mimicking the errors a doctor will occasionally do. Recreating such failures might actually be of particular interest to human reliability researchers, but this is not a mainstream thrust of AI research. Rarely do we design systems to fail intentionally and this goal may be a unique aspect of HRA research.

Interaction and Cooperation with Automation. Another interesting practical use is the partial automation of a role previously performed by a person. This is already a part of most computerized systems. Set criteria, such as the temperature reaching a certain level, are made to trigger certain actions, such as the opening of a valve. These are generally taken for granted as part of a computerized system. However, once the system is intelligent enough to consider a large amount of factors before deciding or suggesting to open a valve, we are approaching the cooperation between the operator and an intelligent agent. In some cases the degree of automation in systems have reached such a high degree that work-roles that previously were manual now mainly consist of monitoring an automated system.

Human-automation interaction or human-automation cooperation has become one of the popular topics in human factors. A review of all papers published in 2015–2016 in the journal *Human Factors,* [24] found automation, including both human-automation interaction and cooperation, to be the third most popular topic and only beaten by driving and physical workload. Though driving, in at least a few instances, overlaps with human-automation interaction.

Gaming AI. The medical field is known for their meticulous efforts in recording and publishing progress and research. However, game development is on the other side of the scale where knowledge and skills are generally transferred through other more informal channels (or kept as trade secrets within the company) [20].

Introducing simplified AI to games has been an important part of game design ever since Pac-Man implemented intelligent agents (often referred to as non-playable character (NPCs) in games) through non-playable characters that chased and ran away from the player [20]. Earlier games, such as Space Invaders have NPCs in the form of enemy ships, but Pac-Man included decision making where the enemies chose a route at each junction, combining an effort to achieve their goal—which, depending on the situation, meant chasing or escaping—and an element of randomness to keep things interesting [20]. This was done through a simple set of rules and a random number generator. Later games added aspects such as perception to their NPCs, as exemplified in Metal Gear Solid and Goldeneye. This perception provided each NPC with limited knowledge about what was going on in the game and would only react after they “saw” or “heard” something. Another interesting AI element, strategic AI, was introduced at about the same time, where the NPCs would employ a number of different strategies to defeat or cooperate with the player [20]. Since the introduction of these cognition-like elements, new and improved versions have been created to suit the need for each individual game. Today the quality of the AI is often a highlighted aspect in modern day videogame reviews, which can result in the monetary success or failure of a videogame launch.

Although there are naturally many differences between the purpose of AI elements in games and what we are trying to achieve with the VHRAM there are certainly overlaps. The inclusion of a human element in a simulated scenario is in many ways the same as including a game character in a game world. A decision making system is required, the VHRAM should base its decisions on the information it has observed, and it should follow a strategy. Perhaps the largest difference is that we do not intend to introduce a human player to the system, rather let the VHRAM "play" by itself in the virtual world.

4 Implementing the Virtual Analyst

The VHRAM is still in development, and changes can still occur in both the general solution and the details of how we have chosen to include it. The symbolic approach to AI describes it as being made up of two components, knowledge and reasoning [20]. Knowledge is often a database, and reasoning is how this knowledge is used.

In a similar manner, the VHRAM will consist of two main aspects: (1) relevant information prepopulated though task categorization and autopopulated from the simulation—knowledge—and (2) the algorithms—reasoning—that use the inputs to determine the HEP, time spent on the task, and the decisions on which path to take in a dynamic scenario.

4.1 Autopopulation

The autopopulated input is automatically gathered from the information already present in the simulation. Examples of the autopopulation are [14, 15]:

- Total size of the task or scenario
- Number of tasks per time
- Time in scenario
- Number of procedures used by the operator
- Number of page shifts done by the operator in the procedures

As the VHRAM improves it is likely that more and more aspects are included as autopopulated inputs. However, it is also likely that some information that could be relevant to human reliability will not be available in the simulation, such as the human-machine interface quality or teamwork problems. If specific aspects like these are focus areas of an analysis it should be possible to inform the model through prepopulated factors connected to either the scenario or specific tasks.

4.2 GOMS-HRA

In many of the traditional HRA methods it is not specified to which level a task should be decomposed before it is quantified [25]. Depending on the method and situation, quantification could be done anywhere from on a very high level (e.g. Depressurize segment A) to a very low level (Press button A). Methods that do not specify the level at which quantification should occur will have more flexibility, but as the quantification level can influence the results, it is also a source of lower reliability [25].

As the VHRAM is an automated approach it was decided that it would quantify at a low level, a level defined as the subtask level. GOMS-HRA [26–28] is currently in development as a method for standardizing tasks at a subtask level. The subtask level of analysis is suitable for modeling time series activities of operators in dynamic HRA. GOMS-HRA provides task level primitives, which are meant to be universal types of tasks performed by humans. Because activities in NPPs are generally highly proceduralized, it is also possible to map procedure steps to their underlying task GOMS level primitives [29].

4.3 HEP Equation

The traditional output of an HRA in the evaluation of a task (in addition to any qualitative descriptions and recommendations) is the HEP. The use of the term "human error" is controversial in human factors and safety research [30]. Some argue that the term implies that the human is to blame for the error [31], others that the term is misleading, as the actions made by the operator can be reasonable to the operator at the time only to be considered an error retrospectively [32]. In HRA the term HEP is simply the probability that the operator will not continue on the intended path that avoids an accident from occurring, without blaming the operator for the mistake. In fact, as most of the PSFs included in many HRA methods (e.g. [3, 8]) are factors external to the operator HRA is often mainly concerned about which external factors could cause the operator to fail.

Currently the VHRAM is based on a stochastic multiple regression with each input as a variable. In the future hopefully empirical data, from simulators or actual installations, can be used either to calibrate the coefficients of the stochastic multiple regression equation, or modify the approach if a more suited model is found.

4.4 Decision Making

The HEP value can be used as a simple form of decision making, through having human error occur at the probability calculated and have the scenario developed based on this. This would however limit the decision making to a binary success or failure for each junction. A dedicated decision making algorithm will enable more nuanced decisions which can include more than two outcomes at each junction.

Several different forms of decision making algorithms, like those seen in both clinical AI and game AI, are being considered at the moment to be able to fully integrate the VHRAM into a dynamic scenario where it can contribute to the evolution of the scenario.

4.5 Including PSFs

The approach of the VHRAM is to start with a simple version and build upon that to include more aspects. The first PSF that was introduced to the model was complexity [14, 15]. Complexity is included in most HRA methods as part of the quantification leading to the HEP [33]. This fits well with our intuitive understanding of complexity and the role it can have in the likelihood of successfully conducting a task. The fact that complexity is a multifaceted concept also means that while it is often modeled as a single PSF it has many different aspects where the inputs can be collected from several different parts of the simulation.

Currently a second PSF, procedures, is being modeled for autopopulation and inclusion to the VHRAM. Procedures will in the same way as complexity, to inform the model with aspects that are included in HEP calculations and the decision making algorithm. However, procedures also hold another very interesting potential. If the VHRAM includes a text-mining approach that can break down procedures into a standardized unit size (such as GOMS-primitives) they can serve as an input directly to the VHRAM [26–29]. This would be an important step in the direction of a model that can run automatically on any scenario where procedures exist.

5 The Way Forward

The way forward for the VHRAM is to continue adding new elements and improving its performance as an automatic human reliability analyst. It is a promising path of research, but there are still challenges that need to be solved. The potential value will depend on the quality of the VHRAM, but also the quality of the virtual twin. In an attempt to create a virtual twin, attempts are made to model every aspect of a system virtually. Naturally, in a complex system there will always be discrepancies between the actual system and the virtual twin. As this discrepancy increases, the relevance of a VHRAM, and other risk analysis performed using the virtual twin, will naturally drop in terms of what you can learn about the real system.

This paper has chosen to describe two examples, clinical AI and game development AI. These were not chosen randomly; rather, they both represent aspects that we want to include in the VHRAM approach. In clinical AI an intelligent agent is created to learn from clinical data and treatment procedures. We want to include this diagnostic element but the clinical data is replaced by empirical or simulated plant and operator performance data and treatment procedures are replaced by operating procedures. We also want to include a decision making algorithm much like the ones developed for the clinical AI applications. The primary difference between the clinical AI and the work here is that the VHRAM model will include the simulation of human error as one of the key aspects. The inspiration for the human error element stems from other fields that developed intelligent agents with inherent limitations as to how well they can perform. For entertainment purposes, an intelligent agent opponent in a game has to provide the player with a challenge, without performing so well that the player is without a chance to win. A chess match between a human and the chess computers of today would not be entertaining, nor would neither a soccer game where every shot made by the opponent is a goal, nor a shooting game where the opponent shoots you repeatedly through the walls. The gaming industry has dealt with these challenges for many years and they provide valuable guidance for how these elements can be included in a simulation of the human component in HRA research. The simulated human has to perform realistically, but that also means it needs to fail realistically, which represents a prominent challenge.

In the future, there could also be other uses for a VHRAM than HRA of a virtual twin. One potential use could be for a combined approach between traditional HRA and VHRAM where the aspects that are autopopulated by the VHRAM could be used as part of the information collected by the analyst conducting the traditional HRA. Another possibility is that a VHRAM is running in real-time at a NPP anticipating when the actual operator will encounter a situation where the PSFs are implying that he or she has an increased chance of making a mistake, as a type of risk monitoring system.

6 Conclusion

This paper presented the ideas around the ongoing development of the VHRAM. We believe it is an approach that will have value through adding a human component to probability risk analysis simulations, and other forms of simulations, where it has been historically under-represented thus far. Furthermore, it is an approach that can have impacts outside of this field by contributing to traditional HRA and risk monitoring systems in physical systems, such as NPPs.

Acknowledgments. This paper was written as part of the Risk Informed Safety Margin Characterization (RISMC) research pathway within the U.S. Department of Energy's Light Water Reactor Sustainability (LWRS) program that aims to extend the life of the currently operating fleet of commercial nuclear power plants. The research presented in this paper aims to feed into the current work on the CoBHRA approach (previously abbreviated CBHRA) called Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER; [16, 34, 35]).

References

1. Boring, R.L.: Fifty years of THERP and human reliability analysis. In: Proceedings of the PSAM 11 and ESREL, Helsinki (2012)
2. Swain, A.D., Guttmann, H.E.: Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications. Final report. U.S. NRC, Washington, DC (1983)
3. Gertman, D.I., Blackman, H.S., Marble, J.L., Byers, J.C., Smith, C.: The SPAR-H Human Reliability Analysis Method. U.S. NRC, Washington, DC (2005)
4. Williams, J.C.: A data-based method for assessing and reducing human error to improve operational performance. In: Conference Record for 1988 IEEE Fourth Conference on Human Factors and Power Plants, pp. 436–450. IEEE (1988)
5. Spurgin, A.J.: Human Reliability Assessment: Theory and Practice. CRC Press, Boca Raton (2010)
6. Rasmussen, M.: The development of performance shaping factors for the PetroHRA method: a human reliability method for the petroleum industry. Doctoral dissertation, Norwegian University of Science and Technology, Trondheim, Norway (2016)

7. Report to the President National Commission on the BP Deepwater Horizon Spill and Offshore Drilling: Deep Water: The Gulf Oil Disaster and the Future of Offshore Drilling. Washington, DC (2011)
8. Bye, A., Laumann, K., Taylor, C., Rasmussen, M., Øie, S., van de Merwe, K., Øien, K., Boring, R.L., Paltrinieri, N., Wærø, I., Massaiu, S., Gould, K.: The Petro-HRA Guideline (IFE/HR/F-2017/001). Institute for Energy Technology, Halden
9. Kim, J.W., Jung, W., Jang, S.C., Wang, J.B.: A case study for the selection of a railway human reliability analysis method. In: International Railway Safety Conference, 22–27 October 2006, Belfast (2006)
10. Kirwan, B., Gibson, H.: CARA: a human reliability assessment tool for air traffic safety management — technical basis and preliminary architecture. In: Proceedings of the 15th Safety-Critical Systems Symposium, Bristol, 13–15 February 2007, Bristol (2007)
11. Glaessgen, E.H., Stargel, D.S.: The digital twin paradigm for future NASA and U.S. air force vehicles. In: 53rd Structures, Structural Dynamics, and Materials Conference: Special Session on the Digital Twin, pp. 1–14, Honolulu (2012)
12. Rabiti, C., Alfonsi, A., Mandelli, D., Cogliati, J., Martineau, R., Smith, C.L.: Deployment and Overview of RAVEN Capabilities for a Probabilistic Risk Assessment Demo for a PWR Station Blackout Deployment and Overview of RAVEN Capabilities for a Probabilistic Risk Assessment Demo for a PWR Station Blackout (INL/EXT-13-29510). Idaho National Laboratory, Idaho Falls (2013)
13. Anders, D., Berry, R., Gaston, D., Martineau, R., Peterson, J., Zhang, H., Zhao, H., Zou, L.: RELAP-7 Level 2 Milestone Report: Demonstration of a Steady State Single Phase PWR Simulation with RELAP-7 (INL/EXT-12-25924 RELAP-7). Idaho National Laboratory, Idaho Falls (2012)
14. Rasmussen, M., Boring, R.L.: The implementation of complexity in computation-based human reliability analysis. In: Walls, L., Revie, M., Bedford, T. (eds.) Risk, Reliability and Safety: Innovating Theory and Practice. Proceedings of the 26th European Safety and Reliability Conference, ESREL 2016, Glasgow, Scotland, 25–29 September 2016. CRC Press, Croydon (2017)
15. Rasmussen, M., Boring, R.L.: Auto-Calculation of Complexity in Computation-Based Human Reliability Analysis (in review)
16. Boring, R.L., Mandelli, D., Rasmussen, M., Herberger, S., Ulrich, T., Groth, K., Smith, C.: Light Water Reactor Sustainability Program Integration of Human Reliability Analysis Models into the Simulation- Based Framework for the Risk- Informed Safety Margin Characterization Toolkit (INL/EXT-16-39015). Idaho National Laboratory, Idaho Falls (2016)
17. Boring, R.L., Joe, J.C., Mandelli, D.: Human performance modeling for dynamic human reliability analysis. Lect. Notes Comput. Sci. **9184**, 223–234 (2015)
18. Turing, A.M.: Computing machinery and intelligence. Mind. **59**, 433–460 (1950)
19. Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice-Hall International, London (1995)
20. Millington, I., Funge, J.: Artificial Intelligence for Games. CRC Press, Boca Raton (2016)
21. Bennett, C.C., Hauser, K.: Artificial intelligence framework for simulating clinical decision-making: a markov decision process approach. Artif. Intell. Med. **57**, 9–19 (2013)
22. Bennett, C.C., Doub, T.W.: Temporal modeling in clinical artificial intelligence, decision-making, and cognitive computing: empirical exploration of practical challenges. In: Proceedings of the 3rd SIAM Workshop of DMMH, Philadelphia (2014)
23. Bellazzi, R., Zupan, B.: Predictive data mining in clinical medicine: Current issues and guidelines. Int. J. Med. Inform. **77**, 81–97 (2006)

24. Laumann, K., Rasmussen, M., Boring, R.L.: A literature study to explore empirically: what is the scientific discipline of human factors and what makes it distinct from other related fields. In: Boring, R.L. (ed.) Advances in Human Error, Reliability, Resilience, and Performance. AISC, vol. 589, pp. 63–73. Springer, Cham (2017)
25. Rasmussen, M., Laumann, K.: The impact of decomposition level in human reliability analysis quantification. In: Walls, L., Revie, M., Bedford, T. (eds.) Risk, Reliability and Safety: Innovating Theory and Practice. Proceedings of the 26th ESREL 2016, Glasgow, Scotland, 25–29 September 2016. CRC Press (2017)
26. Boring, R.L., Rasmussen, M.: GOMS-HRA: a method for treating subtasks in dynamic human reliability analysis. In: Walls, L., Revie, M., Bedford, T. (eds.) Risk, Reliability and Safety: Innovating Theory and Practice. Proceedings of the 26th European Safety and Reliability Conference, ESREL 2016, Glasgow, 25–29 September 2016. CRC Press (2016)
27. Boring, R., Rasmussen, M., Ulrich, T., Ewing, S., Mandelli, D.: Task and procedure level primitives for modeling human error. In: Boring, R.L. (ed.) Advances in Human Error, Reliability, Resilience, and Performance. AISC, vol. 589, pp. 30–40. Springer, Cham (2017)
28. Ulrich, T., Boring, R., Ewing, S., Rasmussen, M.: Operator timing of task level primitives for use in computation-based human reliability analysis. In: Boring, R.L. (ed.) Advances in Human Error, Reliability, Resilience, and Performance. AISC, vol. 589, pp. 41–49. Springer, Cham (2017)
29. Ewing, S.M., Boring, R.L., Rasmussen, M., Ulrich, T.: Text mining for procedure-level primitives in human reliability analysis. In: Boring, R.L. (ed.) Advances in Human Error, Reliability, Resilience, and Performance. AISC, vol. 589, pp. 239–249. Springer, Cham (2017)
30. Reason, J.: A Life in Error: from Little Slips to Big Disasters. Ashgate Publishing Limited, Surrey (2013)
31. Whittingham, R.B.: The Blame Machine: Why Human Error Causes Accidents. Elsevier Butterworth-Heinemann, Oxford (2004)
32. French, S., Bedford, T., Pollard, S.J.T., Soane, E.: Human reliability analysis: a critique and review for managers. Saf. Sci. **49**, 753–763 (2011)
33. Rasmussen, M., Standal, M.I., Laumann, K.: Task complexity as a performance shaping factor: a review and recommendations in standardized plant analysis risk-human reliability analysis (SPAR-H) adaption. Saf. Sci. **76**, 228–238 (2015)
34. Boring, R.L., Mandelli, D., Rasmussen, M., Herberger, S., Ulrich, T., Groth, K., Smith, C.: Human unimodel for nuclear technology to enhance reliability (HUNTER): a framework for computational-based human reliability analysis. In: Proceeding of PSAM 13, 2–7 October 2016, Seoul (2016)
35. Boring, R.L., Mandelli, D., Joe, J., Smith, C., Groth, K.: Light Water Reactor Sustainability Program: A Research Roadmap for Computation-Based Human Reliability Analysis (INL/EXT-15-36051). Idaho National Laboratory, Idaho Falls (2015)

A Dynamic Mechanistic Model of Human Response Proposed for Human Reliability Analysis

Yunfei Zhao(✉) and Carol Smidts

Department of Mechanical and Aerospace Engineering,
The Ohio State University, Columbus, OH 43210, USA
{zhao.2263, smidts.1}@osu.edu

Abstract. A dynamic mechanistic model of human response for human reliability analysis is proposed in this paper. The model is comprised of three main components: information perception, knowledge base, and human reasoning and decision making. The activation based approach, which considers both stimulus-driven and goal-directed activations, is adopted model the impact of information perception on the reasoning process. An operator's knowledge base is represented by a semantic network in the model. Activation propagation theory is applied to model human reasoning and decision making. Illustration of activation propagation through the relief-valve-stuck-open incident in the Three Mile Island accident demonstrates the feasibility of this approach. Additionally, the influences of two significant performance shaping factors, stress and fatigue, are integrated into the model.

Keywords: Human reliability analysis (HRA) · Mechanistic model · Information perception · Human reasoning and decision making · Activation propagation

1 Introduction

In nuclear power plants (NPP) and other complex systems such as air traffic control systems, human response plays an important role in mitigating damages in the case of an incident or accident. Thus, human reliability has been an active area of research in nuclear power plants, and a number of methods have been proposed to characterize human response. These methods can be classified into two categories. The first uses a set of performance shaping factors (PSFs) or performance influencing factors (PIFs) to represent the context within which human response takes place, and then to adjust the base-line probability of human error. Examples of this type of methods include THERP [1] and SPAR-H [2]. Although some new techniques, such as fuzzy sets [3], have been introduced to improve these methods, the framework remains the same. The second category of methods tries to build an integrated mechanistic model of human response. The model includes how a person perceives information, how he/she diagnoses events, and how he/she makes decisions. The models lay their foundation on results of state-of-the-art research in psychology, human behavior, human cognition, etc.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_26

The representative of this type of methods is the IDAC (Information-Decision-Action-Crew) model [4]. It is easy to see that, compared with the first category of methods, mechanistic models analyze human reliability at a much higher level of detail. Therefore, mechanistic models are capable of providing the researcher more detailed information about human response, such as how and why an operator behaves in a particular way.

As a representative of mechanistic models of human response, the IDAC model has evolved through various generations of models. The initial model is called IDA [5, 6]. It contains three modules: information module, problem solver/decision maker, and action module. For information handling, it includes different types of filters that may prevent the operator from receiving incoming alarms, such as an external filter and an internal filter. For problem solving and decision making, the IDA model identifies three high level goals and eight strategies. Based on these initial efforts, some new insights were introduced to improve the capability of the model. For example, in the IDAC model introduced in [4], modeling of the response of a crew rather than an individual operator was included, and a more comprehensive hierarchical structure of PSFs was provided. In the latest version of the IDAC model [7], the focus of the design shifted to a more mechanistic modeling perspective. Efforts in several important areas were made, such as representation of an operator's knowledge base using a semantic network and improved modeling of human reasoning.

This paper takes advantage of the progresses achieved in human reliability analysis so far, but extends the research further by leveraging results obtained in studies of psychology, human cognition, etc. The model proposed in this paper includes information perception, human knowledge base, and human reasoning and decision making. Information surrounding a person is perceived based on the activation level of each piece of information, which combines the salience of the information and the person's attention on the information. A person's knowledge base is represented by a semantic network, which is comprised of the concepts about the system of interest and the relationships between the concepts. In response to the perceived information, human reasoning and decision making will be triggered by way of activation propagation in the knowledge base. At the same time, the distribution of an operator's attention, which has been shown to have a significant influence on human reasoning and decision making, is calculated as the activation level of each concept in the knowledge base.

The structure of this paper is as follows. Section 2 describes the framework of the proposed model of human response. Sections 3 and 4 include details on information perception as well as human reasoning and decision making. Discussion and concluding statements are found in Sect. 5.

2 Framework of the Model

The framework of the mechanistic model of human response is shown in Fig. 1. In addition to the main components in the IDAC model, the mechanistic model also refers to the model of situation of awareness in dynamic decision making, which was proposed by Endsley [8], and the human reasoning model, which was proposed by Guarino et al. [9]. As stated in Sect. 1, the model is comprised of three main

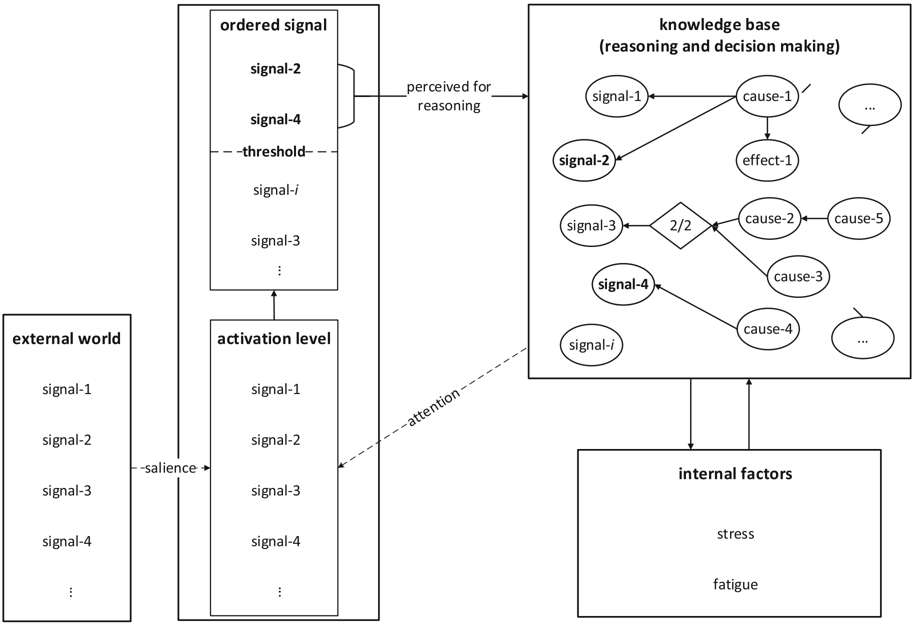


Fig. 1. Framework of the mechanistic model.

components: information perception, knowledge base, and human reasoning and decision making. Each component will be given a brief introduction before a more detailed description is outlined in the following sections.

Human reasoning and decision making begin with the information the person attends to in the external world. One feature of nuclear power plants is their large scale, which results in a large number of parameters needing to be monitored. Because of this, it is not possible for an operator to perceive all the plant signals at the same time. In the model proposed in this paper, the activation based approach [10] is applied to resolve this problem. Whether a signal can be perceived is determined by its activation level. This approach has been widely used in computer vision [11]. The activation of a specific signal involves two contributory factors: the salience of the signal itself and the person's attention on the signal. Only the signals with an activation level higher than a threshold will be perceived and processed for reasoning and decision making.

As shown in Fig. 1, a person's knowledge base in the model is represented by a semantic network, which consists of the main concepts about the system of interest and the relationships between these concepts. Using a semantic network to represent a person's knowledge base was also adopted in [7]. Because the knowledge base serves as the basis for human reasoning and decision making, the two components will be described together in Sect. 4.

Human reasoning and decision making rely on activation propagation in the knowledge base. The calculated activation level represents a person's attention on or belief in each concept. The propagation is directed by the relationships between the concepts in the knowledge base. Specifically, the relationships will not only help the person identify the causes of the perceived signals and then the corresponding effects, but also direct the person to check other relevant concepts. In addition, human reasoning and decision making are subject to internal factors, such as the two significant factors, stress and fatigue, which are included in the model. The reasoning and decision making processes will affect the states of stress and fatigue of an individual, which will in turn affect the two processes mentioned.

3 Information Perception

As stated in [12], data overload is a significant problem in many situations. In these situations, data overload creates a need for information intake that quickly outpaces the ability of a person's sensory and cognitive system to support that need. People can only take in and process a limited amount of information at a time. In order to determine the information that will be perceived by an operator among a large amount of data, the activation based approach is proposed in the model.

3.1 Activation Based Approach

In the activation based approach to information perception, the activation level of one item is influenced by two factors: the salience of the item and the person's attention on the item. Correspondingly, there are two processes in a person's information perception: a stimulus-driven (or bottom-up) process, and a goal-directed (or top-down) process [10, 11, 13]. This approach is supported by results in human brain research [14]. In addition, in [10], the authors applied the activation based approach to some of the basic findings in visual research and the result showed that the approach could reproduce a wide range of the findings.

In a stimulus-driven process, an item's salience is usually defined as its difference with its neighbors in terms of several specified features such as color, shape, orientation, and so forth. In a goal-directed process, a person's attention on an item is usually defined as the item's relevance to current cognitive activities, such as a person's expectation, current goal, and so on.

After the activations from the two sources are obtained, they are summed up according to their respective weights. In visual search, the summation will generate an activation map, where the hills of higher activation mark locations receiving substantial bottom-up or/and top-down activation. Then locations with activation above a threshold are searched in order of decreasing activation level.

3.2 Application of the Activation Based Approach in the Mechanistic Model

When applying the activation based approach to information perception in the mechanistic model, the problem can be decomposed into the following sub-problems: what is the specific information in the application, what are the features of the information, and how to calculate the activation level of each piece of information.

In a nuclear power plant, external information mainly refers to the monitoring signals of the plant, such as primary system pressure, water level in the reactor pressure vessel, radiation level in the secondary loop, and so forth. These signals are shown on the panel in the main control room.

The features of each signal, which can also be seen as the dimensions of salience of the signal, are identified preliminarily as: whether it is an alarm, its variation including the change rate and fluctuation, and its importance. Whether a signal is an alarm or not can be determined by comparing its value with the threshold of the alarm. A signal's variation can be calculated based on its change over a time period. Finally, the importance of a signal refers to its importance as it is being emphasized in an operator's mind through training, experience, etc. For example, primary system pressure is usually more important than the state of a backup system.

The activation level of each signal is calculated through the following equation:

$$Ac_i = \omega_{Al} \cdot Al_i + \omega_{Va} \cdot Va_i + \omega_{Im} \cdot Im_i + \omega_{At} \cdot At_i + \varepsilon \tag{1}$$

where, Ac_i is the activation level of signal i; Al_i, Va_i, Im_i and At_i represent whether signal i is an alarm, its variation, its importance, and the operator's attention on signal i, respectively; ω_{Al}, ω_{Va}, ω_{Im} and ω_{At} are the weighting factors of Al_i, Va_i, Im_i and At_i respectively; ε is the noise added to the process of information perception. From the introduction in this section, it is easy to see that the first three components on the right side in Eq. (1) are related to the bottom-up or stimulus-driven process, while the fourth component on the right side is related to the top-down or goal-driven process.

All signals are ordered according to their activation levels. To determine the signals that can be perceived and further processed in the reasoning and decision making, two criteria are proposed, which is slightly different from the application of the approach in visual search. A threshold is used as the first criterion, as is done in the field of visual search. Only those signals with activation levels above the threshold can be perceived. The second criterion is the maximum number of signals that can be processed at a time. This criterion is set based on the fact that the capacity of human's short-term or working memory is essentially limited and can hold approximately seven plus or minus two chunks of information at the same time [12]. In the end, the perceived signals, together with their activation levels, enter the knowledge base and initiate the reasoning and decision making processes.

4 Human Reasoning and Decision Making

Perceiving information from the environment triggers human reasoning and decision making by way of activation propagation in the knowledge base. This section first introduces the knowledge base and the activation propagation theory, followed by an illustration of activation propagation in the knowledge base through an example. At last, the influences of an operator's internal performance shaping factors on human reasoning and decision making are described.

4.1 Knowledge Base

In the mechanistic model, a person's knowledge base is represented by a semantic network, which consists of an operator's concepts about a system and the relationships between the concepts. This form of knowledge base can be illustrated in part through the incident of "pilot-operated relief valve (PORV) stuck open" in the Three Miles Island (TMI) accident [15]. In TMI, if the pressure in the primary system dropped below a threshold, the PORV was expected to close. Unfortunately, there was a problem with the relief valve, which led to the valve being stuck open but the indicator of the valve showed that it was closed. The consequence of the valve being stuck open was that the temperature downstream of the valve was higher than the maximum temperature allowed.

This knowledge is shown in the network in Fig. 2. The network in Fig. 2 includes three main concepts: *PORV* (i.e. a system component), *primary system pressure* (i.e. a process variable), and *PORV downstream* (i.e. a system location). Each concept is refined further. Specifically, *PORV* possesses two characteristics: *status* and *reliability*. *Status* has two mutually exclusive states: *open* and *closed*. *Reliability* also has two mutually exclusive states: *high* and *low*. *Primary system pressure* has two mutually exclusive states: *high* and *low*. *PORV downstream* possesses one characteristic, i.e. *temperature*, and it has two mutually exclusive states: *high* and *low*. The *low* state of the *primary system pressure* and the *high* state of the *reliability* of the *PORV* are connected to the *closed* state of the *PORV status* through a *2/2* (2 out of 2) logic gate. The *open* state of the *PORV status* is connected to the *high* state of the *temperature* of *PORV downstream* through an *if-then* logic gate. A sub-set of the state nodes, such as the states of *PORV status* and *primary system pressure*, can be identified through the sensors installed in the plant. This information corresponds to the signals the operator perceives from the environment.

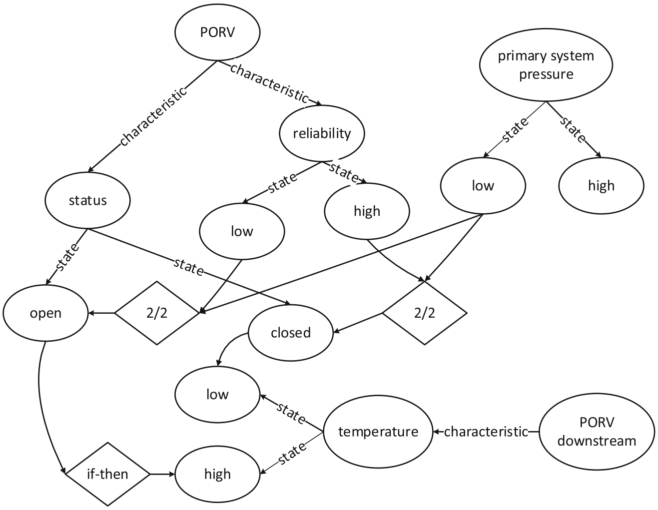


Fig. 2. Illustration of the semantic network through the TMI accident.

4.2 Activation Propagation

Spreading activation theory was proposed by Anderson [16] and applied in his ACT (adaptive control of thought) [16] and ACT-R [17] models. In the ACT-R model, the activation level of one node in the semantic network of the knowledge base is calculated according to the following equation:

$$A_i = B_i + \sum_j W_j \cdot S_{ji} \tag{2}$$

where A_i is the activation level of node i, B_i is the base-level activation of node i, W_j is the activation level of source node j, and S_{ji} is the strength of association from node j to node i.

The expression in Eq. 2 can be easily adapted to the application in this paper. In this paper, the base-level activation is the activation level that has already propagated to the node during the preceding time period, or the activation level of the perceived signals. In the first case, the existing level of activation is also subject to decay with time. Source node j of node i here refers to the node connected to node i in the semantic network. The strength of association between two nodes is in part determined by the type of their relationship and the direction of propagation of the activation.

4.3 Illustration of Activation Propagation

The propagation of the level of activation in the knowledge base is illustrated through the relief-valve-stuck-open incident, as shown in Fig. 3.

In Fig. 3, the perceived signals include *PORV status*, *primary system pressure*, and *PORV downstream temperature*. Their corresponding states are identified as *closed*, *low*, and *high* respectively. These signals are marked with thicker borders in the figure. Part of activation propagations in the network are shown by dashed lines and yellow bubbles in Fig. 3. In this case, the *closed* state of the *PORV status* and the *low* state of the *primary system pressure* activate the *high* state of the *reliability* of the *PORV* through the *2/2* logic gate. The *high* state of the *PORV downstream temperature* activates the *open* state of the *PORV status*, and then activates the *low* state of the *reliability* of the *PORV* through the *2/2* logic gate. Because both the *open* and *closed* state of the *PORV* are activated, and they are mutually exclusive, there is uncertainty with respect to the judgement of the actual state of the *PORV*. This was reflected by the operator's confusion during the accident. In fact, because of the emphasis on the *state* of the *PORV*, the activation level of perceived *status* of the *PORV* is usually higher than the activation level of the perceived *PORV downstream temperature*. As a result, the activation level of the *closed* state of the *PORV status* is higher than the one of the *open* state, which can explain in part the operator's judgment during the accident.

With respect to decision making, a decision is made based on cost-benefit evaluation of alternative actions.

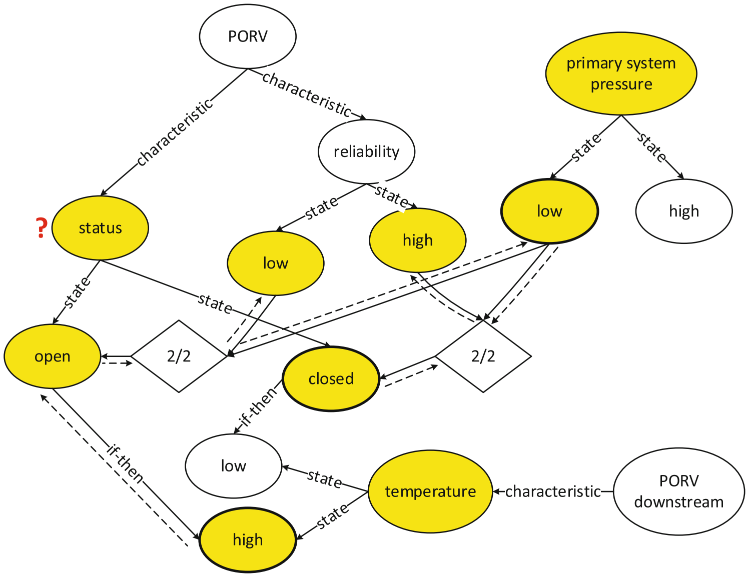


Fig. 3. Illustration of activation propagation in the knowledge base.

4.4 Integration of Internal Factors

Two significant internal factors, stress and fatigue, are integrated into the mechanistic model. These factors influence the process of human cognitive response, including information perception and reasoning and decision making, and vice versa the cognitive response influences the state of these factors.

An operator's stress level can be the result of the level of activation of particular nodes in the knowledge base, which is a consequence of his/her experience of adverse conditions. An operator's fatigue level can be expressed by the time during which the operator is subject to an event, as well as the levels of activation of perceived information (which represent the data load the operator is exposed to).

As for the impact of these factors on the human cognitive response, under high stress people tend to narrow their attention to include only a limited number of central aspects, a phenomenon which is called attentional narrowing [8]. This effect can be implicitly represented in the model by changing the weights for different factors (goal-directed and stimulus-driven) in the information perception process. Specifically, under high stress, the weight for attention in Eq. (1) is increased while the other weights are decreased. Under high stress, the decision may be made without exploring all information available, which is called premature closure [8]. This effect can be represented by the deterioration of the propagation of activation in the knowledge base. For example, when one parent node connected to a 2/2 logic gate is perceived by the operator under high stress, the activation may just propagate to the child node without propagating to the other parent node, which may lead to an error of omission.

For fatigue, research shows that a high fatigue level results in a reduction in goal-directed attention, leaving subjects performing in a more stimulus-driven fashion [18]. This effect can be represented in a way similar as to stress. Besides, with high fatigue, a person's attention or memory decays faster [7], which can be represented by changing the decay constant.

5 Discussion and Conclusion

This paper introduced a mechanistic model of human response, which will be used in human reliability analysis. The model includes three main components: information perception, knowledge base, and reasoning and decision making. The activation based approach is applied in the mechanistic model to determine the activation level of the signals, and then to determine the signals that will be perceived by the operator. An operator's knowledge base is represented by a semantic network in the model, which consists of concepts and the relationships between them. Activation propagation theory is applied to model human reasoning and decision making. Two significant internal factors are integrated into the model, stress and fatigue.

One significant advantage of the model is that it is able to simulate a person's cognitive process, and would be able to answer questions like how and why the operator behaves in one particular way. Examination of the cognitive process can be used to guide HRA related experiment design and data collection. Also based on examination of the cognitive process, human reliability can be improved from new perspectives, such as optimized human-machine interfaces to improve information perception, and more targeted training of the operator to increase the base-level activation of the less noticeable (but equally or more important) elements in the knowledge base. This model can also be compared with current methods. In THERP [1], incorrect human outputs are classified into errors of commission and errors of omission. With information perception based on signal activation and human reasoning and decision making based on the propagation of the level of activation, errors of omission can be modeled appropriately.

It needs to be noted that this paper is limited to the qualitative description of the model. To complete the model, further efforts are necessary, such as developing efficient algorithms for the propagation of the level of activation, quantification of the parameters, and comparison of the model simulation with experimental results.

Acknowledgments. This research is funded by U.S. Nuclear Regulatory Commission under Grant NRC-HQ-60-15-G-0002.

References

1. Swain, A.D., Guttmann, H.E.: Handbook of Human-Reliability Analysis with Emphasis on Nuclear Power Plant Applications, Final Report. Sandia National Labs., Albuquerque, NM, USA (1983)

2. Gertman, D., Blackman, H., Marble, J., Byers, J., Smith, C.: The SPAR-H Human Reliability Analysis Method. US Nuclear Regulatory Commission (2005)
3. Zio, E., Baraldi, P., Librizzi, M., Podofillini, L., Dang, V.N.: A fuzzy set-based approach for modeling dependence among human errors. Fuzzy Sets Syst. **160**(13), 1947–1964 (2009)
4. Chang, Y.H., Mosleh, A.: Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents: Part 1: overview of the IDAC model. Reliab. Eng. Syst. Saf. **92**(8), 997–1013 (2007)
5. Shen, S.H.: IDAC: A Cognitive Model of Nuclear Power Plant Operator Response during Abnormal Conditions. Ph.D. dissertation, University of Maryland, College Park (1994)
6. Smidts, C., Shen, S.H., Mosleh, A.: The IDA cognitive model for the analysis of nuclear power plant operator response under accident conditions. Part I: problem solving and decision making model. Reliab. Eng. Syst. Saf. **55**(1), 51–71 (1997)
7. Li, Y.: Modeling and Simulation of Operator Knowledge-Based Behavior. Ph.D. dissertation, University of Maryland, College Park (2013)
8. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors J. Hum. Factors Ergon. Soc. **37**(1), 32–64 (1995)
9. Guarino, S.L., Pfautz, J.D., Cox, Z., Roth, E.: Modeling human reasoning about meta-information. Int. J. Approximate Reasoning **50**(3), 437–449 (2009)
10. Wolfe, J.M.: Guided search 2.0 a revised model of visual search. Psychon. Bull. Rev. **1**(2), 202–238 (1994)
11. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. **20**(11), 1254–1259 (1998)
12. Endsley, M.R., Bolte, B., Jones, D.G.: Designing for Situation Awareness: An Approach to User-Centered Design. Taylor & Francis Inc., New York (2003)
13. Rybak, I.A., Gusakova, V.I., Golovan, A.V., Podladchikova, L.N., Shevtsova, N.A.: A model of attention-guided visual perception and recognition. Vis. Res. **38**(15), 2387–2400 (1998)
14. Corbetta, M., Shulman, G.L.: Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. **3**(3), 201–215 (2002)
15. Rogovin, M.: Three Mile Island: A Report to The Commissioners and to The Public (No. NUREG/CR-1250, vol. 1. Nuclear Regulatory Commission, Washington, DC (1979)
16. Anderson, J.R.: A spreading activation theory of memory. J. Verbal Learn. Verbal Behav. **22** (3), 261–295 (1983)
17. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An integrated theory of the mind. Psychol. Rev. **111**(4), 1036–1060 (2004)
18. Boksem, M.A., Meijman, T.F., Lorist, M.M.: Effects of mental fatigue on attention: an ERP study. Cogn. Brain. Res. **25**(1), 107–116 (2005)

Human Reliability and Human Factors Research

A Framework for Understanding Operator Decision Making in Simulated Nuclear Power Plant Cyber Attacks

Susan Adams(✉), Stacey Hendrickson, and Phillip Turner

Sandia National Laboratories, Albuquerque, NM, USA
{smsteve, smhendr, plturne}@sandia.gov

Abstract. Malicious cyber-attacks are becoming increasingly prominent due to the advance of technology and methods over the last decade. These attacks have the potential to bring down critical infrastructures, such as nuclear power plants (NPP's), which are so vital to the country that their incapacitation would have debilitating effects on national security, public health, or safety. Despite the devastating effects a cyber-attack could have on NPP's, there is a lack of understanding as to the effects on the plant from a discreet failure or surreptitious sabotage of components and a lack of knowledge in how the control room operators would react to such a situation. In this project, the authors are collaborating with NPP operators to discern the impact of cyber-attacks on control room operations and lay out a framework to better understand the control room operators' tasks and decision points.

Keywords: Human factors · Decision making · Nuclear power plants · Cyber attack

1 Introduction

Malicious cyber-attacks are becoming increasingly prominent due to the continual advancement of technology and adversary methods. These attacks can bring down critical infrastructures, such as nuclear power plants (NPP's). These plants are critical national assets that are so vital to the country that their incapacitation could have debilitating effects on surrounding populations as well as the national economy [1]. Despite the devastating effects a cyber-attack could potentially have on a NPP, there is a lack of understanding as to the effects within the plant from a discreet failure or surreptitious sabotage of components. How control room operators would perceive and react to such a situation is also a question. Would they be able to keep the plant safe?

In this project, the authors are collaborating with NPP operators to better understand the impact of cyber-attacks on control room operations. The authors followed the methodology used by Stevens-Adams et al. [2] to better understand the control room operators' tasks and decision points.

NPP operators are interviewed to gain a better understanding of the daily control room tasks and decisions required of them. The authors used Applied Cognitive Task Analysis (CTA) and Critical Decision Method to obtain this information.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_27

The NPP operators and subject matter experts (SME's) are also interviewed to better understand how expertise may play a role in responding to cyber-attacks. Expertise is domain specific [3] and there is reason to believe that less experienced operators may respond very differently to an attack than will experienced ones.

All of the information collected will inform scenarios (and resulting experiment) which will simulate plant conditions during a cyber-attack, component failure or insider sabotage. A scenario-based experiment with NPP operators will then be executed to determine if the operators can recognize the cyber-attack and how the operators respond to these attacks. The operators' performance will be measured and workload assessments will be collected. The results of the experiment will aid in the understanding of operator decision making during cyber-attacks and provide a platform to promote enhanced cybersecurity for the nuclear power industry.

2 Applied Cognitive Task Analysis

Applied Cognitive Task Analysis is a methodology in which interview methods are used to extract information about the cognitive demands and skills required for a task. This method is composed of three different techniques to elicit different aspects of cognitive skill: task diagram, knowledge audit and simulation [4]. The task diagram provides the researcher with a broad overview of tasks and highlights difficult cognitive portions of the task. The knowledge audit identifies ways in which expertise is used in a domain and provides examples based on actual experience. Finally, the simulation interview is based on presentation of a challenging scenario to subject matter experts, and asking the expert to identify major events, including judgments and decisions.

2.1 Task Diagram

The task diagram provides the interviewer with a broad overview of tasks and highlights difficult cognitive portions of the task. The interview consists of a series of questions, such as:

- "Think about when you complete a task. Can you break this task down into less than six, but more than three steps?"
- "Of the steps that you have just identified, which require difficult cognitive skills? Skills include judgements, assessments, problem-solving and thinking skills."

2.2 Knowledge Audit

The knowledge audit identifies ways in which expertise is used in a domain and provides examples based on actual experience. As each aspect of expertise is uncovered, it is probed for concrete examples in the context of the job, cues and strategies used, and why it presents a challenge to novices. The knowledge audit consists of a series of probes for different topics, including:

- Past & Future: Is there a time when you walked into the middle of a situation and knew exactly how things got there and where they were headed? Why would this be difficult for a novice to do?
- Big Picture: Can you give me an example of what is important about the big picture for this task? What are the major elements that you have to know and keep track of? Why would this be difficult for a novice to do?
- Noticing: Have you had experience where part of a situation just “popped” out at you, where you noticed things going on that others did not catch? What is an example? Why would this be difficult for a novice to do?
- Job Smarts: When you do this task, are there useful ways of working smart or accomplishing more with less that you have found especially useful? Why would this be difficult for a novice to do?
- Opportunities/Improvising: Can you think of an example when you have improvised in this task or noticed an opportunity to do something better? Why would this be difficult for a novice to do?
- Self-Monitoring: Can you think of a time when you realized that you would need to change the way you were performing in order to get the job done? Why would this be difficult for a novice to do?
- Equipment Difficulties: Have there been times when the equipment pointed in one direction but your own judgment told you to do something else? Or when you had to rely on experience to avoid being led astray by the equipment?

2.3 Simulation Interview

The simulation interview allows the interviewer to better understand the SME’s cognitive processes within the context of an incident. The interview is based on presentation of a challenging scenario to the SME. The SME is then asked a series of questions, such as:

- As the job you are investigating in this scenario, what actions, if any, would you take at this point in time?
- What do you think is going on here? What is your assessment of the situation at this point in time?
- What pieces of information led you to this situation assessment and these actions?
- What errors would an inexperienced person be likely to make in this situation?

3 Critical Decision Method

The Critical Decision Method [5] is an interview methodology that is implemented to better understand situation awareness and decision-making in non-routine situations. This approach is especially valuable for examining skilled performance under time pressure, which is likely a critical element in cyber-attacks.

The procedure employed for the critical decision method is as follows:

- Step 1 – Select an incident. The control room operators are asked to think about non-routine incidents in which parameter readings were not conforming to expectations.
- Step 2 – Obtain unstructured incident account. The control room operators are asked to describe the incident from the time they received the first alarm until the time that the incident was judged to be under control.
- Step 3 – Construct an incident timeline. After the incident was relayed, a sequence and duration of each event was established.
- Step 4 – Decision point identification. During the timeline construction, specific decisions were identified for further probing.
- Step 5 – Decision point probing. Follow-up questions are asked about specific decisions. Different probe types were used, including:
  - Cues (what were you seeing, hearing?)
  - Knowledge (what information did you use?)
  - Analogues (were you reminded of a previous experience?)
  - Goals (what were your goals at the time?)
  - Options (what other courses of action were considered?)
  - Experience (what specific training or experience was necessary?)
  - Time pressure (how much time pressure was involved in making the decision?)

4 Expertise

Expertise has been studied in numerous domains using a wide variety of tasks, from chess to air traffic control tasks [4, 5], physicians and clinical diagnosis tasks [6–8], music [9], and weather forecasting [10] among many others. Expertise research can be categorized by skill acquisition [11] or knowledge acquisition [12] or whether it is important to differentiate between several levels of expertise [13] or whether the differentiation between expert and non-expert is acceptable. In trying to define who or what an expert is, the theories are wide and varied. Ericsson suggests an expert must be able to select superior actions, generate rapid reactions, and control movement production [14]. Weiss and Shanteau claim there are four different categories of expertise, including expert judges, experts in prediction, expert instructors, and expert performers, yet each of these types of experts are bound by the fundamental cognitive ability of evaluation [15]. Finally, the Dreyfus model of skill acquisition [13] defines five different levels of expertise including novice, advanced beginner, competent, proficient, and expert. While dissecting the nuances of the complete expertise literature is beyond the scope of this effort, it was the authors' goal to understand the main characteristics that define an expert, agnostic of domain, and how best to define what makes an expert in the NPP control room.

The fact that expert reasoning is specific to a domain [3] is a widely-accepted statement regarding expertise and speaks to the importance of the current study. Certainly an airline pilot with thousands of hours of flight time in a particular aircraft would be considered an expert in that aircraft; however, that same individual would not be considered an expert if he suddenly found himself doing brain surgery. While that is a drastic example, other attributes that are theorized to define what an expert is are more nuanced. For example, while extensive experience of activities in a particular domain is necessary to reach exceptional levels of performance, that experience does not necessarily translate to expert levels of achievement [11]. In fact, it appears as though one's ability to reach expert levels of achievement are constrained by individual characteristics such as abilities, mental capacities, and innate talents [11] and more specifically constrained by information processing ability and working memory capacity [16] of the individual. Furthermore, some research shows "knowledge gained through extensive experience of activities in a domain" is not a differentiator between experts and novices [17, 18]. In other words, while expertise can clearly be defined as being domain-specific, simply time spent working in a domain is not the sole factor in determining expertise.

Given this understanding of what it means to be an expert, how does one go about determining who is an expert in a particular organization or domain? Experts must have extensive experience in a domain, though this is not predictive of expert levels of achievement [11]. Social acclimation [19] is the agreement of professionals regarding who is an expert in the domain. Typically, multiple individuals will not elect the same individual who is not an expert. However, it is also true that occasionally a nominated individual's performance is found to be lacking [20, 21]. Additionally, amount of overall experience in a domain needs to be distinguished from deliberate practice and the number of years spent on relevant activities is not strongly related to performance [22].

Given all of this, the authors plan to collect information pertaining to years in the field, years in the job, academic experience, training experience, and other relevant experience for control room operators. In addition, the authors discuss with each participant who s/he believed was the most expert person at her/his work location, and why.

5 Future Steps

The information obtained from the Applied Cognitive Task Analysis, Critical Decision Method, and Expertise discussions will be used to develop realistic scenarios that will be used in a future simulator experiment.

5.1 Future Simulator Experiment

The underlying NPP architecture for the emulated environment will be derived from plant drawings and plant and industry SMEs. A cyber emulation of a digital control system will be developed and coupled with a generic pressurized water reactor (GPWR) training simulator. A facility with an existing GPWR simulator platform will be used to run the experiments with licensed operators. The operators will participate in the experiment and will be asked to complete a series of scenarios on the simulator. Of particular interest is how long it takes the operator to notice that there is an issue (if the operator notices at all), what course of actions they take to solve the problem, and how cognitively taxed the operators are during the scenarios.

This research will help inform the nuclear power industry how to take appropriate actions to keep power plants safe from a cyber event. Key insights will be gained that will help inform improved operator training and procedures. The modeling capability will also provide a platform for further research and analysis of the cyber security controls within the plant and their efficacy for protecting against various types of threats. This will provide a means for advanced analysis of plant response and cyber security controls that will directly improve power plant safety and help reduce overall costs by focusing security efforts. Experiments with NPP operators will be carried out over FY2018 and results of the research are expected by the end of FY2018.

Acknowledgments. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000, SAND2017-2674 C.

References

1. Lloyd's: Business Blackout: The Insurance Implications of a Cyber Attack on the US Power Grid. Emerging Risk Report (2015). https://www.lloyds.com/news-and-insight/risk-insight/library/society-and-security/business-blackout
2. Stevens-Adams, S., Cole, K., Haass, M., Warrender, C., Jeffers, R., Burnham, L., Forsythe, C.: Situation awareness and automation in the electric grid control room. Procedia Manufact. **3**, 5277–5284 (2015)
3. Voss, J.F., Greene, T.R., Post, T.A., Penner, B.C.: Problem-solving skills in the social sciences. In: Bower, G.H. (ed.) The Psychology of Learning and Motivation: Advances in Research and Theory, vol. 17, pp. 165–213. Academic, New York (1983)
4. Militello, L.G., Hutton, R.J.B.: Applied cognitive task analysis: a practitioner's toolkit for understanding cognitive task demands. Ergonomics **41**(11), 1618–1641 (1998)
5. Klein, G.A., Calderwood, R., MacGreggor, D.: Critical decision method for eliciting knowledge. IEEE Trans. Syst. Man Cybern. **19**(3), 462–472 (1989)
6. Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., Wang, Y.: Expertise in a complex skill: diagnosing X-ray pictures. In: Chi, M.T.H., Glaser, R., Farr, M.J. (eds.), The Nature of Expertise, pp. 311–342. Erlbaum, Hillsdale (1988)
7. Woods, N.N.: Science is fundamental: The role of biomedical knowledge in clinical reasoning. Med. Educ. **41**(12), 1173–1177 (2007)
8. Feltovich, P.J., Barrows, H.S.: Issues of generality in medical problem solving. In: Schmidt, H.G., de Volder, M.L. (eds.) Tutorials in Problem-Based Learning, pp. 128–142. Van Gorcum, Maastricht (1984)
9. Glenberg, A.M., Epstein, W.: Inexpert calibration of comprehension. Mem. Cogn. **15**, 84–93 (1987)
10. Hoffman, R.R., Trafton, G., Roebber, P.: Minding the Weather. How Expert Forecasters Think. MIT Press, Cambridge (2005)
11. Ericsson, K.A.: The influence of experience and deliberate practice on the development of superior expert performance. In: The Cambridge Handbook of Expertise and Expert Performance, pp. 685–705 (2006)
12. Anderson, J.R., Corbett, A.T., Koedinger, K.R., Pelletier, R.: Cognitive tutors: lessons learned. J. Learn. Sci. **4**(2), 167–207 (1995)

13. Lester, S.: Novice to expert: the dreyfus model of skill acquisition (2005). http://devmts.org.uk/dreyfus.pdf. Adapted from Dreyfus, S.E.: Formal models vs. human situational understanding: inherent limitations on the modeling of business expertise, USAF, Office of Scientific Research, Report ORC-81-3 (1981). http://www.dtic.mil/dtic/tr/fulltext/u2/a097468.pdf
14. Ericsson, K.A.: The search for general abilities and basic capacities. In: The Psychology of Abilities, Competencies, and Expertise, pp. 93–125 (2003)
15. Weiss, D.J., Shanteau, J.: Empirical assessment of expertise. Hum. Fact. **45**(1), 104–116 (2003)
16. Jipp, M.: Expertise development with different types of automation: a function of different cognitive abilities. Hum. Fact. **58**(1), 92–106 (2016)
17. Camerer, C.F., Johnson, E.J.: The process-performance paradox in expert judgment: how can the experts know so much and predict so badly? In: Ericsson K.A., Smith, J. (eds.) Towards a General Theory of Expertise: Prospects and Limits, pp. 195–217. Cambridge University Press, Cambridge (1991). https://www0.gsb.columbia.edu/mygsb/faculty/research/pubfiles/1152/Process_and_performance_of_experts.pdf
18. Shanteau, J., Stewart, T.R.: Why study expert decision making? Some historical perspectives and comments. Organ. Behav. Hum. Decis. Process. **53**, 95–106 (1992)
19. Shanteau, J., Weiss, D.J., Thomas, R.P., Pounds, J.C.: Performance-based assessment of expertise: How to decide if someone is an expert or not. Eur. J. Oper. Res. **136**(2), 253–263 (2002)
20. Ericsson, K.A., Smith, J.: Prospects and limits of empirical study of expertise: An introduction. In: Ericsson, K.A., Smith, J. (eds.) Toward a General Theory of Expertise: Prospects and Limits, pp. 1–38. Cambridge University Press, Cambridge (1991)
21. Vicente, K.J., Wang, J.H.: An ecological theory of expertise effects in memory recall. Psychol. Rev. **1**(105), 33–57 (1998)
22. Ericsson, K.A., Krampe, R.T., Tesch-Romer, C.: The role of deliberate practice in acquisition of expert performance. Psychol. Rev. **100**, 363–406 (1993)

A Sound Investment: Managing Human Error in the Finance Sector

Andrew Wright(✉) and Clare Borras(✉)

Corporate Risk Associates, Vox Studios, 1-45 Durham Street,
London SE11 5JH, UK
`{awright,cborras}@crarisk.com`

Abstract. This paper presents the human factors analysis of the development processes, tools, guidance and organisational culture of a Tier 1 Finance institution to quantify and qualify human error incidents. Baseline analyses were performed to identify key themes behind human error in the institution's Infrastructure Support Team for server-provisioning tasks. A number of recurring themes were identified during the baseline analyses that took the analysis beyond operator error, requiring an expansion to the original scope in order to consider wider issues such as the organisational processes, policies, personnel and working culture. A number of common themes were identified across the analyses that were considered to be key contributors to incidents of human error within the organisation.

Keywords: Human factors · Human error · Global finance · Human factors integration

1 Introduction

Within the Global Finance sector there is increasing recognition of the role that human error plays in compounding incidents in terms of financial risk, process delays, organisational security and reputation and in terms of employee motivation and retention. In particular, human error that results in a cyber-security breach can lead to consequences far wider than a single organisation.

Traditionally, the Finance sector has used technology to respond to incidents of human error. Adopting Human Factors (HF), HF Integration (HFI) and a risk-based, human-centred focus to understand the *'real situation'* represents a radical approach to incident analysis for the sector. HF and HFI are firmly established in Defence, Nuclear and Aviation sectors, which provide the opportunity for sectors such as Finance, to exploit tried and tested processes, tools, guidance and expertise to develop tailored approaches that meet sector-specific needs.

This paper describes a practical approach taken to apply HF within Finance. As the assessment progressed, it became necessary to tailor approaches to fit emerging requirements and constraints with directly accessing systems and tools for analysis. Significant challenges were encountered due to target audience and Stakeholder availability.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_28

2 Project Scope

Increased organisational awareness of the impact of human error within the Server Provisioning function and, in particular, Server Design and Build (D&B), led to the Human Error Reduction (HER) Project within a global investment institution. The project scope was designed to take a *'slice through'* the institution to consider the types of human error impacting business and also make a more detailed analysis within the business area responsible for Server Provisioning, namely Infrastructure Support (IS). To this end, the project was split into two distinct phases. Phase 1 baselined human error within the institution using qualitative and a quantitative assessment of recorded incidents. Phase 2 used these data as an input to a more detailed analysis within Server Provisioning processes, tools and guidance.

3 Phase 1 – Baseline Analysis

Initially the baseline analysis aimed to quantify instances of human error recorded in two databases of approximately 800,000 incidents. The databases were designed for the Information Technology Support Desk, with incident details captured and logged by Support Desk personnel with a broad knowledge of Information Technology equipment. The databases were used to organise and prioritise technical responses to incidents to resolve equipment failure, and included information of the reporting date, equipment location and support status, together with a free text description of each incident. As the baseline analysis progressed, it became apparent that the databases included ad hoc, incomplete and inconsistent references to human error.

To quantify recorded human error incidents, the approach taken was to filter records based on the weight of information provided within specific categories. This was followed by a key word search in the free text column for terms considered to have no relevance to human error. The methodology reduced the dataset to approximately 33,000 records. At this point, a project review considered the practicalities to further analyse a data set of this size combined with concerns related to the reliability of how instances of human error had been logged.

From this, the decision was taken to re-direct the baseline study to analyse thirteen Post Problem Review (PPR) records. PPRs are initially logged in the incident databases, but are upgraded to PPRs to reflect the increased risk posed by the incident to the business. PPRs provide a more structured analysis of incidents, and directly consider human error. Analysis of the PPR records included a line by line document review, followed up by a talk-through with stakeholders to explore these case studies in more detail. Identified themes providing an initial insight into the existing gaps in organisational processes, policy, training, cultural and tools - with regard to supporting and mitigating instances of human error during server provisioning - are summarised as:

- Instances of human error were wider than operator error.
- Organisational processes and training were not designed to support human behaviour.

- Solutions were implemented to resolve human error rather than support human behaviour.
- Error mitigation processes were reactive and failed to consider the impact of modifications farther upstream, or downstream.
- Wider organisational processes between teams and business areas were not integrated, resulting in issues to communication flow and software updates.
- Guidance was not supported by sufficient 'how to' processes.
- A lack of integration and interoperability between and across the IS team impacting build processes.
- Technical workarounds that were intended as temporary measures often became permanent features rather than addressing the underlying issues.
- Errors during the input of complex data strings.

4 Phase 2 – Detailed Analysis on Server Provisioning Capability

Server Provisioning refers to the development of virtual and hardware systems that securely store, process and transfer data around the globe. Servers are developed in-house by the D&B Team to interact with systems internal and external to the organisation. Delays, errors and inefficiencies in Server Provisioning will impact the wider business processes, both upstream and downstream, and have been found to lead a range of high profile incidents. Using the findings generated in the baseline phase, the project progressed to analyse in detail the primary tool used to support D&B tasks and the processes that govern Server Provisioning.

4.1 Phase 2a – Detailed Usability Analysis on Server Provisioning Tool

The server-provisioning tool is an in-house bespoke software system that supports the development and management of active servers throughout the organisation in Europe, the Middle East and Africa (EMEA). It was developed to improve the Server Provisioning process by providing automation, management tools and a detailed history of all live builds in EMEA. The initial purpose of Phase 2a was to perform a rigorous assessment of the server-provisioning tool based on a checklist of core usability principles [1–3] and the HF UK Defence Standard [4]. Any subsequent usability shortfalls would then be mapped onto a usability compliance matrix in order to pinpoint specific issues. Due to limitations in access to the software (as a result of security restrictions), the scope was subsequently modified away from a rigorous assessment of specific aspects of the tool, and toward a broader assessment based on established HF design principles.

Methodology. The methodology used included a stakeholder led talk-through and walk-through of a D&B scenario, an observation exercise undertaken in the real working environment with the target audience performing real builds and semi-structured interviews with a questionnaire undertaken by HF Analysts with the target audience. Interviews and follow up questions using email facilitated the collection of subjective feedback about the tool in terms of its functional scope, interoperability issues and usability. The data collection methodology was supported by the predefined checklist of established HF design principles to enable analysts to capture the scope of usability issues impacting the human-computer-interaction. This checklist is presented in Table 1.

Table 1. Checklist of established usability principles.

#	Usability principle
1	Consistency in terminology, layout and use of commands
2	Feedback on actions performed proportionate to task importance
3	Means to reduce/remove possibility of error
4	Visibility of system status, controls and information.
5	Reversibility of actions
6	Provision of dialog to yield closure of processes.
7	Constraints on user actions
8	Integrated shortcuts for advanced users
9	Minimalist design
10	Mapping of controls to the real world
11	Affordance of controls
12	Low short-term memory load
13	Match between system and real world

Stakeholder Workshop. The stakeholder workshop provided an opportunity to develop analyst familiarisation of the domain and the Server Provisioning processes used within the organisation. HF Analysts were joined by a team leader from D&B to talk through the software, processes and to provide a flavour of some of the potential HF issues within the processes as perceived by the D&B Team. The workshop used open-ended questions to prompt exploration of emerging issues and directed subsequent technical activities.

Observation Exercise. The D&B team comprised three persons, all of whom participated in the Observation exercise. The exercise took place on-site over a period of three hours with personnel performing real tasks in the normal working environment. Participants were primed to raise issues and processes as they occurred during the session to elicit feedback and provide explanations as necessary.

User Questionnaires. Following the Observation Exercise, an analyst-led semi-structured interview was conducted with participants. Participant anonymity was assured and an opportunity to provide further information and provide additional information after the session was provided. The session was designed to elicit participant feedback concerning:

- The role the tool played in D&B Server Provisioning tasks.
- How well the tool supported the real D&B task flow.
- The degree to which the tool supported D&B Server Provisioning tasks.
- Known issues that could impact error detection and mitigation.
- Those functions considered most and least useful to support D&B tasks.

4.2 Phase 2b – Detailed Analysis on Server Provisioning Processes

The aim of Phase 2b was to review the processes that governed Server Provisioning and D&B to support human error detection and the consideration of human factors issues. The aim was to build on issues identified from the preceding analyses that recurred and also to list any additional issues that emerged. In consultation with the stakeholder, the decision was taken to conduct the process analysis on a review of three candidate processes. The process considered to reflect current best practice within D&B was underpinned by a generic process referred to as DMAIC (Design, Measure, Analyse, Improve, and Control). Using the Design, Analyse, Improve and Control (DAIC) development stages as a baseline, a draft tailored approach was developed; the Infrastructure Technology Timeline Process (IT Timeline Process). The IT Timeline Process was the focus of the Phase 2b assessment.

Methodology. Taking into consideration the highly pressured and complex working environment, the methodology used to analyse the IT Timeline Process was iterative, agile and minimised intrusion where possible. The methodology included the following:

- Comparability and Gap Analysis: Using a process mapping exercise to identify the gaps and potential linkages between the IT Timeline Process, Systems Engineering and the UK Defence HFI Process.
- Task analysis of nine incidents attributed to human error.
- Case study analysis of two non-standard server builds.
- Stakeholder Workshop to pinpoint and plot issues identified in preceding analyses on the TI Process Timeline.

Approach. Three technical activities were conducted to identify and explore issues. These activities included a Process Review, a Case Study Analysis and a Workshop to plot identified issues along the D&B development process.

Process Review and Mapping Exercise. An analysis of the technical activities embedded within System Engineering, HFI and the TI Timeline Process enabled the mapping of process stages that broadly aligned. Figure 1 demonstrates the first cut analysis to broadly demonstrate process stage alignment.

Case Studies Analysis. Two project based case studies were selected for detailed analysis. Both case studies had been rated as high priority incidents because of the extent of the delays and costs incurred due to human error and human factors issues.

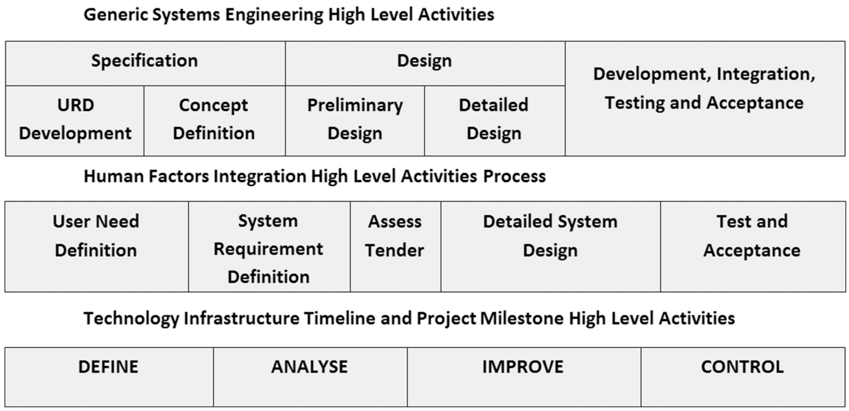


Fig. 1. High level development process mapping

Case Study 1 was concerned with an update to the existing '*stencil*' (template) portfolios used to define the variables and parameters of an automated server build. System updates elsewhere in the organisation led to an inconsistent implementation of new variables in the stencil portfolio. Delays were further compounded by regional variations in resources and software.

Case Study 2 was a project based on a specialist business requirement that fell outside of the institution's standard global pricing processes. The requirement drove a need to integrate specialist systems from new suppliers and the lack of suitable in-house subject matter expertise led to additional complexity and required specialist expertise to support the project that was no longer available within the institution.

The case studies indicated that difficulties were encountered with standard builds due to issues with incomplete requirements and assurance processes that only tested the validity of variables and parameters without considering the operational context. Furthermore, existing processes lacked the agility to support non-standard builds and, as projects moved beyond standard processes, the likelihood of meeting project milestones and budgets decreased and the scope for human error increased.

The identified issues were distilled into a checklist. This checklist was then used in a Stakeholder workshop to plot these, and any emerging issues, based on a consensus of opinion along the TI Process timeline.

Emerging Issues. Table 2 represents the issues identified and distilled during preceding analyses. It is important to note that issues identified in Phase 1, reappeared in Phase 2.

Plotting Issues and Gaps along the TI Process Timeline. The Stakeholder Workshop was driven by two analysts with expertise in Human Factors and Operational risk. Six participants from IS at the Chief Technology level were involved in the process to plot the identified checklist of distilled issues and gaps in Table 2 along the TI Process Timeline.

Table 2. Emerging issues from Phase 2b analysis

Issue number	Description
1	Environment of rapid deployment
2	SME availability and workload
3	Forecasting and '*Just in Time*' provisioning
4	Annual change freeze and funding rollover restrictions.
5	Quarter on quarter server purchases lacked consistency
6	Financial approval processes were complex
7	Vendor selection and negotiation for non standard projects
8	Process for non standard projects
9	Resourcing of expertise for non standard projects
10	'*Can Do*' culture can impact review processes
11	Capture and reporting of operator issues
12	Cultural differences
13	Transatlantic scheduling
14	Inconsistent Project Management and ownership
15	Timeliness, accuracy and relevancy of information flow
16	'Monitoring' solutions to track build progress
17	Requirements for equipment that does not comply with standards
18	Requirements: Specification, Overestimate, Management, Tracking and Creep
19	Quality Assurance does not assess the operational context
20	Operators can override automated checks
21	Gaps between Quality Control 1 and Quality Control 2
22	Error rates increase with manual data entry
23	Fault diagnosis is manual and operator-driven
24	Stencil approval, adding functionality and upgrades
25	Cloning and cleansing of stencils and templates
26	Domain Naming System updates can delay process for forward/reverse look up
27	No process to support early engagement of D&B
28	Process guidance lacks 'how to' information
29	Training gaps
30	No means of preventing incorrect resource allocation

The aim was to further develop the high level development process mapping undertaken previously and place issues and gaps where they occurred along the timeline. Figure 2 demonstrates cross cutting and phase specific issues and gaps revealed from preceding analyses. This activity provided the basis for further work to consider aspects of HFI that could be integrated to improve D&B processes for further work.

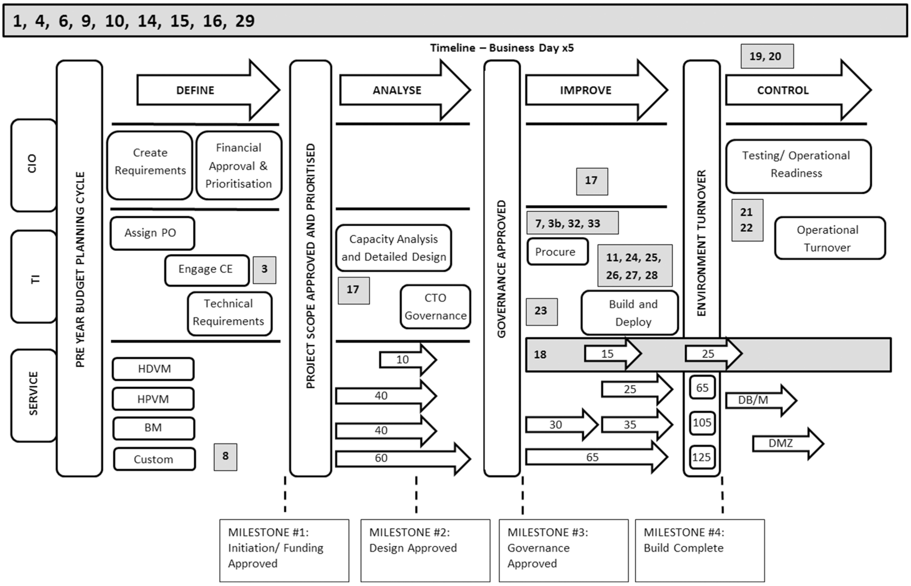


Fig. 2. Identified issues and gaps located on the TI process

5 Findings

Triangulation of the Phase 1 and 2 analyses revealed a number of common human factors issues and themes that significantly contributed to instances of human error, and identified issues and gaps in development process, policy and working culture.

5.1 Process

The scope of issues identified with the TI Timeline process included the extent to which D&B activities were not optimally supported. Shortfalls were identified in terms of the failure to trigger early D&B engagement and non-optimal integration of external processes to Server Provisioning. D&B had no formal system to monitor development progress and formal project management approaches were not followed; as well as there being no formal means to consider the human component during development.

Processes also failed to consider the impact of changes further up and down stream in terms of development, personnel, assurance and change management. Processes were not found to be sufficiently flexible to support non-standard or novel projects, in terms of both resourcing and requirements. A lack of clear ownership, active management and expertise for non standard projects was identified, which contributed to a number of high profile incidents within the organisation. Evidence was found where informal workaround tools and spreadsheets were constructed to plug gaps in existing tools, processes and systems.

5.2 Error Detection

D&B tasks required the aggregation and processing of a large amount of information. This information was based on requirements that did not adhere to formal requirements engineering (e.g. specific, measurable, acceptable, realistic and timely (SMART)) and were incomplete and inconsistent in quality and detail. D&B personnel pulled information from a number of systems resulting in a complicated user interface with tasks moving between various applications resulting in interruptions to task flow. The build process required complex manual data input tasks that did not comply with established standards. This issue was further exacerbated by the requirement to gather this information from a number of sources for every server provisioning build.

Fault diagnosis was manual and operator driven. Quality control processes assessed the validity of build parameters and variables but not build relevance to the operational context or against formal requirements. Of significant concern was the practice of build modifications being made after the second quality control process had completed, or builds going live without the second process having been initiated.

Given that a server build can vary greatly depending on client needs, this gap in quality control was considered to represent a complete gap in mitigation against human error, where actions with severe operational consequences can be performed without any error detection/task validation from the server provisioning tool.

5.3 Task Flow in Design

Extraneous functionality was added to the Server Provisioning tool without consultation with the target audience, resulting in confusion within the team. It was determined later in the project that the additional functionality was introduced as part of a future software upgrade. However, this was not communicated, and it was unknown whether this functionality could negatively impact existing D&B tasks and provisioned servers.

5.4 Societal/Cultural Attitudes

The working culture was very much results-oriented with a strong ‘roll your sleeves up’ attitude to working. The working environment was very fast paced with a high workload and the target audience were motivated by quality and efficiency. Added complexity came from different methods of working within different teams. At the time of the assessment, the server provisioning team were transitioning to a more formal project management and requirement management approach to development, whilst other teams remained focused on quick fix solutions that did not adequately consider the impact of changes elsewhere in the organisation or project.

During the project a number of awareness raising sessions were conducted, relating to project progress and information about human error and human factors. During these sessions it was apparent that the target audience could relate to the issues being presented and supported the need for process improvement.

6 Conclusions

This project took a ‘slice through’ a Finance institution and identified the causes of human error within the server provisioning process. The project generated a large number of findings concerning the institution’s processes, policies and tools to support development projects and improve consideration of human issues. The assessments pinpointed where issues occurred during the server provisioning process and developed a number of ‘quick win’ and long-term recommendations in order to improve resilience to human error. The key findings generated a number of overarching recommendations that can be summarised as:

- Improve server provisioning processes to support the full range of server provisioning projects which consider the human component and support project management.
- Integrate, enforce and formalise quality assurance to include operational assessment and requirements management processes, supported by a suitable champion.
- Raise organisational awareness of the human risk and the impact of human error.
- Develop tools that support the allocation of function and error detection at the point of entry, interoperability based on HF task and risk analysis and compliance to established HF standards.

This project has helped initiate the consideration of HF within the organisation and the wider finance sector, and requires validation and further work based on the recommendations generated. Further projects with the institution are under consideration to consider how HFI can be implemented to support Server Provisioning and Project Management.

References

1. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of ACM CHI 1990 Conference, Seattle, WA, 1–5 April, pp. 249–256 (1990)
2. Norman, D.: Design principles for human-computer interfaces. In: Proceedings of SIG CHI 1983 Conference, Boston, MA, 12–15 December, pp. 1–10 (1983)
3. Schneiderman, B.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. 5th edn. Addison-Wesley (2009)
4. Ministry of Defence. Defence Standard 00-250 Human Factors Integration for Defence Systems. Issue 1 (2008)

Adapting the Cognitive Framework of NUREG-2114 for Use in Human Error Analysis for AP1000 Plant Licensing in the United Kingdom

April M. Whaley(✉)

Westinghouse Electric Company LLC, 1000 Westinghouse Drive,
Cranberry Township, PA 16066, USA
whaleyam@westinghouse.com

Abstract. Westinghouse Electric Company is in the process of completing the Generic Design Assessment (GDA) phase of licensing its AP1000® (®AP1000 is a trademark or registered trademark of Westinghouse Electric Company LLC, its affiliates and/or its subsidiaries in the United States of America and may be registered in other countries throughout the world. All rights reserved. Unauthorized use is strictly prohibited. Other names may be trademarks of their respective owners.) Pressurized Water Reactor (PWR) nuclear power plant in the United Kingdom. To address a Human Factors (HF) GDA issue identified in an earlier phase of the GDA process, the Westinghouse HF team updated the human error analysis methodology to comport with lessons learned and advancements in the field of human reliability analysis (HRA). NUREG-2114 [1] provided a cognitive framework as an update to the psychological basis for HRA, but did not provide specific guidance to analysts on how to use this framework in performing HRA. This paper describes how the Westinghouse HF team adapted the cognitive framework in NUREG-2114 for application in the human error analyses performed to resolve the HF GDA issue. Westinghouse HF determined that the adapted cognitive framework was useful for identifying potential human errors in the task analysis and identifying potential design improvements. Results from using NUREG-2114 to inform human error analysis and recommendations for additional development are discussed.

Keywords: Human factors · Human error analysis · Human reliability analysis · NUREG-2114 · Performance shaping factors · Performance influencing factors · Error mechanisms · Proximate causes · HRA · PSFs · PIFs

1 Introduction

The United Kingdom (UK) Office for Nuclear Regulation (ONR) uses the Generic Design Assessment (GDA) process to assess new nuclear power plant designs before granting a licence for site construction to begin. The UK regulatory framework requires duty holders to reduce risks ‘as low as reasonably practicable’ (ALARP) [2]. To ensure this legal requirement is met, the ONR has established Safety Assessment Principles (SAPs) [3] and Technical Assessment Guides (TAGs) to guide inspectors in assessing

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_29

if the plant design conforms with international and national modern safety standards and industry guidance that the ONR considers as relevant good practice (RGP). The ONR uses the SAPs, TAGs and RGP in assessing plant designs across the multi-step GDA process to ensure that the plant and equipment design is robust and provides adequate protection against potential accidents to a degree that meets modern international good practices.

Westinghouse Electric Company completed Step 4 of the GDA process for the AP1000® design in 2010, and achieved an interim design acceptance. The ONR also identified 51 GDA issues for Westinghouse to resolve in order to successfully complete the GDA process and achieve design acceptance and authorization to begin site licencing. One of the GDA issues specifically addressed HF. In 2011, Westinghouse suspended the GDA process while customer finalization occurred.

Upon resumption of the GDA process in late 2014, Westinghouse began work to resolve the Human Factors GDA issue, which concerned the completeness of the human factors safety case, particularly regarding the identification of human error mechanisms in the human reliability analysis (HRA) and the potential for operator misdiagnosis and violation.

At this time, Westinghouse began updating its human error analysis (HEA) methodology to comport with the latest developments in the field of HRA[1], including lessons learned from the International and U.S. HRA empirical studies [4–7] and Good Practices for Implementing HRA [8] regarding the importance of thorough, well-performed qualitative error analysis. Westinghouse also chose NUREG-2114, Cognitive Basis for Human Reliability Analysis [1] to guide the approach for identifying potential error mechanisms and performance shaping factors (PSFs).

This paper describes the process the Westinghouse HF team used to adapt the cognitive framework in NUREG-2114 into a tool that analysts could use in conducting the HEAs performed to address the Human Factors GDA issue, the lessons learned in applying this tool, and recommendations for additional guidance.

2 The Cognitive Framework of NUREG-2114

The authors of NUREG-2114 [1] conducted a wide-ranging review and synthesis of the scientific, cognitive, and psychological research literature to develop an updated technical basis and cognitive model for HRA. Prior to the publication of NUREG-2114, most HRA methods relied on cognitive models dating from the 1980s or earlier. NUREG-2114 provides a model of human macrocognition, or cognition in real-world settings, which emphasizes what people do with their brains rather than the fundamentals of how the brain works. NUREG-2114 presents a detailed discussion of five overlapping and interacting macrocognitive functions:

---

[1] The Westinghouse process for conducting HRA in the UK involves the Human Factors organization performing the qualitative analysis, and the Probabilistic Risk Analysis (PRA) discipline performing the HRA quantification and input into the overall Probabilistic Safety Assessment (PSA) model. For this reason, Westinghouse uses the term HEA to refer to the qualitative portion of the human reliability analysis.

- Detecting and Noticing
- Understanding and Sensemaking
- Decision Making
- Action
- Teamwork

NUREG-2114 identifies for each macrocognitive function the underlying cognitive mechanisms—the psychological processes that enable the function to operate (or, alternately, fail). NUREG-2114 also identifies the relevant PSFs that can influence those cognitive mechanisms toward success or failure, as based on the research literature. Most importantly, NUREG-2114 provides a cognitive framework: a model that links the proximate causes of failure of each macrocognitive function to the underlying cognitive mechanisms and the relevant PSFs that contribute to success or failure. By showing how PSFs influence cognitive errors, this causal tree illustrates how and why macrocognition may fail. Figure 1 shows an example of the cognitive framework structure; note that NUREG-2114 uses the term "performance influencing factor" (PIF) as a synonym to PSF.

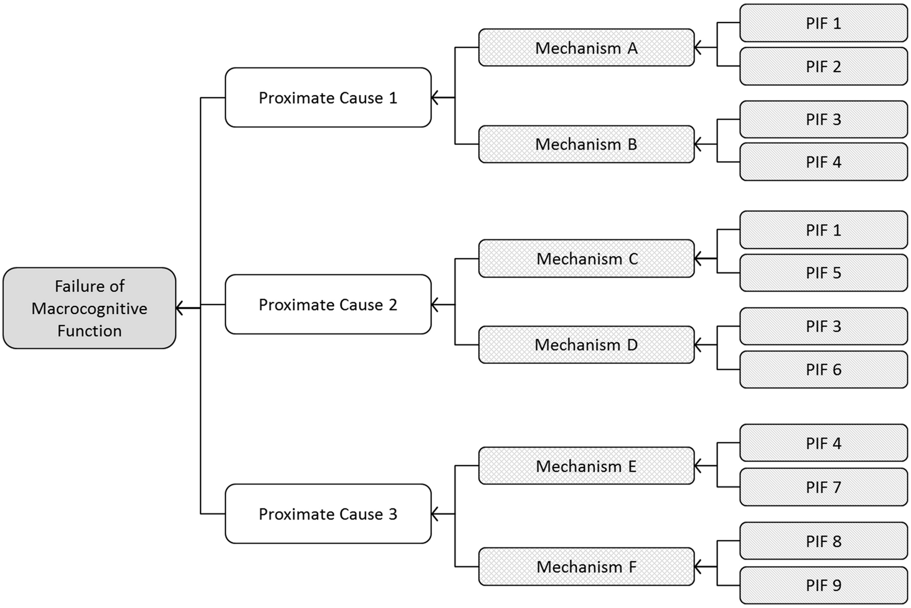


Fig. 1. Example cognitive framework structure from NUREG-2114.

3 Adapting the Cognitive Framework for Human Error Analysis

The authors of NUREG-2114 designed the cognitive framework to be a tool to aid analysts in identifying which PSFs, cognitive mechanisms, and proximate causes to consider investigating for a particular analysis. However, the Westinghouse HF team found that it was not possible to utilize the cognitive framework as is; while detailed AP1000 design information was available, site-specific information was not yet available. Absent a constructed plant, site-specific organizational processes and plant procedures; it was necessary to adapt the cognitive framework to this particular application. The Westinghouse HF team also found that the proximate causes and PSFs identified in the NUREG-2114 cognitive framework are typically measurable, observable, or otherwise identifiable, whereas cognitive mechanisms are not. Knowing the cognitive mechanisms underlying a failure is useful to identify potential improvements in the design or procedures, but it is not feasible to predict failure at the cognitive mechanism level.

The Westinghouse HF team also noted that the proximate causes more closely resemble the kinds of error mechanisms typically assessed in HRA. A more utilitarian view of the framework would consider the *error mechanisms* as arising from a PSF interacting with the psychological processes to produce an error (proximate cause of failure) for the macrocognitive function. In conducting the error analyses performed to address the HF GDA issue, the analysts found the macrocognitive functions, proximate causes and PSFs to be of most use in identifying potential errors, and the cognitive mechanisms as useful for identifying potential design improvements.

The proximate causes NUREG-2114 identifies for the macrocognitive functions are shown in Table 1. See NUREG-2114 for detailed discussion of the cognitive mechanisms. Even at this level, the Westinghouse HF team considered the proximate causes to be more detailed than it is possible to predict for HEAs conducted during GDA.

Table 1. Proximate causes of failure of the macrocognitive functions in NUREG-2114.

Macrocognitive function	Proximate cause of failure
Detecting and Noticing	• Cues and information not perceived • Cues and information not attended to • Cues and information misperceived
Understanding and Sensemaking	• Incorrect data used to understand the situation • Incorrect integration of data, frames, or data with a frame • Incorrect frame used to understand the situation
Decision Making	• Incorrect goals or priorities set • Incorrect internal pattern matching • Incorrect mental simulation or evaluation of options
Action	• Failure to execute desired action (error of omission) • Execute desired action incorrectly
Teamwork	• Failure of team communication • Error in leadership/supervision

It is worth noting that the authors of NUREG-2114 acknowledge that the factors identified in the cognitive framework are not necessarily the *only* potentially relevant factors; other factors could play a role. Furthermore, the authors of NUREG-2114 note that the teamwork framework is limited to communication and leadership factors, but it does not include organizational factors that influence team performance, such as openness and democracy of the team, procedure compliance policy, and communications protocols.

For these reasons, the Westinghouse HF team simplified the NUREG-2114 framework for use in the HEA methodology, and the HF analysts were given license to identify other error modes or factors outside of the NUREG-2114 framework, as applicable to their respective assessments.

Westinghouse first identified the PSFs most often identified for each macrocognitive function, as shown in Table 2.

Table 2. Number of times NUREG-2114 identifies a PSF as relevant for each macrocognitive function.

PSF	Detecting and Noticing	Understanding and Sensemaking	Decision-making	Action	Teamwork
Human System Interface (HSI) - Output	12	10		10	
Knowledge/Experience/Expertise	7	16	14	7	5
Stress	7				
Task Complexity	7				
Procedure availability and quality		15	7		
Training		12	14		
Time load/pressure			9		6
Non-task load (Addressed by divided attention)				5	
Task Load (Addressed by task complexity)				7	
Leadership style (Addressed by Conduct of Operations assumptions)					4

Westinghouse HF also reviewed PSFs identified by the UK Health and Safety Executive (HSE) as important for tasks [9]), as shown in Table 3.

Based on the review of the PSFs from NUREG-2114 and the above factors identified by HSE as important for tasks, the 2015 Westinghouse HEA methodology identified the following list of PSFs for use in assessments:

- Human System Interface (HSI) design
- Knowledge/experience/expertise (training)
- Stress
- Task complexity
- Attention

Table 3. Factors identified by HSE as important for tasks [9].

Job - Attention to task
Job - Attention to surroundings
Job - Clarity of signs, signals, instructions and other information
Job - System/equipment interface (labelling, alarms, error avoidance/tolerance)
Job - Difficulty/complexity of task
Job - Routine or unusual
Job - Divided attention
Job - Procedures inadequate or inappropriate
Job - Preparation for task (e.g. permits, risk assessments, checking)
Job - Time available/required
Job - Tools appropriate for task
Job - Communication, with colleagues, supervision, contractor, other
Job - Working environment (noise, heat, space, lighting, ventilation)

- Procedures
- Time pressure
- Communication
- Task preparation (for maintenance tasks only)
- Availability of required tools (maintenance tasks only).

The Westinghouse HF team then consolidated the macrocognitive functions into diagnosis and execution/action, with Detecting and Noticing and Understanding and Sensemaking aligning to diagnosis, and Decision Making and Action Implementation aligning to action execution. Figure 2 shows the simplified tool for identifying which PSFs to focus on for diagnosis and action, based on the NUREG-2114 framework.

Westinghouse considers teamwork to be a part of every action; the entire crew is responsible for failure or success of an action. Additionally, the NUREG-2114 framework for teamwork identifies many of the same factors as are in the individual-level macrocognitive functions. Due to this overlap and the acknowledged weakness in the NUREG-2114 teamwork framework, as discussed above, the Westinghouse HF team decided to address teamwork via generic assumptions about currently unavailable site-specific information that relate directly to the teamwork error mechanisms and PSFs identified in NUREG-2114. Specifically, Westinghouse assumes that the site licensee will:

- Implement a best-practice nuclear safety culture comporting with industry standards,
- Implement Westinghouse's AP1000 standard plant conduct of operations,
- Require verbatim adherence to procedures,
- Require the use of human performance tools,
- Institute clear roles and responsibilities and a hierarchical leadership style,
- Employ a systematic approach to training, and
- Comply with UK regulations.

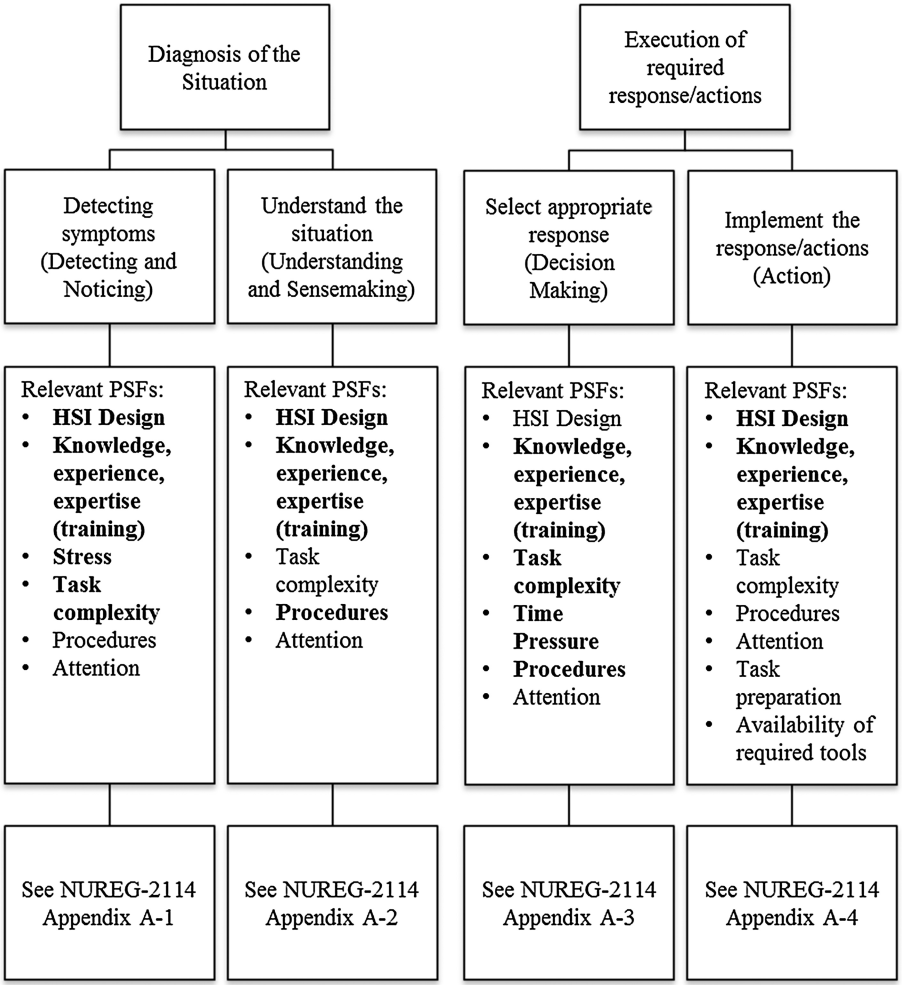


Fig. 2. Identification of most relevant PSFs for each macrocognitive function. Bolded text indicates PSFs identified by NUREG-2114 for a macrocognitive function; plain text indicates PSFs that Westinghouse also deemed relevant based on experience and analyst judgment.

Keeping with HRA best practice [7], the HEA documentation explicitly describes these assumptions so they can be validated with the site licensee upon commencement of the site licensing and plant construction process.

4 Results and Discussion

The Westinghouse HF analysts used the adapted NUREG-2114 framework as described above to inform the HEAs performed to address the Human Factors GDA issue. Specifically, analysts used the adapted framework while conducting the task analyses to identify relevant PSFs and potential errors and error mechanisms across the subtasks of the action. The operator actions analyzed to address the Human Factors GDA issue included post-fault operator actions modeled in the PSA, operator actions that have the potential to cause initiating events if performed incorrectly, maintenance actions, and actions taken in the case of a prolonged station blackout.

The adaptation of the NUREG-2114 cognitive framework into the HEA process enabled Westinghouse to address the Human factors GDA issue and demonstrate that Westinghouse's approach comports with relevant good HRA practice, in compliance with ONR expectations for best practice methodology.

In implementing this adapted cognitive framework into the HEA process, the Westinghouse HF team made several observations about the utility of NUREG-2114. Specifically:

- NUREG-2114 provides a thorough discussion of human psychology and the origin of cognitive errors, and is useful for training analysts who do not have a background or formal training in psychology.
- The PSFs identified in the framework are useful for identifying which PSFs should be evaluated in an analysis.
- The proximate causes are useful for identifying potential errors that can occur at particular subtasks.
- The cognitive mechanisms are most useful for informing recommendations for design or procedure changes to reduce the opportunity for error.
- The cognitive framework of NUREG-2114 is informative theoretically, but needs adaptation for specific applications. In this case, the Westinghouse HF team observed that:
  - The PSFs included in the cognitive framework may not cover all possible situations, for example, unique maintenance considerations such as rigging, lifting, or accessibility, or considerations for beyond-design basis situations such as a prolonged station blackout or severe flooding. In such cases the PSFs to be considered need to be expanded to include additional specific factors relevant to the situation under assessment.
  - The teamwork structure needs additional development to include aspects such as crew cohesion, distribution of workload, conduct of operations, roles and responsibilities, and communication protocols. Westinghouse addressed GDA by making assumptions about site-specific organizational and teamwork factors. A more developed teamwork structure will be useful when the HRA is conducted during site licensing.
  - The proximate causes, while useful to identify types of errors, do not always correspond well to error taxonomies commonly used in HRA methods. The Westinghouse HF team often identified other types of errors in its analyses beyond those listed in the proximate causes. The HF team recognizes that the

proximate causes lie in a causal structure, and that many error taxonomies are at a descriptive rather than causal level; this means that the proximate causes could underlie many error types. Therefore, the Westinghouse HF team recommends that it would be useful to perform a mapping exercise to identify how the proximate causes relate to and potentially cause the kinds of errors in common HRA error taxonomies.

5 Conclusions

In summary, the Westinghouse HF team found the NUREG-2114 cognitive framework to be valuable in developing a best-practice HEA methodology, and while application-specific modifications are necessary to the cognitive structure to enable utility, Westinghouse considers the adapted framework structure as useful not only for performing HEA but also to assist in the identification of potential design improvements. The adapted NUREG-2114 cognitive framework was a significant contributor to Westinghouse's ability to address successfully the Human Factors GDA issue. Additional development of the teamwork structure, enhancements for a wider range of scenarios, and additional discussion of how the proximate causes relate to commonly used error taxonomies would be beneficial in increasing the utility of the cognitive framework in HRA.

References

1. Whaley, A.M., Xing, J., Boring, R.L., Hendrickson, S.M.L., Joe, J.C., LeBlanc, K.L., Morrow, S.L.: Cognitive Basis for Human Reliability Analysis (NUREG-2114). U.S. Nuclear Regulatory Commission, Washington DC (2016)
2. Office for Nuclear Regulation: Guideline on the Demonstration of ALARP (As Low As Reasonably Practicable). ONR Technical Assessment Guide NS-TAST-GD-005, Rev. 7 (2015)
3. Office for Nuclear Regulation: Safety Assessment Principles (SAPs). http://www.onr.org.uk/saps/
4. Forester, J., Dang, V.N., Bye, A., Lois, E., Massaiu, S., Borberg, H., Braarud, P.Ø., Boring, R., Mannisto, I., Liau, H., Julius, J., Parry, G., Nelson, P: The International HRA Empirical Study: Lessons Learned from Comparing HRA Methods Predictions to HAMMLAB Simulator Data (NUREG-2127). U.S. Nuclear Regulatory Commission, Washington DC (2014)
5. Lois, E., Dang, V.N., Forester, J., Borberg, H., Massaiu, S., Hildebrandt, M., Braarud, P.Ø., Parry, G., Julius, J., Boring, R., Mannisto, I., Bye, A.: International HRA Empirical Study – Phase 1 Report: Description of Overall Approach and Pilot Phase Results from Comparing HRA Methods to Simulator Performance Data (NUREG/IA-0216, vol. 1). U.S. Nuclear Regulatory Commission, Washington DC (2009)

6. Bye, A., Lois, E., Dang, V.N., Parry, G., Forester, J., Massaiu, S., Boring, R., Braarud, P.Ø., Broberg, H., Julius, J., Mannisto, I., Nelson, P.: International HRA Empirical Study – Phase 2 Report: Results from Comparing HRA Method Predictions to Simulator Data from SGTR Scenarios (NUREG/IA-0216, vol. 2). U.S. Nuclear Regulatory Commission, Washington DC (2011)
7. Forester, J., Liau, H., Dang, V.N., Bye, A., Lois, E., Presley, M., Marble, J., Nowell, R., Broberg, H., Hildenbrandt, M., Hallbert, B., Morgan, T.: The U.S. HRA Empirical Study – Assessment of HRA Method Predictions against Operating Crew Performance on a U.S. Nuclear Power Plant Simulator (NUREG-2156). U.S. Nuclear Regulatory Commission, Washington DC (2016)
8. Kolaczkowski, A., Forester, J., Lois, E., Cooper, S.: Good Practices for Implementing Human Reliability Analysis (NUREG-1792). U.S. Nuclear Regulatory Commission, Washington DC (2005)
9. Health and Safety Executive: Development of a Working Model of How Human Factors, Safety Management Systems, and Wider Organisational Issues Fit Together (RR543). (2007)

Effects of Information Acquisition Method on Diagnostic Task Performance Using Digitalized Interfaces

Xuansheng Ding[1], Zhizhong Li[2(✉)], and Manrong She[2]

[1] Economics and Development Research Institute,
SINOPEC, 100029 Beijing, People's Republic of China
[2] Department of Industrial Engineering, Tsinghua University,
100084 Beijing, People's Republic of China
zzli@tsinghua.edu.cn

Abstract. This study investigated the effects of information acquisition method on performance and behavior of diagnostic tasks in a simulated nuclear power plant (NPP) system. Four different displays were designed. Each type of display had one specific information acquisition method (namely, query, filter, and query + filter, as well as conventional visual search). Experimental results revealed that visual search display gave the shortest completion time, lowest subjective workload, and shortest time to get the first correct hypothesis. However, the Query display, Filter display and the Query + Filter display did not show anticipated strength in aiding diagnosis. Meanwhile, high time pressure led to shorter completion time, shorter time to get the first diagnosis, and higher subjective mental workload. The results enrich the evaluation for information acquisition methods and may provide cues of interface design for complex industrial systems.

Keywords: Information acquisition method · Interface design · Time pressure · Diagnosis performance · Diagnosis process

1 Introduction

1.1 Background

The digitalization of the Human System Interfaces (HSIs) has largely changed the roles of operators in control rooms. Many benefits could be realized, such as large storage space, strong computing power, and various information representation methods. However, new problems could also show up, such as massive information leading to overload, and deep system causing loss of situation awareness. These problems would be harmful to task performance, especially for knowledge-based tasks such as diagnosis.

Diagnosis is one difficult type of task for main control room (MCR) operators. To diagnose the failure of a component or system, operators have to observe much information, and rely on their operating experience and system knowledge. Industrial systems like nuclear power plants (NPPs) have complicated structures and numerous parameters, thus the difficulty of diagnosis is quite high. Various diagnosis aids have

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_30

been designed, such as diagnostic algorithms and expert systems. However, the core role of operators cannot be replaced by these aids. Thus, display design for operators to better interact with the system is always essential.

1.2 Literature Review

Diagnosis is a task process that operators collect parameter values, perceive operating status of a system, and then find out the reason of the malfunction. The diagnosis task is a knowledge-based task, according to the skill-/rule-/knowledge-based task classification [1]. The whole diagnosis process can be divided into three stages, i.e. information acquisition, hypothesis formation and hypothesis confirmation [2].

In the studies of NPP operator behaviors, researchers have also established the process of fault diagnosis. The process is almost the same—obtaining system operating signals, recognizing symptoms of the system, and getting the reason of fault [3]. Researchers have also found that, in complicated systems like NPPs, operators tend to diagnose based on their experience [4]. Various diagnosis aids have already been developed. There are roughly three categories of such aids: display aids, algorithm aids and knowledge aids.

Display Aids. The primary design goal of diagnosis display aid is to show the relationships of the parameters on side as many as possible. In this way, the thinking process could be simplified, the chance of error could be reduced, and the performance could be improved. Such aids can be one single graph or the whole display design. One single graph represents the relationships of several parameters (see [5, 6]), or marks the range of parameters. The whole display aid features in that the whole display area is designed for thoroughly showing the status of the system and the relationships between parameters. One typical design is Ecological Interface Design (EID). EID groups parameters not only based on the physical relationships, but also on their functional relationships [7]. EID designers have designed various graphs, such as configural graphs, to represent these relationships.

Algorithm Aids. This category of aids helps diagnosis by providing algorithms [8]. One kind of algorithm aids is based on system models. The malfunction of the system can be founded by comparing the difference between actual parameter value and the computing result. Another kind of algorithm aid is based on signals. The algorithm is to extract the features of the parameter values, and then the reason of fault could be found. This aid category is applied to the systems that could be described in mathematical language.

Knowledge Aids. This category of aids uses operating history information as knowledge to help operators diagnose the fault of the system [9]. They could be qualitative or quantitative. Qualitative methods include expert systems and qualitative trend analysis (QTA); these methods are experience-based. Quantitative methods include statistical methods like PCA (Principal Component Analysis), and non-statistical methods such as neural network approach. Knowledge aids could be applied to complicated systems that are very hard to model, but have accumulated abundant historical data.

Despite the existing diagnostic aids, supporting diagnosis by assisting information acquisition from displays is still necessary. Acquiring information is the first step of cognitive task. The obtained information is the basis for environment perception, fault diagnosis, and other cognitive activities. Computerization of control room HSIs greatly assists operators: Various flexible graphs can be presented, complex computation can be implemented, and massive data can be stored. However, the limited size of screens and the large amount of data are in a dilemma. This dilemma may cause "keyhole effect", which reduces operators' perception of the whole system and may cause operating errors [10]. Acquiring the right information in time and reducing the negative effects of massive information are also important to improve diagnosis performance.

Solutions focusing on information acquisition in MCRs have been developed. Navigation systems are designed so that operators can access key displays through buttons on displays. In complicated systems like NPPs, a navigation display is necessary. In other application areas like document databases and online maps, where massive data also exist, the query function is widely used. The query function usually contains a query box for a quick access to target and related information. Advanced query functions could help users obtain information by several categories, or get more related information by fuzzy search. Complicated industrial systems have the similar problem of large information amount, and query functions may be applied. The operator could directly access target parameters instead of entering into a navigation system.

Chen [11] found that information acquisition method and information quantity had an interactive effect. With small information quantity, the participants that used a search acquisition method achieved high diagnosis accuracy; while with large information quantity, the participants that used a query acquisition method achieved high accuracy. This study proved the effect of a query function in NPP MCR diagnosis. However, other information acquisition methods such as filtering have not been studied yet.

In real task circumstances, the available time for accomplishing tasks, especially diagnosis under emergencies, is mostly limited. When the available time is close to or less than the needed time, and a psychological stress comes out, time pressure comes into being [12]. Time pressure could change people's behaviors and further affect their task performance. Under high time pressure, people may tend to use simpler ways to solve a problem [13], to use one method instead of more ones to solve the problem [14], or to consider less information to accomplish the task [15]. High time pressure may cause low task accuracy, high mental workload, and people's lower confidence in their judgement [14]. In a study on the students' decision-making situations, researchers found that an inverted U-shape curve existed between information overload and time pressure [16].

Time pressure may have interaction effects with display design. It was found that under mid-range task time, participants using graphical displays achieved lower accuracy than those using text displays; but under tight task time, participants using graphical displays achieved higher accuracy (e.g., [17]). The task circumstances for NPP operators are mostly under limited time, and it is important for them to accomplish diagnosis or other tasks quickly and effectively, to reduce the bad consequences in safety and economy. This study was aimed to investigate the effects of several information acquisition methods under different time pressures.

2 Materials and Methods

A simulated simplified NPP system was developed for the experiment. Displays with specific information acquisition methods were designed, which were named *Query*, *Filter*, and *Query plus Filter*, as well as conventional *Search* (used as the control group). Participants used either of the designed displays to diagnose the abnormalities occurring in the simulated system. Several performance indicators were recorded to evaluate the effectiveness of the design.

2.1 Displays with Different Information Acquisition Methods

The simulated NPP system contains one heat generator, two hot-legs, two cold-legs, two steam generators, one pressurizer, one steam turbine, one electric generator, one condenser, one heater, and all the necessary pipelines and valves that connect the above components. The system can be divided into two parts, the primary loop (i.e., the coolant circulation) and the secondary loop (i.e., the cooling water circulation). Parameters such as temperatures, liquid levels, pressures and flows, are all presented on the display to indicate the status of the system. The tasks were to diagnose the leaking locations of the simulated system.

The four displays adopt different information acquisition methods. The conventional Search display consists of three pages: an overview page of the simulated plant, a page with process diagram of the primary loop, and a page with process diagram of the secondary loop. The other three displays have a similar layout, as Fig. 1 shows. The information acquisition panel provides the access of parameter data, while the process and instrument diagram (P&ID) panel only shows the components and the names of the available parameters. Thus, the value of a certain parameter can be accessed only by using one specific acquisition method in the information acquisition panel. All the possible leakage locations were listed on the ‘diagnosis panel’ where the participants can choose one option as his diagnostic hypothesis. The top right corner lays the countdown clock that shows the remaining time during one scenario. This countdown was expected to put time pressure on the participants.

The *Query* display has one query box and one watch list, as Fig. 2 shows. The participants could type in the first Chinese character of the parameter name, and see the parameters with the same first character. One or more parameters of interest can be added into the watch list, and the participants can trace the real-time value of these parameters.

The *Filter* display has one filtering zone as well as one watch list, as Fig. 3 shows. The filtering zone has two filter boxes. One classifies the parameters by sub systems (e.g., coolant circulation or nuclear reaction), and the other classifies the parameters by parameter types (e.g., temperature or pressure).

The *Query + Filter* display contains all the features of the Query and Filter displays. The participants could use either the query box or the filtering zone to acquire parameter values. The parameters of interest could be added in the watch list.

2.2 Experimental Design

A two factor (4 × 2) mixed experimental design was used. Information acquisition method was the between-subjects factor with four levels (i.e., Query, Filter, Query + Filter, and Search), and time pressure was the within-subjects factor with two levels (i.e., low and high).

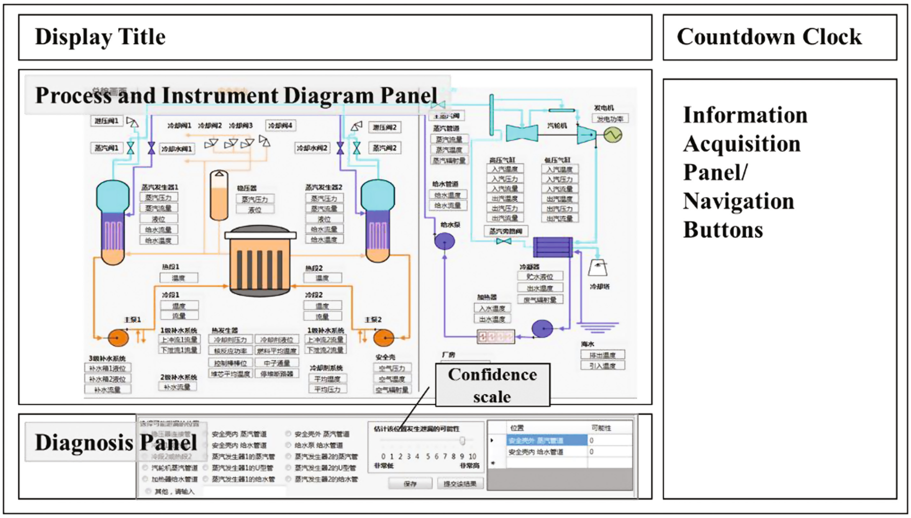
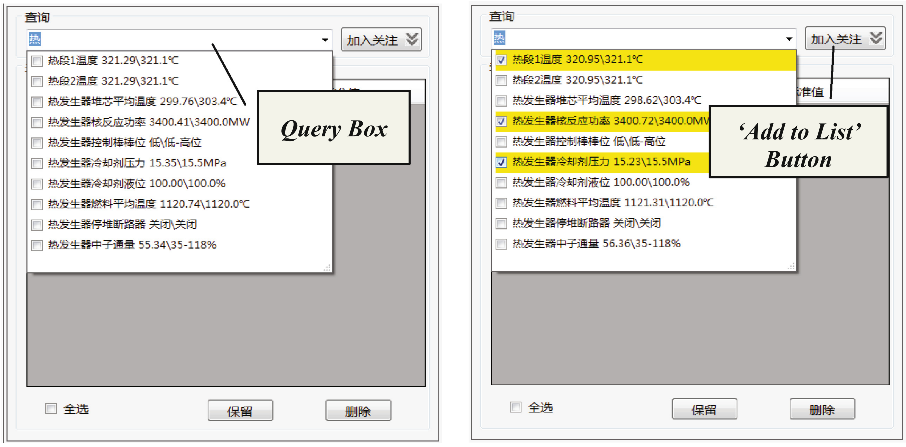


Fig. 1. The general display layout.



(a)

热

☐ 热段1温度 321.29\321.1℃
☐ 热段2温度 321.29\321.1℃
☐ 热发生器堆芯平均温度 299.76\303.4℃
☐ 热发生器核反应功率 3400.41\3400.0MW
☐ 热发生器控制棒棒位 低\低-高位
☐ 热发生器冷却剂压力 15.35\15.5MPa
☐ 热发生器冷却剂液位 100.00\100.0%
☐ 热发生器燃料平均温度 1120.74\1120.0℃
☐ 热发生器停堆断路器 关闭\关闭
☐ 热发生器中子通量 55.34\35-118%

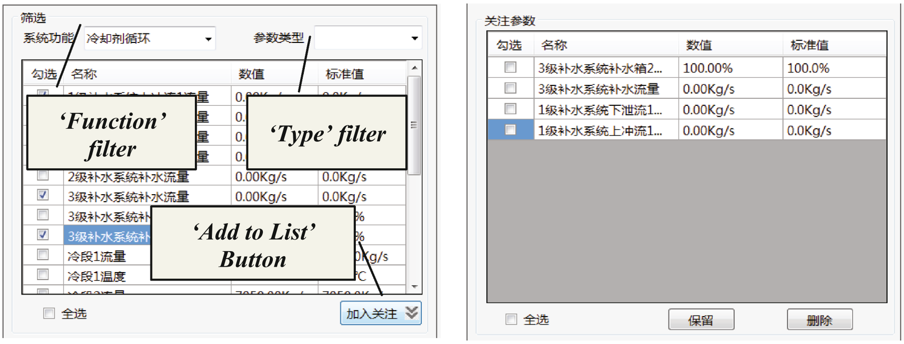
(b)

热

☑ 热段1温度 320.95\321.1℃
☐ 热段2温度 320.95\321.1℃
☐ 热发生器堆芯平均温度 298.62\303.4℃
☑ 热发生器核反应功率 3400.72\3400.0MW
☐ 热发生器控制棒棒位 低\低-高位
☑ 热发生器冷却剂压力 15.23\15.5MPa
☐ 热发生器冷却剂液位 100.00\100.0%
☐ 热发生器燃料平均温度 1121.31\1120.0℃
☐ 热发生器停堆断路器 关闭\关闭
☐ 热发生器中子通量 56.36\35-118%

(a) The matching function (b) Multiple-choice function

Fig. 2. The Query acquisition panel



系统功能 冷却剂循环
参数类型

勾选	名称	数值	标准值
☐	2级补水系统补水流量	0.00Kg/s	0.0Kg/s
☑	3级补水系统补水流量	0.00Kg/s	0.0Kg/s
☐	3级补水系统补		%
☑	3级补水系统补		%
☐	冷段1流量		0Kg/s
☐	冷段1温度		C

☐ 全选

勾选	名称	数值	标准值
☐	3级补水系统补水箱2...	100.00%	100.0%
☐	3级补水系统补水流量	0.00Kg/s	0.0Kg/s
☐	1级补水系统下泄流1...	0.00Kg/s	0.0Kg/s
☐	1级补水系统上冲流1...	0.00Kg/s	0.0Kg/s

☐ 全选

(a) The parameter store list (upper) (b) The watch list (lower)

Fig. 3. The upper and lower areas of filter acquisition panel

The dependent variables included three categories, which are diagnostic performance, subjective evaluation, and diagnostic process records. Diagnostic performance indicators included diagnosis time (the duration from the trial start to the participants' submission of final diagnosis), diagnosis accuracy (the number of correctly diagnosed trials divided by the number of all trials). Subjective evaluations included mental workload (measured in NASA-TLX scale) and confidence in diagnosis (subjectively rated by the participants in a scale ranging from 0 to 10). Diagnostic process records included the evaluation of diagnosis process (here named EDP, measured in a score given by the experimenter), number of hypotheses, and the time of first correct hypothesis (here named FCH time).

Each participant was required to diagnose the leakage locations in eight scenarios. Four scenarios were under high time pressure and the other four were under low time pressure. The performance data of the four scenarios with the same time pressure were averaged.

2.3 Participants

Eighty participants were recruited through the leaflets posted on bulletin boards on the campus. The participants were male undergraduate students of Tsinghua University, and they all came from science and engineering majors. This was to ensure that they had the similar knowledge basis and would be able to accomplish the tasks after a short training. As the operators in Chinese NPPs are all men, this study only recruited male participants. The participants were randomly allocated into four groups, and each had 20 participants. Each group used one of the four displays with associated information acquisition methods. Within each group, the orders of the two time pressures and the orders of the task groups were counterbalanced.

2.4 Experimental Procedure

The experiment consisted of three stages, introduction, training, and formal experiment. The whole process lasted about 2 to 3 h.

At the very beginning, the experimenter introduced the experiment and notices of the experiment. The participants signed the informed consent form voluntarily. After the introduction, the experimenter throughly explained the principle, structure, and parameters of the simplified NPP system. The participants had to pass a test to ensure their mastery of the necessary knowledge. Then the experimenter instructed the participants to get familiar with the experiment platfrom. The participants also practiced one scenario using the platform, which was to help them better understand the experiment.

After a short break, the participants entered into the formal experiment stage. At the end of the experiment, a short interview was given to each participant in order to collect their experience on the display. At last, the participants got their reward.

3 Results and Analysis

The means and standard deviations of all the dependent variables were listed in Table 1. The normality and the homogeneity of variance were first tested. Diagnosis time, subjective mental workload, and FCH time all passed these test, thus a repeated measure analysis of variance was used to analyze the effects of information acquisition method and time pressure. Diagnosis accuracy, diagnosis confidence, and EDP did not pass these tests, thus nonparametric tests were used.

Table 1. Descriptive statistics of dependent variables, in M(SD)

	Display					Time pressure	
	N	Search	Query	Filter	Query + Filter	High	Low
Diagnosis time (s)	80	221.66 (67.12)	284.26 (75.23)	269.01 (82.83)	289.50 (66.82)	234.88 (56.11)	297.37 (83.27)
Diagnosis accuracy	80	0.63 (0.29)	0.53 (0.27)	0.66 (0.23)	0.55 (0.32)	0.56 (0.32)	0.63 (0.25)
Mental workload	80	29.93 (9.20)	37.26 (9.12)	34.20 (7.49)	37.88 (7.75)	36.35 (9.35)	33.28 (8.25)
Diagnosis confidence	80	8.19 (1.65)	8.60 (1.41)	8.16 (1.98)	7.70 (2.10)	8.03 (1.90)	8.30 (1.73)
FCH time (s)	73	199.80 (66.90)	250.33 (87.06)	235.92 (82.82)	264.99 (77.37)	205.80 (65.63)	265.55 (85.02)
Hypotheses number	80	1.16 (0.29)	1.16 (0.37)	1.38 (0.51)	1.18 (0.30)	1.14 (0.32)	1.30 (0.43)
EDP score	80	3.53 (1.49)	3.09 (1.37)	3.79 (1.15)	3.31 (1.65)	3.20 (1.56)	3.09 (1.37)

The effect of diagnosis time was significantly affected by time pressure ($F_{(1,\ 76)} = 34.12$, $p < 0.001$). Under the high time pressure, diagnosis time was shorter. Meanwhile, information acquisition method also had a significant effect ($F_{(3,\ 76)} = 8.76$, $p < 0.001$). The participants using the Search display were much quicker than those using the Query, Filter, and Query + Filter displays. However, the latter three displays did not have significantly different effects on diagnosis time. No significant interaction effects were found.

Data of diagnosis accuracy did not satisfy normality but satisfied the homogeneity of variance. In each display type, a Wilcoxon signed rank test was conducted. In all conditions, diagnosis accuracy under the low time pressure was higher than that under the high time pressure, but only in Query + Filter display the difference was marginally significant (p = 0.065). Under each time pressure, a multiple-independent-samples Kruskal-Wallis test was used and no significant effects were found ($p_{low\ time\ pressure}$ = 0.463, $p_{high\ time\ pressure}$ = 0.168).

A repeated measure of variance analysis was used to analyze mental workload. Time pressure had a significant influence. Mental workload under high time pressure was higher than that under low time pressure ($F_{(1,\ 76)}$ = 11.82, p = 0.001). Information acquisition method also showed significant effect, with p = 0.004. In the Search display, the subjective mental workload was the lowest. However, the differences among Query, Filter and Query + Filter were not significant.

A Box-Cox transformation was used for diagnosis confidence, and the normality was almost satisfied. Time pressure had a marginal significant effect ($F_{(1,\ 76)}$ = 3.408, p = 0.069). The diagnosis confidence was lower under high time pressure. Information acquisition method and its interaction with time pressure did not show any significant effects.

A mixed model analysis of variance was conducted for the analysis of FCH-time. This was because that the number of samples was not the same in all the groups, as some participants did not give any correct hypothesis. Time pressure gave significant influence ($F_{(1,\ 71.453)}$ = 34.12, p < 0.001). Under high time pressure the FCH time was shorter. Information acquisition method also showed significant effect ($F_{(1,\ 71.399)}$ = 4.942, p = 0.004). The participants using the Search display were quicker in coming out of the first correct hypothesis. The other three displays did not show significant difference on first correct hypothesis time.

A non-parametric test was used in the analysis on number of hypotheses. The interaction between the two factors cannot be tested and thus the factors were respectively tested. A Wilcoxon signed rank test was used to check the effect of time pressure; and it had a significant influence (p < 0.001). The participants gave fewer hypotheses under high time pressure.

The normality of EDP score was mostly fulfilled; thus a repeated measure of variance analysis was conducted. Time pressure again significantly affected the score ($F_{(1,\ 76)}$ = 6.399, p = 0.013). The score under high time pressure was lower than that under low time pressure. Information acquisition method and its interaction with time pressure did not show any significant effects.

4 Discussion

In summary, information acquisition method and time pressure both had significant effects on the process and performance of the diagnosis task. However, no interactive effects were found between these two factors. The Query display and the Query + Filter display both led to poor diagnosis performance and process results. The Filter display did not show much better than these two displays, either. Meanwhile, the Search display users had the shortest diagnosis time, the shortest FCH time and the lowest mental workload. This study showed the newly designed information methods were not superior to the conventional visual search method.

These conclusions did not fully replicate Chen's experimental results [11], in which the query display led to shortest diagnosis time and highest accuracy under high display complexity. In addition, although advantages of function information and functional grouping of information have been verified (as previously presented in studies like [18, 19]), the Filter display did not support diagnosis strongly. These may be caused by the design of query functions or the limited knowledge of the participants. Interaction actions like 'selecting' and 'adding into list' may hinder the participants' diagnosis process by consuming more time in typing the parameter names. As the experiment platform is a simplified NPP system, the amount of parameters and displays is not really large. Therefore, the advantage of the query function could be weakened. In addition, the participants were students who were trained to get familiar with principles of a NPP. Although with enough knowledge to perform the tasks, they were not as fluent in the parameter names and system structures as real NPP operators, and thus inputting target parameters was less convenient than direct visual searching.

The effect of time pressure was consistent with previous studies. Under high time pressure, the participants accomplished the diagnosis faster, showed higher mental workload, had lower confidence for diagnosis result, generated fewer hypotheses, and used less time to get the first correct hypothesis. The tighter time forced the participants to accelerate the process of information searching, to generate less hypotheses, and to be less confident in their diagnosis.

5 Conclusion

The digitization of displays in complicated industrial systems extends the possibility of enhancing operating safety and performance. This study applied the widely used information acquisition method, query and filter, into the design of interfaces for a simulated NPP. The designs were intended to help relieve operators' navigations through displays and thus to improve diagnosis performance. The experiment found that the currently widely used search display did better than other displays in most performance indicators. In the overview display of a system, such a method should be adopted and well designed. To observe the advantages of the other information acquisition methods, further studies are required under more complex simulated systems or even in industrial field settings.

Acknowledgments. This study was supported by the National Natural Science Foundation of China (project No. 71371104).

References

1. Rasmussen, J.: Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Trans. Syst. Man Cybern. SMC **13**(3), 257–266 (1983)
2. Weber, E.U., Böckenholt, U., Hilton, D.J., Wallace, B.: Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. J. Exp. Psychol. Learn. Mem. Cogn. **19**(5), 1151–1164 (1993)

3. Kim, J.H., Lee, S.J., Seong, P.H.: Investigation on applicability of information theory to prediction of operator performance in diagnosis tasks at nuclear power plants. IEEE Trans. Nucl. Sci. **50**(4), 1238–1252 (2003)
4. Zhou, Y., Zhang, L.: Analysis of operators' diagnosis tasks based on cognitive process. Nucl. Power Eng. **33**(5), 69–75 (2012)
5. Vicente, K.J., Moray, N., Lee, J.D., Rasmussen, J., Jones, B.G., Brock, R., Djemil, T.: Evaluation of a rankine cycle display for nuclear power plant monitoring and diagnosis. Hum. Factors J. Hum. Factors Ergon. Soc. **38**(3), 506–521 (1996)
6. Burns, C.M., Hajdukiewicz, J.: Ecological Interface Design. CRC Press, Boca Raton (2004)
7. Vicente, K.J.: Ecological interface design: progress and challenges. Hum. Factors J. Hum. Factors Ergon. Soc. **44**(1), 62–78 (2002)
8. Gao, Z., Cecati, C., Ding, S.X.: A survey of fault diagnosis and fault-tolerant techniques part i: fault diagnosis with model-based and signal-based approaches. IEEE Trans. Ind. Electron. **62**(6), 3757–3767 (2015)
9. Gao, Z., Cecati, C., Ding, S.X.: A survey of fault diagnosis and fault-tolerant techniques part ii: fault diagnosis with knowledge-based and hybrid/active approaches. IEEE Trans. Industr. Electron. **62**(6), 3768–3774 (2015)
10. Kim, J.H., Seong, P.H.: Human factors engineering in large-scale digital control systems. In: Seong, P.H. (ed.) Reliability and Risk Issues in Large Scale Safety-critical Digital Control Systems, pp. 163–195. Springer, London (2009)
11. Chen K.J.: Diagnosis task performance influencing factors in digitalized industrial systems. Ph.D. dissertation, Tsinghua University, Beijing (2015)
12. Svenson, O., Edland, A.: Change of preferences under time pressure: choices and judgements. Scand. J. Psychol. **28**(4), 322–330 (1987)
13. Svenson, O., Benson, L.: Framing and time pressure in decision making. In: Svenson, O., Maule, A.J. (eds.) Time Pressure and Stress in Human Judgment and Decision Making. Springer, New York (1993)
14. Edland, A., Svenson, O.: Judgement and decision making under time pressure. In: Svenson, O., Maule, A.J. (eds.) Time pressure and stress in human judgement and decision making, pp. 27–40. Springer, New York (1993)
15. Braun, R.L.: The effect of time pressure on auditor attention to qualitative aspects of misstatements indicative of potential fraudulent financial reporting. Acc. Organ. Soc. **25**(3), 243–259 (2000)
16. Hahn, M., Lawson, R., Lee, Y.G.: The effects of time pressure and information load on decision quality. Psychol. Mark. **9**(5), 365–378 (1992)
17. Coury, B.G., Boulette, M.D.: Time stress and the processing of visual displays. Hum. Factors J. Hum. Factors Ergon. Soc. **34**(6), 707–725 (1992)
18. Ham, D., Yoon, W.C.: The effects of presenting functionally abstracted information in fault diagnosis tasks. Reliab. Eng. Syst. Saf. **73**(2), 103–119 (2001)
19. Wu, X.J., Gao, Q., Song, F., Dong, X.L., Li, Z.Z.: Display design for AP1000 nuclear power plant: FBTA method vs. traditional method. In: Paper Presented at the 11th International Conference on Probabilistic Safety Assessment and Management & Annual European Safety and Reliability Conference 2012 (PSAM 2011 & ESREL 2012), Helsinki, Finland (2012)

Cognitive Processes and PEMs of TSA in Digitized MCRs of NPPs

Pengcheng Li[1,2](✉), Licao Dai[1], Li Zhang[1], Yu Jiang[1],
and Xiaofang Li[1]

[1] Human Factor Institute, University of South China,
Hengyang 421001, Hunan, China
lipengcheng0615@163.com
[2] Systems Engineering and Engineering Management,
University of North Carolina at Charlotte, Charlotte 28223, USA
pli6@uncc.edu

Abstract. In order to identified macro-cognitive processes and psychological error mechanisms (PEMs) of TSA in digital NPPs. Firstly, the components of TSA are identified according to team task analysis, which including individual situation awareness (ISA), shared situation awareness (SSA) and mutual awareness (MA). Furthermore, the macro-cognitive processes of TSA are identified respectively from ISA, SSA and MA on the basis of simulation experiments, and the process model of TSA is established to illustrate the forming mechanism of TSA. Finally, the PEMs of TSA errors are identified, it is very useful to prevent and reduce TSA errors.

Keywords: Team situation awareness · Macro-cognitive process · Psychological error mechanism · Nuclear power plants

1 Introduction

Team Situation Awareness (TSA) is more important for team tasks. It is attributed to the monitoring of operating state of systems and the disposition of abnormal state are completed by team in complex industrial systems such as nuclear power plants (NPPs). Team are important units of organizational work because they bring diverse expertise, skills, and resources to complex tasks that may be too large or complex for a single individual to undertake [1]. Researches show that system safety are more dependent on team performance rather than individual performance in complex dynamic systems such as NPPs and air traffic control [2], and team performance is positively correlated with TSA [3]. Furthermore, if individuals make SA errors, it may be detected, corrected and restored by other team members. Consequently, there is no doubt that the TSA is also critical to the successful execution of team tasks.

So far, there are some researches on cognitive model and error mechanism of TSA, but few researches in digital NPPs. Salas et al. [4] point out that, due to the cognitive nature of team SA, research into the construct is difficult, deficient and complex. As a consequence, team SA suffers from a similar level of contention as ISA theory does. It leads to the absence of a uniform and known-well structure or model to describe the cognitive processes or mechanisms of TSA. Although there were some TSA model to

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_31

illustrate the interactive processes of a team, for example, Salas et al. [5] proposed a framework/ model of TSA, suggesting that it comprises two critical processes, ISA and team processes. Salmon PM et al. [6] established a TSA model on the basis of predecessors' work, which includes environment data, SA development, ISA, team processes and TSA. Although the literature discussed above made contributes to analysis of cognitive process and error mechanism of TSA, they only analyzed or partly illustrated the constituent elements and interaction process of TSA, and not analyzed much aspects of error mechanism of TSA, such as TSA error classification/modes, performance shaping factors (PSFs) and psychological error mechanisms (PEMs) etc, especially cognitive processes and PEM of TSA in digital NPPs. Therefore, in terms of the characteristics of context in digital MCR of NPPs, we need to analyze and identify TSA cognitive processes and error mechanisms in detail, clearly identify possible error modes, PSFs and PEMs of TSA, this work is of great significance for the prevention of TSA errors.

2 Forming Mechanism of TSA in Digital MCRs

According to previous researches [6, 7], we think that TSA should include three basic elements: ISA, SSA and MA in digital MCR of NPPs. The forming processes model of TSA is established as shown in Fig. 1 on the basis of the results of literature research and our simulator experiment observation. The model mainly includes the forming processes of TSA, PSFs influencing TSA and macro-cognitive processes of TSA. When an abnormal event occurs, one or more alarms and information will appear and display on the shared display screens and computer-based workstation, the operators will observe, check and understand the occurred alarms and information to form individual aware of system/unit state by ISA processes. In order to confirm this system/unit state, team members will form TSA by team interactive processes such as communication, collaboration, cooperation and mental processes. The following aspects will illustrate other components of the model in detail to explain the forming mechanism of TSA on the basis of the field observations during operators' retraining, simulation experiments and team task analysis from a perspective of human factors.

2.1 Cognitive Processes of ISA

Endsley [8] views ISA as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". According to Endsley's definition, SA process includes a series of cognitive activities. The classical Endsley's SA model [8] demonstrates the cognitive processes of ISA on the basis of information processing theory. However, it merely demonstrates the macro-cognitive processes without explaining the cognitive activities involved in each macro-cognitive process. If we can find the more specific cognitive functions or activities, it is useful to make specific preventive measures for preventing SA error. According to the previous research [9], the operator's cognitive activity is monitoring/detection by "seeing/listening", "information searching or

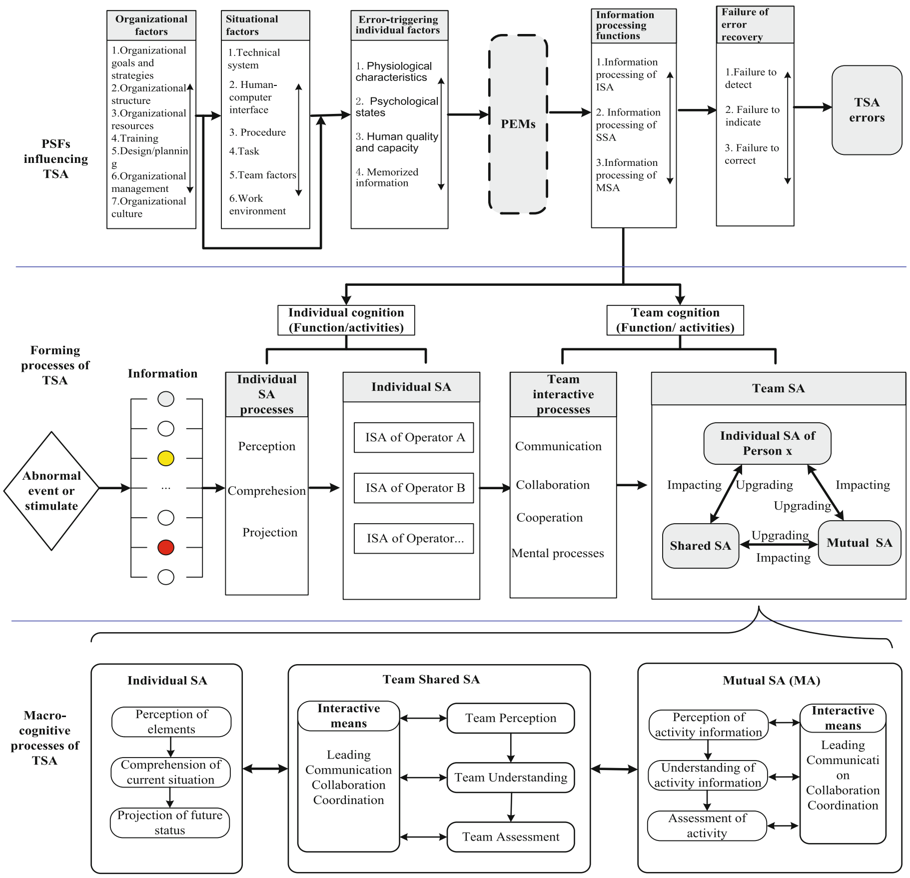


Fig. 1. The forming processes of TSA

locating", "information recognition" and "information verification" for single information, and their activities are information filtering, screening, etc. For more information, which will be combined into a cognitive function that is called "multiple information gathering". The cognitive functions of operator in situation assessment process also involve "information comparing", "information integrating", "state explanation", "cause identification" and "state projection". Therefore, the main cognitive activities are identified.

2.2 Cognitive Processes of SSA

Jones and Endsley [10] used the term SSA to describe the extent to which different team members have the same information about an aspect of the task that is relevant to them. How does SSA develop? Generally speaking, the formation of SSA between operators (such as RO1 and RO2) by communication is rarely observed in digital MCR of NPPs, this may be due to the design features of the digital control system (for example, shared information help them form SSA), the complexity of tasks and severity of incidents (for example, the task is more simple, they don't need shard SA) etc. However, in the case of more serious emergencies (such as SGTR, small LOCA, SGTR+loss of power), it needs a SSA for a team in order to better mitigate and handle accidents. We find that there are three main macro-cognitive stages for the forming of SSA of team on the basis of observation data, namely team perception, team understand, and team assessment, and these macro-cognitive stages are completed by the means of leading, communication, collaboration and coordination. In each macro-cognitive stages, there are various cognitive activities. ISA and Mutual SA (namely MA) provide support for the forming of SSA, and they dynamically interact with each other as shown in Fig. 1.

Team Perception. Team perception is the first stage to form SSA for handling incidents/accidents. We define team perception as a organized process includes information collection, perception and pre-disposing actions taken by teammates in order to build their own knowledge [11, 12]. In the stage of team perception, each operator recognizes an occurrence of a abnormal event when a alarm or lots of alarms occur, then they will check the alarm(s), and collect the main information related to the alarm(s) through 4 large shared display screens. If a more serious accident occurs, the purple alarm will appear, operators must execute SOP to collect information related to system state and implement relevant operation using computer-based workstations and large display screens. However, If a more complex accident occurs, which may not be covered by SOP, then it is difficult to solve this issue such as superimposed accidents. The operation team must best solve it according to the serious accident handling plan and strategy. At this time, shift supervisor or other leaders carry out the arrangement of organizational elements (members, information etc.) and provide guidance, instruction, direction and leadership to other members of the team for the purpose of achieving a common goal (such as forming a SSA to mitigate accidents). The perception process is regulated by team leader. The leader of team will organize the other members to collect required information and discuss the related issues, including cognitive activities of convening of members, arranging of related resources, determining the topic, determining of the priority of issues, controlling the time of discussion, correcting the mistakes made by other people, guiding the discussion. These work of the process are mainly prepared by leader of the team, the other people exclusively follow the command of the leader to prepare related elements (for example, their own information collection, the pre-disposing of their knowledge and organizing of their viewpoint, information and view expression).

Team Understanding. After a team achieves the collection and handling of relevant information, the team needs to integrate their information to build team knowledge for understanding system state. According to the viewpoint provided by Klein et al. [13], team sensemaking (understanding) is defined as the process by which a team manages and coordinates its efforts to explain the current situation and to anticipate future situations, typically under uncertain or ambiguous conditions. If successful, the outcome of the team sensemaking process is collective understanding of the situation. In the stage of team understanding, which is a kind of process of team knowledge building, here the team also has three primary processes: First, the process of team information exchange involves passing relevant information to appropriate teammates at appropriate times. Second, the process of team common understanding involves shard data or knowledge analysis, explanations and interpretations shared data between team members or with the team as a whole to identify possible system state by relying on rules, procedures and pre-established criteria etc. Third, to anticipate possible occurrence of event and state to assess and validate the results of team understanding.

Team Assessment. Furthermore, after team understanding, team may have reached a preliminary result for the discussion issue, then operating team should assess whether the preliminary result is in line with the actual situation and give a logical and reasonable explanation for forming a consistent and correct cognition of the state of unit/system. Team's assessing and validating of a consistent SA mainly refers to a process that team members are to analyze, test and validate their mental model to gradually converge for forming a unifying and consistent mental model. In assessing process, team members will carry out various actions to validate the provided preliminary hypotheses to reach a agreement. The activities contain provision of hypotheses of a state of unit/system, analysis of rationality of hypotheses, validation of hypotheses by information collecting, and agreeing with and confirming the assumption and reporting it among team.

Team Interactive Means. Team cognition or group think is realized by various team interactive means, which includes leading, communication, collaboration, coordination. O'Connor, et al [14] have explained the meaning of communication, collaboration and coordination, and indicate that communication, collaboration, and coordination are main ways of teamwork. In order to transfer individual knowledge into team knowledge, the team knowledge building process is a process of information processing, exchanging and updating [15]. Whereas information processing and mental model updating are internal processes, the communication of information occurs external to the team members and, therefore, represents an observable component of the convergence process [16]. According to the interactive model illustrating communication processes between members [17], we can see that the cognitive activities such as information sending, information receiving, information conveying using media tools such as telephone, information understanding, information feedback, are involved in the team's communication. In addition, team members also need to cooperate to quickly achieve a SSA in the communication process, which include searching additional information, understanding information, monitoring state of unit/system, information feedback etc. Therefore, there is no doubt that the individual's cognitive processes and activities produce complex interactions between team members to form team cognitive processes and activities.

2.3 Cognitive Processes of MA

Under the emergency of accident occurrence, the activities of team members not only characterize ISA and their SSA, but also reflect the MA in team's communication and coordination. For example, in the handling of many accidents such as SGTR in digital

NPPs, it is necessary that the RO1 of primary loop require the closely coordination of RO2 of secondary loop to mitigate successfully accidents, so they need MA. Shu and Furuta [18] think that MA refers to the awareness that individuals of a cooperative entirety have of each other's activities, beliefs, and intentions. This definition is similar to the concept of workspace awareness (WA) provided by Gutwin et al [19]. WA is the up-to-the-moment understanding of another person's interaction with a shared workspace. WA involves knowledge about where others are working, what they are doing, and what are going to do next. This information is useful for many of the activities of collaboration such as coordinating action, managing coupling, talking about the task, anticipating other's actions. Therefore, understanding of MA (or workspace awareness) is an essential element in cooperative activities and plays an important role in the fluidity and naturalness of collaboration.

MA between team members will occur by oral communication and non-verbal one (visual observation, movement changes, gestures, etc.) to obtain other people's information. For example, in the condition of an accident, the shift supervisor (SS) will move to the operator's workstation, he or she will monitor or observe what RO1 or RO2 are doing to understand their current work situation. If SS find that operators implement a inappropriate task/action or make a mistake, he or she will timely remind them to correct it. Again, in the process of RO1 mitigating an accident in accordance with procedures, the operator will ask RO2 to coordinate his or her operating and controlling activities on the basis the requirements or commands of SOP, and will tell RO2 what he or she should do in this or next time. On the basis of the simulation experiments observation, the macro-cognitive processes of MA can be divided into three stages, namely perception of activity information, understanding of what other members are doing (understanding of activity information), assessing of what they are doing and predicting of what they will do in the future, and he or she is ware of beliefs of other people(assessment of activity). For example, in assessment stages, he or she will ask some questions in his or her mind: Why are they doing this? Is this correct? What do they think? How much should I believe them? etc.

2.4 The Performance Shaping Factors of TSA

Even though a number of researchers have paid attention to TSA because of its contribution for team performance, there has not been much consideration of the influencing factors of TSA that can show a causal relationship between it and TSA, it may due to research complexity and difficulty. So far, there are few literature to more comprehensively study PSFs impacting SA. For example, Endsley's three-level model of situation awareness analyzed performance shaping factors (PSFs) influencing ISA. [8]. Lee et al [20] specified what factors affect the shared mental models, which include team knowledge, team skills, team Attitudes, team dynamics, team environment. But the methods do not analyze in detail the PSFs impacting TSA as their main research purpose. According to our previous study [9], operator's SA mainly is influenced by individual factors, situational factors and organizational factors based on the established organization-oriented "structure-behavior" model for HRA. We think that TSA should also be impacted by these PSFs more or less. The forming processes of TSA errors and the PSFs are shown in Fig. 1.

3 PEMs of TSA Errors in Digital MCRs of Digital NPPs

There are many classification methods of human error. Reason [21] divided classification of human error into three kinds of categories: behavior-based, context-based and concept-based classification. We can classify the TSA errors into three categories: Individual SA errors; Shared SA errors; MA errors according to the composition elements of TSA. We also can classify TSA errors based on TSA processes, TSA cognitive functions etc. We think the process-based classification of TSA error is very useful to the identification of error origins or stages and error mechanisms of TSA. However, the product-based classification of TSA is also very useful to the measurement of TSA errors. Therefore, we adopt the latter to determine the classification of TSA errors, involve individual SA errors, shared SA errors and MA errors, and they are classified into more concrete sub-classification by combining with the key words such as omission, inadequate, wrong and delayed. As for psychological error causes or mechanism of TSA, we think this is the occurrence of human cognition bias and issue, so we identify these PEMs on the basis of human information processing model [22] and the combination with previous research results [21–26].

Different researchers provided different classification of error mechanisms, which is different at the detailed level of classification, abstraction level, size, and criteria of classification, and there is no uniform term related to PEMs. Therefore, in order to identify PEMs of TSA, we will classify the PEMs according to the elements of information processing model provided by Wickens et al. [22]. The error mechanisms will be classified as attention-related, memory-related, knowledge-related (long-term memory), process-related (perception, comprehension etc), other person-related characteristic, interaction-related, and terminology will be unified. These error mechanisms are classified, merged and integrated as shown in Table 1. Furthermore, the main PEMs of ISA and SSA are identified. The PEMs of MA is similar to PEMs of SSA and ISA, so we do not discuss it further. For example, coordinator monitors RO1's manipulation, simultaneously RO1 understands the coordinator is monitoring his work, so this MA can be seen as a teamwork from the perspective of team. When RO2 tells RO1 what he is doing, then RO1 know it, but RO2 do not know RO1 what he is doing, so it can be seen as individual information collection.

3.1 Psychological Error Mechanisms of ISA

According to Endsley's viewpoint [9], ISA is produced by three macro-cognitive processes: perception, comprehension and projection. Errors will happen in each cognitive stage because of the negative influencing of PSFs. Then, the proximate causes, PEMs are analyzed as listed in Table 2 according to event reports, small deviation event analysis and previous research results [10, 26]. In the stage of perception, the possible reasons mainly are derived from outside information and individual capability of

Table 1. Special error mechanism

PEMs classification	Special PEMs (some are similar)		
Attention-related	●Distraction/Interference/interruptions ●Unable to maintain vigilance; ●Divided attention	●Attentional capture/selective attention ●Attention-missing a change in cues ●Improper control of attention	●Reduced intentionality ●Overattention (Omissions; Repetitions Reversals) ●Preoccupation
Memory-related	●Memory block ●Slip of memory ●Negative transfer ●Lost memory of goal ●Forgetting an intention	●Memory failure ●Lapse of memory ●Working memory limitations; ●Loss of information in working memory	●Failure of prospective memory ●Memory capacity over ●Working memory capacity overflow;
Knowledge-related	●Default/lack of knowledge; ●Inadequate mental model ●Mental manipulation of the information is inadequate, inaccurate, or otherwise inappropriate ●Incorrect or inappropriate frame	●Incorrect knowledge; ●Incorrect mental model ●Misuse of knowledge ●Loss of activation ●Insufficient learning ●Learned carelessness/ frequential simplification	●Incomplete knowledge ●Inexperience ●No frame/mental model ●Mislearning
Process-related (cognitive function)	●Perceptual confusions ●Perceptual tunneling ●Confirmation bias ●Biased reviewing ●False assumption ●Misinterpretation ●Failure of understanding ●Misunderstanding ●Lack of awareness ●Cognitive overload ●Information/Stimulus overload ●Risk recognition failure ●familiar pattern not recognized ●information-detection failure ●Expectation bias ●Integration failure ●Decision freeze ●Functional confusion ●Mis-anticipation	●Failure to consider side-term effects ●Mistake alternatives ●Cognitive fixation ●Prioritisation failure ●Manual variability ●Habit intrusion ●Inappropriate intonation ●Misarticulation/Description ●Dysfluency ●Cross talks; ●Lack of correct rules ●Misfiring of good rules ●Encoding deficiencies in rule ●Misapplication of good rules ●Dissociation between knowledge and rules	●Illusory correlation Problems with complexity (hindsight bias) ●Spatial confusion ●Place-losing error ●False triggering ●Failure to trigger ●Mismatch between expected and actual cues ●Data not properly recognized, classified, or distinguished ●Improper integration of information or frames ●Improper comparison
Other person-related characteristic	●Human variability; ●Deviation of motor skills ●Manual variability; ●Halo effects; ●Short-cut ●Bounded rationality; ●Salience bias ●Recency bias	●Stereotype fixation ●Stereotype take-over ●Infrequency bias/frequency gambling; ●Tunnel vision; ●Freezing; ●Capture error	●Overconfidence; ●Similarity interference/matching ●Strong habit intrusions ●Faulty heuristics; ●Oversimplification;
Interaction-related	●Source error of omission ●Target error of commission ●Failure to verify that other operators have correctly performed their responsibilities	●Source error of commission ●Incorrect timing of communication ●Failure to consider information communicated by an individual	●Target error of omission ●Decision making failures ●Failure to iterate the communication process sufficiently

Table 2. The main PEMs of ISA

Occurrence stage	Proximate cause	Psychological error mechanism	Example
Perception error	●No, inadequate, wrong information ●Information salient/important/priority ●Operator's vigilance ●Operator's memory	●No attention ●Inadequate attention ●Unable to maintain vigilance ●Slip of memory, ●Lapse of memory ●Memory loss ●Information/stimulus overload ●Perceptual confusions	Operator's vigilance—When the alarm of failure of electrical disk occurred, operator did not respond, later he check and find the loss of power of entire electrical disk
Comprehension error	●Failed to understand information ●Failed to integrate information	●Memory failure, ●Memory capacity over ●Lack of knowledge and experience ●Lack of or incomplete mental model ●Incorrect knowledge ●Incomplete knowledge ●Cognitive overload ●Misunderstanding ●Integration failure	The Internal leakage of VVP101VV is detected, but because coordinator's vision is narrow or inadequate knowledge, he think that it is triggered by the failure of GCT121VV valve. It resulted in the cause leading to the leakage of VVP101VV is not detected
Projection error	●Failed to project situation into near future	●Biased reviewing ●False assumption ●Mis-anticipation ●Expectation bias ●Lack of Knowledge and experience ●Lack of or incomplete mental model	Click on RCP001VP valve, click 2 to 3 times to open it, but operator do not see the valve opening degree changed, so he directly use the button of fast open, the result shows that the valve directly whole open

information perception. In the stage of comprehension, the possible reasons mainly are derived from individual capability of information processing such as information integrating, reasoning. In the stage of projection, the possible reasons mainly are also derived from individual capability of information processing such as inference, forecasting. For example, an alarm occur, if the operator is lack of vigilance, he will pay no attention to this alarm because of his careless.

3.2 Psychological Error Mechanism of SSA

SSA is formed by complex interactive processes, in order to better explain this process, it can be approximately divided into three stages: team perception, team understanding, team assessment on the basis of simulation experiments. These processes are achieved by communication and cooperation of team members. In the stage of team perception, they share their viewpoints and information. Therefore, team perception errors are derived form shared information and interactive issues. In the stage of team understanding, team need to discuss what is going on with the system according to team perception information. Team understanding errors are mainly derived from common understanding of knowledge, rebuilding and team interactive issues in case of the team perception information is valid. In the stage of team assessment, team need to validate their judgment, so they need to further gather information, cues and evidences to confirm what they identified the state of system. Team assessment errors are mainly derived from obtained information, common understanding of knowledge, rebuilding and team interactive issues. The identified PEMs of SSA are listed in Table 3.

Table 3. The main PEMs of SSA

Occurrence stage	Proximate cause	Psychological error mechanism	Example
Team perception error	●Incomplete shared information ●Wrong shared information ●Too much shared inference information ●Inadequate information exchange	●No attention ●Inadequate attention ●Slip of memory ●Lapse of memory ●Memory loss ●Information/stimulus overload ●Incorrect knowledge ●Incomplete knowledge ●Incorrect mental model ●Misperception ●Inadequate interaction	When US left MCR, 3RRI001PO tripped, and 3RRI002PO has been in the isolated state because of overhaul. After that, he returned to the MCR, but the US did not collect information by himself to independently verify equipment status, he directly to ask RO2 the fault issue and discuss the cause of fault

(*continued*)

Table 3. (*continued*)

Occurrence stage	Proximate cause	Psychological error mechanism	Example
Team understanding error	●Failed to understand viewpoint of other person ●Inadequate information rebuilding ●Failed to project situation into near future	●Lack of Knowledge ●Incorrect knowledge ●Incomplete knowledge ●Lack of or incomplete mental model ●Inadequate interaction ●Misunderstanding ●Failure of understanding	AHP102/202VL valve is online abnormal shutdown, resulting in low efficiency for turbine operating,in the case of team do not know the reasons of the malfunction, they does not systematically use the ERP method to solve this issue, so they do not find the real reasons of the fault or abnormal for a long time
Team assessment error	●Incomplete information; ●Too much shared inference information ●Incomplete interaction; ●Failed to understand causal relationship	●No attention ●Inadequate attention ●Information/stimulus overload ●Lack of Knowledge; ●Incomplete mental model ●Inadequate interaction ●Confirmation bias ●Biased reviewing	SBLOCA accident occurs under the condition of RRA connected in NPPs, the water level of pressure vessel is below the mark of THL. According to the SOP guidelines, team should start SI into the ECP4 procedure; But they think that the procedure guideline is wrong without any analysis and assessment, so they decide to deal with the accident by using ECPR2 procedure

4 Conclusions and Discussions

TSA is critical to team performance in digital MCRs of NPPs. It is very significant that cognitive processes and error mechanisms of TSA are identified to prevent TSA errors. Therefore, we have identified constituent elements, main macro-cognitive processes, PEMs of TSA. Some conclusions are shown as follows.

1. TSA includes individual SA, shared SA and mutual awareness according to task analysis;
2. The macro-cognitive processes of shared SA is composed of team perception, team understanding, and team assessment, and MA includes perception of activity, understanding of activity, and activity assessment on the basis of simulation experiments observation.
3. The psychological error mechanisms of human errors are integrated to six classification: attention-related, memory-related, knowledge-related, process-related, other person-related characteristic, and interaction-related from the perspective of human information processing.

4. The main psychological error mechanism of ISA and SSA errors are identified on the basis of event reports and small deviation reports analysis.

Although the PEMs of TSA errors(or human errors) have been identified and classified, it is necessary to collect more event reports to expand the PEMs of TSA errors and identify the main PEMs to support TSA error prevention in the near future.

Acknowledgments. The financial support by Natural Science Foundation of China (71371070, 71301069, 51674145) and Scientific Research Program supported by Lingdong Nuclear Power Plant (KR70543), the Postdoctoral Science Foundation of China (2016M600633), the Natural Science Foundation of Hunan Province (2017-260), the Humanities and Social Science Projects of Ministry of Education (11YJC630207), Program of China Scholarship Council (CSC) and the Social Science Projects of Hunan Province (16YBA314) are gratefully acknowledged.

References

1. Salas, E., Fiore, S.M.: Team Cognition: Understanding the Factors that Drive Process and Performance. American Psychological Association, Washington DC (2004)
2. Banbury, S., Tremblay, S.: A Cognitive Approach to Situation Awareness: Theory and Application. Routledge, London (2004)
3. Kim, S.K., Park, J.Y., Byun, S.N.: Crew resource management training for improving team performance of operators in Korean advanced nuclear power plant. In: IEEE International Conference on Industrial Engineering and Engineering Management, vol. 44, pp. 2055–2059. IEEE Press, Hong Kong (2009)
4. Salas, E., Muniz, E.J., Prince, C.: Situation awareness in teams. In: Karwowski, W. (ed.) International Encyclopedia of Ergonomics and Human Factors. Taylor & Francis Ltd, Oxford (2006)
5. Salas, E., Prince, C., Baker, P.D., et al.: Situation awareness in team performance: implications for measurement and training. Hum. Factors **37**(1), 123–136 (1995)
6. Salmon, P.M., Stanton, N.A., Walker, G.H., et al.: What really is going on? review of situation awareness model for individuals and Teams. Theor. Issues Ergon. Sci. **9**(4), 297–323 (2008)
7. Yuan, X.H.: Effects of Mutual Awareness on Team Performance in Accident Diagnosis under Computerized Working Environment. Tsinghua University, Beijing (2013). (in Chinese)
8. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors **37**, 32–64 (1995)
9. Li, P.C., Zhang, L., Dai, L.C., et al.: Study on operator's SA reliability in digital NPPs. Part 1: The analysis method of operator's errors of situation awareness. Ann. Nucl. Energy **102**, 168–178 (2017)
10. Jones, D.G., Endsley, M.R.: Sources of situation awareness errors in aviation. Aviat. Space Environ. Med. **67**, 507–512 (1996)
11. Letsky, M., Warner, N., Fiore, S.M., et al.: Macrocognition in complex team problem solving. In: Proceedings of the Twelfth International Command and Control Research and Technology Symposium (12th ICCRTS), Newport, pp. 19–21 (2007)

12. Whaley, A.M., Xing, J., Boring, R.L., et al.: Building a Psychological Foundation for Human Reliability Analysis, NUREG-2114. U.S. Nuclear Regulatory Research, Washington D.C (2012)
13. Klein, G., Wiggins, S., Dominguez, C.O.: Team sensemaking. Theor. Issues Ergon. Sci. **11** (4), 304–320 (2010)
14. O'Connor, P., O'Dea, A., Flin, R., et al.: Identifying the team skills required by nuclear power plant operations personnel. Int. J. Ind. Ergon. **38**(11), 1028–1037 (2008)
15. Fiore, S.M., Smith-Jentsch, K.A., Salas, E., et al.: Towards an understanding of macrocognition in teams: developing and defining complex collaborative processes and products. Theor. Issues Ergon. Sci. **11**(4), 250–271 (2010)
16. McComb, S., Kennedy, D., Perryman, R., et al.: Temporal patterns of mental model convergence: implications for distributed teams interacting in electronic collaboration spaces. Hum. Factors **4**, 264–281 (2010)
17. Foulger, D.: Models of the Communication Process (2004). http://davis.foulger.info/research/unifiedModelOfCommunication.htm
18. Shu, Y., Furuta, K.: An inference method of team situation awareness based on mutual awareness. Cogn. Technol. Work **7**, 272–287 (2005)
19. Gutwin, C., Greenberg, S.: A descriptive framework of workspace awareness for real-time groupware. Comput. Support. Coop. Work **11**, 411–446 (2002)
20. Lee, M., Johnson, T., Lee, Y., et al.: The Conceptual Framework of Factors Affecting Shared Mental Model, pp. 561–565. Association for Educational Communications & Technology (2004)
21. Reason, J.: Human Error. Cambridge University Press, New York (1990)
22. Wickens, C.D., Hollands, J.G., Banbury, S., et al.: Engineering Psychology and Human Performance, 4th edn. Pearson, Upper Saddle River (2012)
23. Shorrock, S., Kirwan, B.: Development and application of a human error identification tool for air traffic control. Appl. Ergon. **33**(4), 319–336 (2002)
24. Kirwan, B.: Human error identification techniques for risk assessment of high risk systems-Part 2: towards a frame work approach. Appl. Ergon. **29**, 299–318 (1998)
25. Kirwan, B.: A Guide To Practical Human Reliability Assessment. CRC Press, Taylor & Francis Group, Boca Raton (1994)
26. Zhang, J.J., Paterl, V.L., Johnson, T.R., et al.: A cognitive taxonomy of medical errors. J. Biomed. Inform. **37**, 193–204 (2004)

The Effect of the Urban Bus Stop on Human-Machine Characteristics in Public Transport

Fengyuan Wang[1(✉)], Xiaoting Chen[1], Lu Wang[2], Shiyuan Yan[1],
and Fuquan Pan[1]

[1] Qingdao University of Technology, Qingdao 266520, China
122290913@qq.com
[2] Tengyuan Design Institute Co. Ltd., Qingdao 266000, China

Abstract. This paper designed and carried a practical vehicle test through using FaceLAB eye tracker, vehicle dynamic performance test system to collect and analyze human-machine characteristics and the variation regulation of natural condition and working conditions were compared. The results showed significant differences. As for perceived characteristics, the average gaze of right and left rear view mirrors in real-car conditions was much more than that in the natural state. As for manipulation characteristics, the natural state of vehicle in the outbound control stage was 0.2 less than in the natural state and the average steering wheel angle was 0.1° less than in the natural state, and the average vehicle speed was 1.9 m/s lower than that in the natural state, and the acceleration was 0.31 m/s^2 lower than that in the natural state. For the fluctuation range of each parameter, the real-vehicle conditions were obviously greater than the natural state.

Keywords: Urban bus · Traffic environment · Station environment setting · Bus driver · Human-Machine characteristics

1 Introduction

With the development of urbanization and motorization, the role of urban public transport in the transportation system has been continuously improved, and the safety of travel seems to be more and more important. According to statistics, 90% to 95% of all traffic accidents are related to driver factors [1], and human-machine characteristics of public transport have also been concerned.

Dario D. Salvucci used the simulation test to study the characteristics of the driver in the process of lane change. The study found that the driver would slightly brake before the lane change, and the lane change process would accelerate until the lane change was completed. The driver's visual attention focused on the current driving lanes, and focused on the target lane and exterior mirrors very little. In the lane change intention, the driver began to increase the attention of the target lane, and began to pay attention to the rear view mirror [2]; Based on the driving simulator, the relationship between steering wheel steering characteristics and sensory characteristics was studied.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_32

Then, the standard deviation of steering wheel angular speed, the standard deviation of steering wheel angle, the steering angle, the steering wheel angle variation parameter and zero speed percent were chosen as the five parameters of the sensing parameters characteristics, and finally the perceptual characteristic detection model was established [3]. Bus operation was the results of driver's perception, decision-making, manipulation and mutual cooperation [4]. At present, domestic and foreign scholars usually studied from a single visual characteristics or manipulation of research, so the scope of study was small.

In this paper, the human-machine characteristic data of public transport was analyzed by theoretical analysis and statistical analysis. The change rule of human-machine characteristic parameters between natural state and real vehicle condition was compared to provide guidance for driving safety.

2 Natural Vehicle Test Design

2.1 Test Platform

The natural road test platform included the test vehicle and test instrument. Bus No. 3 was selected as our test vehicle, and the equipped car was LCK6910G-3, as shown in Fig. 1. Liujiadao-Yumingzui in Huangdao district of Qingdao was chose as

Fig. 1. Experimental car

Fig. 2. Test road

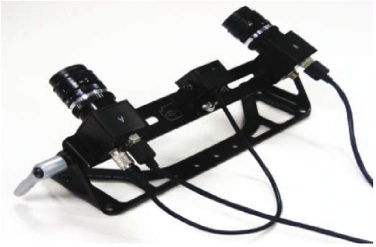


Fig. 3. FaceLAB eye tracker

Fig. 4. INDAS-5000 vehicle test equipment

Table 1. Driver sample

	Standard sample size	Deviation tolerance	Sample size
Optimum value	7	2	48
Minimum	10	5	16

the test line, whose length was about 4 km. There were 8 bus stations in the line and the distance between two stations was about 500 m, as shown in Fig. 2.

The FaceLAB eye tracker, driving behavior acquisition system and vehicle dynamic performance test system was selected to acquire the bus driver's perceptual characteristics, handling characteristics and bus operating characterization parameters in the bus stop process, as shown in Figs. 3 and 4.

2.2 The Experimental Staff

This paper studied on the effect of bus stops on the driver's human-computer characteristics, which focused on the driver's perceptual characteristics, handling characteristics and decision-making. But the degree of the three factors were related in this research. In order to eliminate the influence, the drivers were selected randomly according to the results of the bus company's safety and skills evaluation and they were selected by the relevant samples to introduce the effect of driving skills difference on the test results. In the test, the driver of the sample was selected to carry out the test according to the bus company's existing operation schedule, so as to realize the random distribution of samples. The number of sampled driver samples was shown in Table 1.

2.3 The Test Process

According to the test purpose and precautions, the test was designed. By the way, the time recording and synchronization test work should be taken seriously to ensure the accuracy of data collection. The test flow chart was shown in Fig. 5.

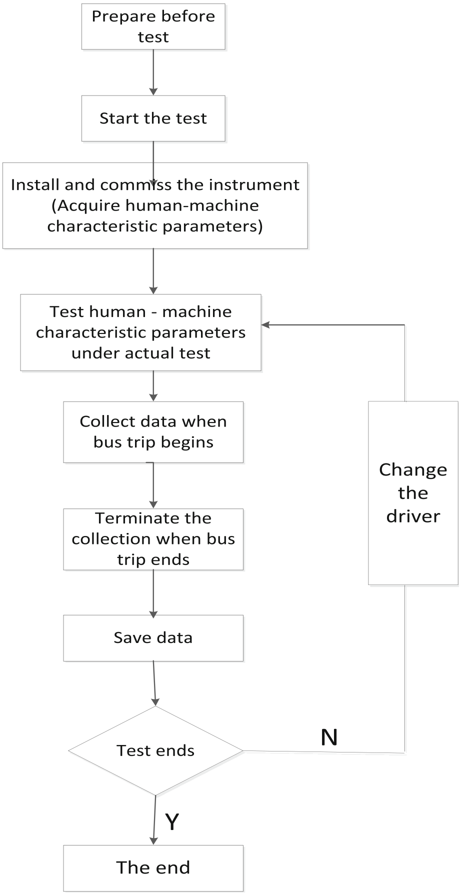


Fig. 5. The real vehicle test flow chart

2.4 Position Coordinate Method

In the natural operation of the bus, the driver mainly judged and made decisions according to the speed and position of buses, the speed of the front and rear cars and the information of road traffic environment. So this paper used the position coordinate method to analyze the change rules of human-machine characteristics of public transport. Since the average distance between the selected sections was 500 m, this paper set the coordinates of stop at the site as 250 m, which was not affected by the intersection.

As shown in Fig. 6, the driver first turned on the turn signal to alert other vehicles, and then selected the appropriate time to turn the steering wheel out of the station, so this article would be divided into the routine process of intent and manipulation of two stages; under the location coordinates, the point which the driver turned on the lights and turned the steering wheel were coincided. In order to facilitate the data analysis, the outbound process was not divided into the intent stage; at the same time, the free-running stage was used to express the region outside the station since it was not the study focus.

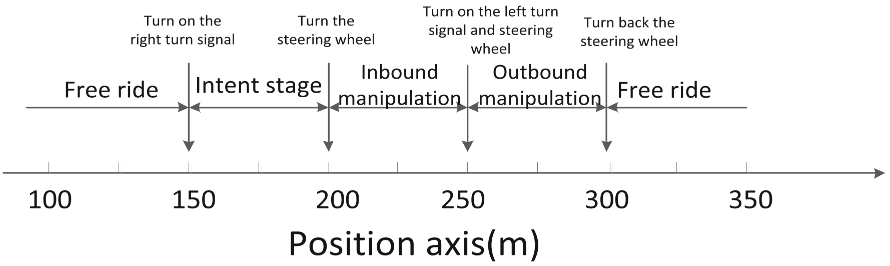


Fig. 6. Location coordinate

3 Analysis of Human - Machine Characteristics of Public Transport by Real Vehicle Test

Road traffic environment carried all pedestrians, non-motor vehicles, motor vehicles through road layout, signal control, traffic rules in a limited space, so to achieve the aim that all the traffic individuals could get orderly movement. When the bus was in the process of entering and leaving the bus station, it would be subject to the station passengers and docked vehicles, vehicles outside the platform and other complex changes in the state of the factors, such as shown in Fig. 7.

During the operation of bus, the driver was expected to complete the driving task in a safe manner. The safety could be achieved by the best matching quality between the driver and the orderly structure of road traffic environment, which was the safe operation of the bus. The intrinsic determinants were the driver's perception, decision-making and action, and the correlation between them. The exterior manifestation was the driver's perceptual characteristics, handling characteristics and bus operating characteristics.

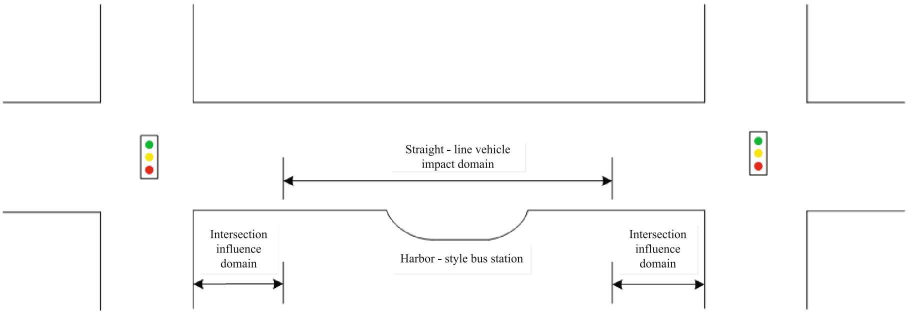


Fig. 7. Traffic influence schematic diagram

3.1 Driver's Perceptual Characteristics Analysis

Perceptual characteristics referred to the changes of sensory organs such as vision, hearing and touch during the operation of the bus. The information acquired by visual system accounted for about 90% of the total amount of information [5]. This paper selected and studied the fixation of left and right rear view mirrors as the perception characterization parameter.

Comparative Analysis of Perception Characteristics Basic Regulation
Fixation of Left Rear View Mirrors. As could be seen from Fig. 8, there was no significant difference in the fixations number of left rear view mirrors in the free ride and intent stage, and the average number of fixations was less than 1.5. Since the platform was arranged on the right side of vehicle, when the bus was in the free ride or in the intent stage, the left side of vehicle impact on driver's was small; in the outbound control stage, the natural state of the average number of fixation in the 0.6–2.0 times, and the number of real vehicle conditions was 0.4–3.0 times, which was obviously more than the former. It could be seen that in the outbound stage, social vehicles with high speed would crash in. In order to improve driving safety, drivers needed to pay more attention to the left rear view mirror in the traffic changes and took preventive measures.

Fixation of Right Rear View Mirrors. It could be seen from Fig. 9 that the fixation of the right rear view mirror in real vehicle conditions was significantly more than the number of natural state. In the intent stage, the average number of right rear view mirror fixation was 1.5–2.0 times and 2.3–2.8 times in the real vehicle condition; the average number of fixation in the natural state was 1.3–2.0 times; in the handling stage, the fixation number of real vehicle was 1.0–2.8 times and there were significant differences between this two conditions. As for the reason, in the inbound stage, the driver needed to pay attention to the right and rear right of the road traffic environment changes about other vehicles and pedestrian in order to get into the bus station safely.

Comparative Analysis of Perception Characteristics Distribution. In order to test whether there were significant differences in the number of fixation of the rear view mirrors on both sides of the vehicle in the natural state and the real vehicle condition,

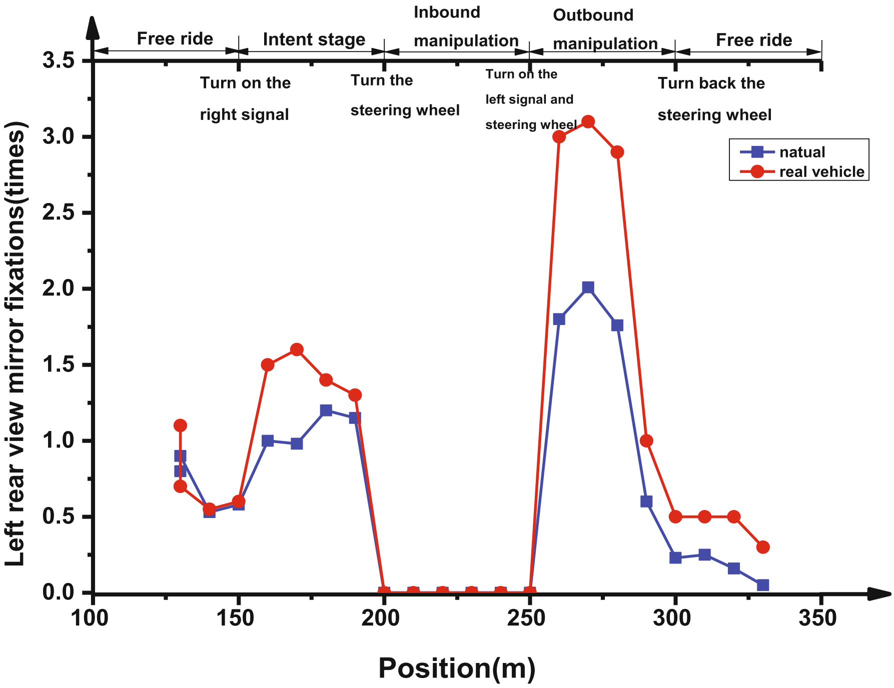


Fig. 8. Rear view left mirror gaze number variation

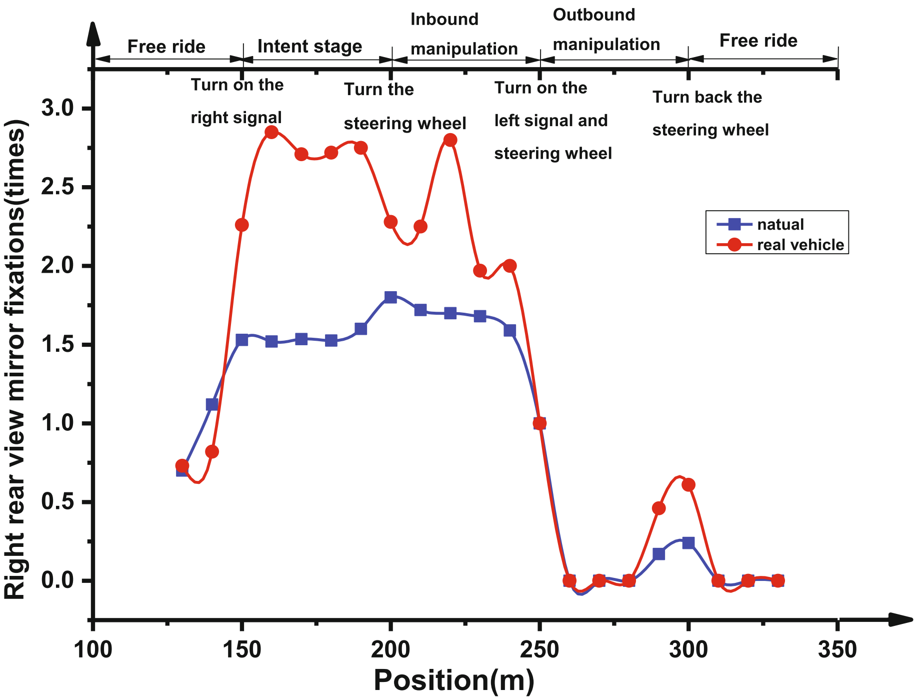
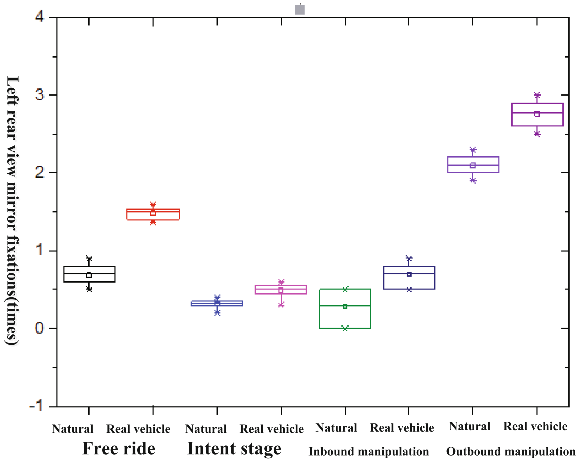
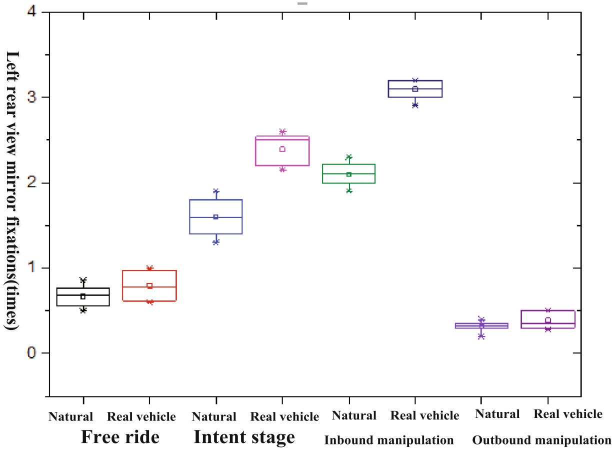


Fig. 9. Rear view right mirror gaze number variation

the distribution regulation of the number of fixation of the rear view mirrors on both sides was established according to the real vehicle test sample data, as shown in Fig. 10. It could be seen that the average number of fixations of the left rear view mirror in real vehicle conditions was more than that of the natural state, and the difference was especially obvious in the outbound stage. The difference of the average fixation times was 0.7 times. The mean number of fixations in the rear view mirror was similar to that in the rear view mirror. There were significant differences in the mean number of fixations between the two conditions in the intent and operation stage, which indicates that there was a significant difference in the perception characteristics of driver between the natural state and the real vehicle condition.



(a) Left rear view mirror



(b) Right rear view mirror

Fig. 10. The average of rear view mirror gaze number distribution

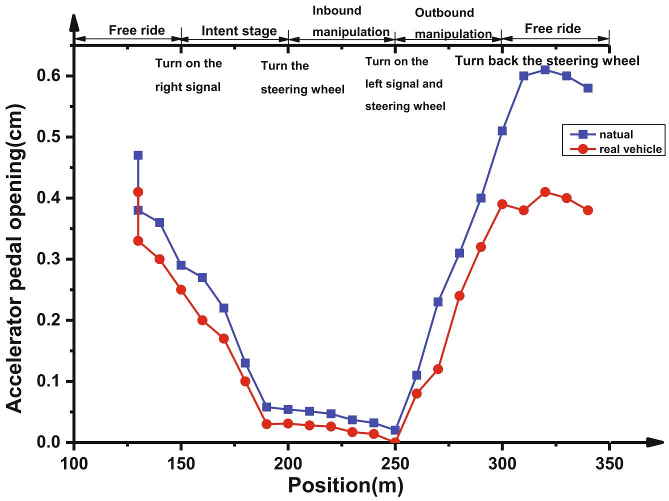


Fig. 11. Accelerator pedal opening variety regulation

3.2 Driver's Handling Characteristics Analysis

Handling characteristic was that drivers make their own decisions according to the steering wheel, accelerator pedal, brake pedal rotation angles and opening variation, then leaded the steering, acceleration and braking respectively into a continuous variation. This paper chose steering wheel and accelerator pedal opening angle as the characterization parameters to study the change regulation of handling characteristics.

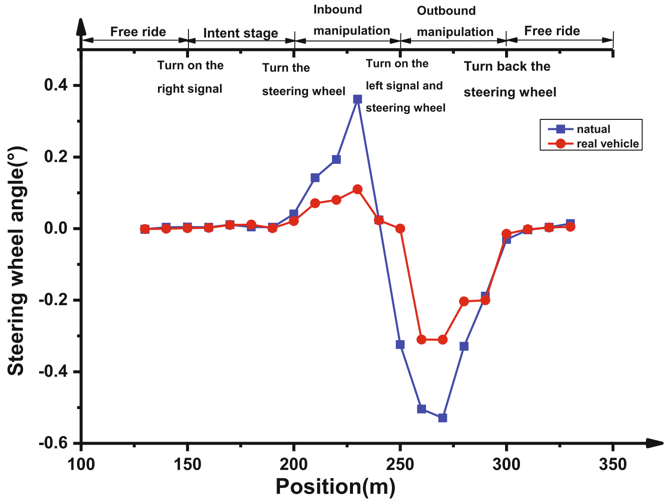
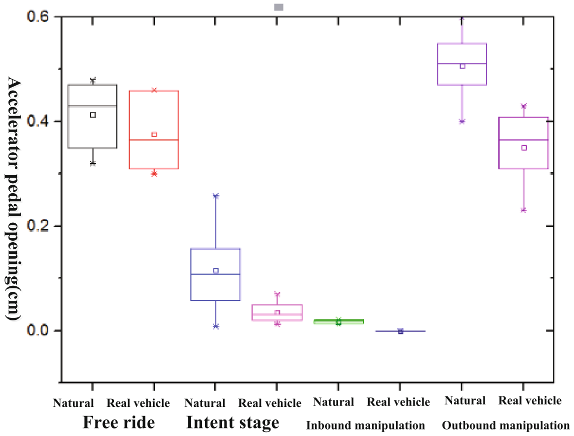


Fig. 12. Steering wheel angle variety regulation

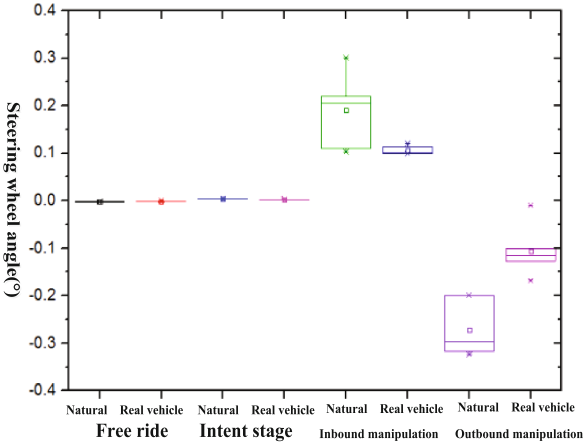
Comparative Analysis of Handling Characteristics Basic Regulation

The Opening of Accelerator Pedal. For the opening variation of accelerator pedal, as shown in Fig. 11, overall, the various stages in and out of the bus station, in fact, the average opening of accelerator pedal was smaller than that in natural state. This was because that in real vehicle conditions, driver would see platform layout, traffic flow, pedestrian and other factors, and their vigilance would calm down significantly. In terms of vehicle handling, drivers tended to touch the accelerator pedal for a relatively high frequency, to make the bus driving in a stable and controllable state, at the same time, to ensure traffic safety and to improve the efficiency of services.

The Angle of Steering Wheel. The variation of the number of right rear view mirror could be seeing from Fig. 12. Overall, the various stages in and out of the bus station, the average angle of the steering wheel under real vehicle condition were less than that



(a)Average accelerator pedal opening



(b) Average steering wheel angle

Fig. 13. Characteristics of driver manipulation distribution

in the natural state. Particularly, in the progress of in and out of the bus station, the difference in average angle of the steering wheel could be up to 0.3°.

Comparative Analysis of Handling Characteristics Distribution. As shown in Fig. 13, it could be seen that average the accelerator pedal opening under real vehicle conditions was smaller than that in natural state, and there was particularly noticeable phase difference in outbound control stage, in which the difference could be up to 0.2.

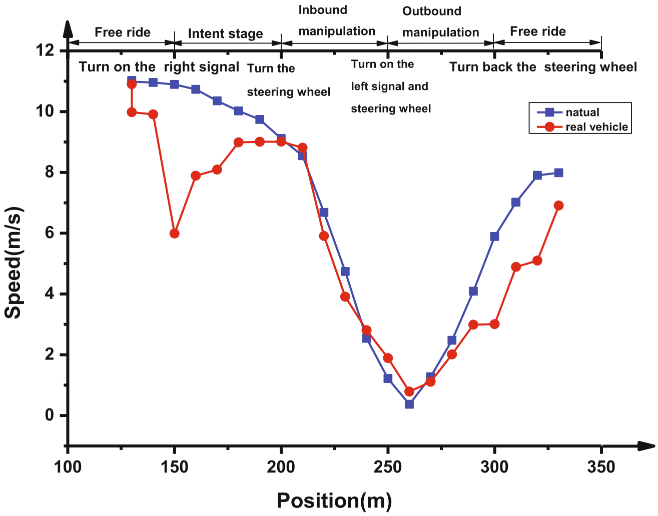


Fig. 14. Bus running speed variety regulation

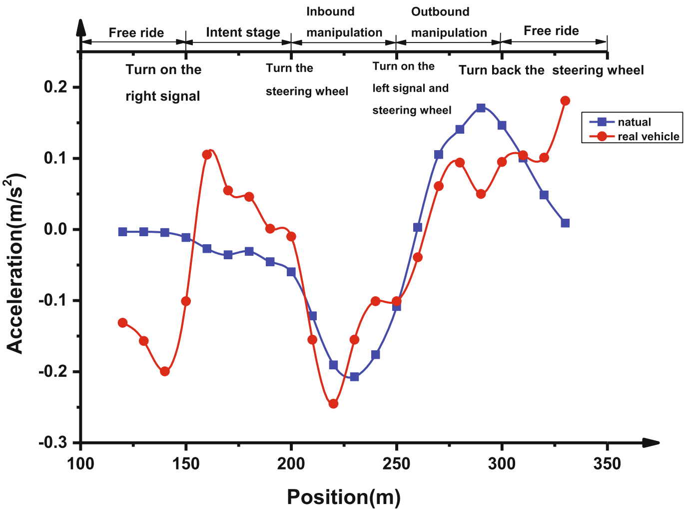


Fig. 15. Bus running acceleration variety regulation

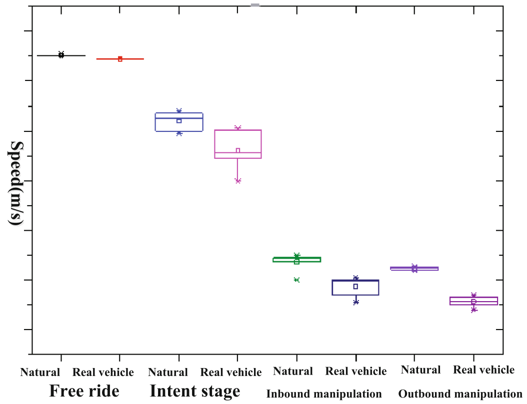
The steering wheel angle under real vehicle condition was also smaller than that in natural state, and there was noticeable phase differences in inbound and outbound operation stage, and the difference could be up to 0.1°.

3.3 Buses' Operating Characteristics

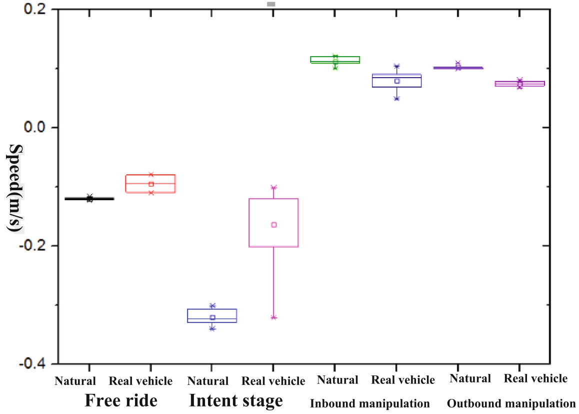
The buses operation characteristics pointed to the change regulations of bus under the driver's manipulation of speed, acceleration and so on [6]. This paper chose buses running speed and acceleration as the characteristic parameter to study the vehicle operation characteristics distribution.

Comparative Analysis of Operation Characteristics Basic Regulation

The Vehicle Speed. As shown in Fig. 14 the speed range under real vehicle conditions was larger than that in the natural state, and the speed fluctuation under real vehicle condition was significantly greater than in the natural state, especially in the stop



(a) Average speed



(b) Average acceleration

Fig. 16. Bus running distribution

intention stage. Under the natural state, the bus speed was always in steady decline, on the contrary, under the real vehicle condition, the bus speed changed from 10 m/s to 5.8 m/s. Compared with trial video, it could be seen that in that point, a social vehicles with high speed suddenly crashed into the space near the platform, which happened to correspond to the speed of mutation.

The Vehicle Acceleration. As shown in Fig. 15, the acceleration range under natural state was greater than that in the real vehicle condition. In stages, the acceleration fluctuation under real vehicle condition was significantly greater than that in the natural state, especially in the controlling stage. Under the natural state, the acceleration changed smoothly, on the contrary, under the real vehicle condition, when other vehicles with fast speed drove near platform, the acceleration rapidly changed to −0.29, with a large volatility.

Comparative Analysis of Operation Characteristics Distribution. As shown in Fig. 16. It could be seen that the average speed under real vehicle condition was less than that in the natural state, during the period of the outbound manipulation, the difference of speed was up to 1.9 m/s. The average acceleration under real vehicle condition was also less than that in the natural state, and the difference of acceleration in the outbound stage was up to 0.23 m/s2. It showed that there existed obvious difference in the vehicle operation characteristics between the natural state and real vehicle condition.

4 Conclusions

1. In terms of perceptual characteristics, the driver's average gaze on the left and right rearview mirrors under the real vehicle condition was more than that in the natural state; as for handling characteristics, the average opening angle of the accelerator pedal and the steering wheel under real vehicle condition were lower than those in the natural state; in terms of the bus operating characteristics, the average speed and average acceleration of vehicle were lower than those in the natural state.
2. Bus company can improve the relevant regulations during the driver training and the operation of the vehicle. For example, when the bus is in and out of the bus station normally, the driver is required to inspect the left rear view mirror more than 2.5 times; as for the accelerator pedal tread and steering wheel angle controlling, drivers should handle high-frequency and flexible. The accelerator pedal maximum opening value shall not exceed 0.7 cm and the absolute value of the steering wheel angle is not greater than 0.6°.

Acknowledgments. The research is sponsored by Shandong Province Natural Science Fund (ZR2016EEM14).

References

1. Treat, J.R.: A study of crash factors involved in traffic accidents. HSRI Res. Rev. **10**(6), 1–35 (1980)
2. Salvucci, D.D., Liu, A.: The time course of a lane change: driver control and eye-movement behavior. Transp. Res. Part F **5**, 123–132 (2002)
3. Lisheng, J., Kelong, L., Qingning, N.: A new method for detecting driver fatigue using steering performance. Traffic Inf. Secur., 5 (2014)
4. Solomn D.: Accidents on Main Rural Highways Related to Speed, Drivers, and Vehicle, pp. 11–17. Bureau of Public Roads, Washington (1964)
5. Fengyuan, W., Jieyu, F.: Study on the visual physiology of driver safety in urban traffic condition. J. Qingdao Univ. Technol. **7**, (2012)
6. Vadeby, A.M.: Modeling of relative collision safety including driver characteristics. Accid. Anal. Prev. **36**(5), 909–917 (2004)

Operating Events in Chinese Nuclear Power Plants: A Preliminary Analysis

Lixin Zheng[1], Feng Jiao[1], Peng Liu[2](✉), Jiejuan Tong[3], and Zhizhong Li[4]

[1] Nuclear and Radiation Safety Center,
Beijing 100082, People's Republic of China
{zhenglixin, jiaofeng}@chinansc.cn
[2] Department of Industrial Engineering, Tianjin University, Tianjin 300072,
People's Republic of China
pengliu@tju.edu.cn
[3] Institute of Nuclear and New Energy Technology, Tsinghua University,
Beijing 100084, People's Republic of China
tongjj@tsinghua.edu.cn
[4] Department of Industrial Engineering, Tsinghua University, Beijing 100084,
People's Republic of China
zzli@tsinghua.edu.cn

Abstract. Operating event reporting is important for continuously improving the safety level of nuclear power plants (NPPs). This paper presents the preliminary analysis results based on the data of the licensee operating events (LOEs) occurred in Chinese NPPs during 1991–2015. A total of 805 LOEs were collected. Trends of the number of LOEs and reactor units are presented. It is found that the number of LOEs per unit declined gradually. There were 50% of all LOEs related to human factors and 48% related to equipment factors. More than 50% of human-related LOEs were associated with Personal Work Practices and Written Procedures and Documents. The differences in operating events between conventional and advanced NPPs were analyzed. These results and lessons learned from these LOEs will be helpful for operator error prevention and human reliability analysis in NPPs.

Keywords: Human factors · Licensee operating events · Operator error prevention · Experience feedback · Chinese NPPs

1 Introduction

For safety-critical systems, safety is highly dependent on operator performance. Humans are widely acknowledged to be a major cause of system incidents and accidents [1–3]. It is widely accepted that the contribution of human error to incidents and accidents in safety-critical systems is approximately 30%–80%.

We focus on incidents and accidents in nuclear power plants (NNPs). Analysing operating events benefits at least two human performance related areas. First, it can improve operator error prevention activities and increase system safety. The significant causal operator performance problems are identified in the analysis of operating events.

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_33

The operator error prevention program should place the high priority on them. Second, the findings from operating events analysis can be transferred into human reliability analysis (HRA). Causal analysis can shed light on human error mechanism. It can be used to identify the significant performance shaping factors (PSFs) which are used to predict and explain human error probability.

There are many factors contributing to the safe operation of NPPs. Although the NPP industry places the safety as the first priority in the operation of NPPs, there are still incidents emerged causing unscheduled outages or degradation of systems. Since the 1979 Three Mile Island accident and the 1986 Chernobyl accident, it is increasingly recognized that human performance is a major contributor to risk. The database of operating events provides the opportunity to locate the contribution of human performance problems and their major categories. Bento [4] reviewed 197 scrams and 1759 licensee event reports (LERs) during 1983–1988 in Swedish NPPs and observed that 38% of the scrams and 27% of the LERs were caused by human performance problems. Among them, “work organization”, “work place ergonomics”, “procedure not followed”, “training” and “human variability” were reported to be most significant causal categories. Subramanya [5] observed that human errors caused 15% incidents at India NPPs. Hirotsu et al. [6] investigated the contribution of human errors in design, manufacture, installation, maintenance and operation to incidents occurring at Japanese NPPs between 1970 and 1996, and the ratio of human error cases to total incidents was around 20%. Among them, about 60% of human error incidents occurred during maintenance and about 20% during operation. The Korean regulatory organization for nuclear and radiological systems, Korea Institute of Nuclear Safety (KINS), provides a list of major operating events (e.g., unplanned reactor trips and unplanned initiations of safety systems) occurring in Korean NPPs. Their information is shown on a public website, i.e. the Operational Performance Information System (OPIS) (http://opis.kins.re.kr) [7]. Nuclear events are caused by human errors, mechanical defects, electrical defects, instrumental and control (I&C) defects, and external effects. According to this web-based database, about 18.0% of the Korean nuclear events were caused by human errors during 1978–2017. Defects in I&C and mechanical defects were the two major causes, which contributed to 29.4% and 26.4% nuclear events, respectively.

Up to March 2017, mainland China has 36 nuclear power reactors in operation, 21 under construction [8]. The main impetus for developing nuclear power in China is to against air pollution from coal-fired plants. Several publications have reported the operating events information in Chinese NPPs. For example, Huang and Zhang [9] found that 39% events, including 24-hour events, licensee operating events (LOEs), and internal operating events (IOEs), were classified as human factor events in the first three years commercial operation of Daya Bay NPP during 1994–1996. The main root causes for human error events were operator omission, procedure deficiency, procedure not followed, lack of training, communication failures, and work management inadequacy. Liu et al. [10] applied the accident sequence precursor (ASP) technique to analyse LOEs and IOEs occurring at Daya Bay (during 2000–2009) and LingAo NPPs (during 2003–2009), and suggested that the two major contributors were equipment defects/faults (61%) and human factors (39%). One recent study by Zou et al. [11] reporting human factor events in LingAo II NPP found that the significant types of human errors included work practice, file management, and communication. The summary report of operating events at Chinese NPPs has been issued in 2012 [12].

Operating event reporting and experience feedback are critical aspects of safe operation of NPPs. It is important to communicate these information in the worldwide. The current statistical information from Chinese NPPs and experience feedback are published in domestic publications, increasing the difficulty for sharing the information to the international counterparts. The importance of international operating experience feedback is increasingly recognized. This article presents the overall information of the human contribution to and lessons learned from operating events occurred in Chinese NPPs.

2 Methodology

2.1 Data Sources

Strict nuclear event reporting system has been built in China. In 1995, the Chinese regulatory agency for nuclear and radiological systems, National Nuclear Safety Administration (NNSA), issued the operating event reporting system for nuclear facilities [13] and determined the format and content for NPP owners to submit reports of nuclear events to NNSA. In 2004, NNSA issued the operational safety requirement for nuclear power plant and specified the requirement of operation experience feedback [14]. In China, operation experience feedback is an important measure for maintaining NPP operation safety and improving the quality of surveillance for the regulatory agency.

The reported LOEs should be coded according to the standard of “Event coding of nuclear power plant” [15]. The coding system consists of 8 categories, including consequences, systems, components, station status, station activity, group, direct cause, and root cause. Event consequences include, for example, degraded station operating conditions, station transient, and equipment damage. The direct cause codes are classified into 9 main categories. The human factors direct causes include slip/lapse, mistake, violation, sabotage, and others [3]. The root causes (i.e., causal factors) are coded at three levels. Level 1 has 3 major group, i.e. human performance related, management related, and equipment related. Each group is further subdivided in several sub-groups (Level 2). Each sub-group is also further subdivided in specific factors/causes (Level 3), in order to be more precise in identifying the causes. Usually, it is difficult for incident analysts to attribute operating events to the Level 3 of root causes.

The LOE reports ($N = 805$) from Chinese NPPs during 1991–2015 were summarized as the data sources.

2.2 Data Analysis

The number of LOEs per unit during 1991–2015 was calculated for trend analysis. The Level 1 root causes (i.e., human related, equipment related, and management related) were classified and their percentages in the LOEs were calculated. The Level 2 root causes in the human-related categories were further analyzed.

3 Results and Discussion

3.1 Trend Analysis

During 1991–2015, a total of 805 LOEs were reported. Figure 1 shows the change of the number of LOEs and NPP units. In 1993, the two units of Daya Bay NPP were connected to the grid and started their commercial operation. Thus, it had a significant increase of LOEs in 1993. There also had a significant increase of LOEs in 2002, because of the startup of Ling Ao Phase I and Qinshan Phase III units. These two sharp points reflect the early phenomenon. The number of NPP units in operation was gradually increased.

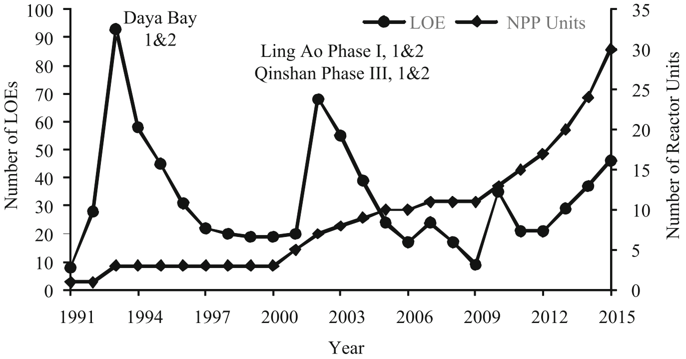


Fig. 1. Number of LOEs per year and reactors per year during 1991–2015.

Figure 2 presents the number of LOEs per unit. Chinese nuclear industry was at the initial stage during the first 10 years (1991 to 2000), and thus it is not surprising to see a high number of LOEs per unit during the first 10 years. The mean number of LOEs per reactor/year in the first 5 years was 21.1 and in the second 5 years it dropped to 7.4. During the second 10 years, it was at the development state of Chinese nuclear industry. Since then, the mean number of LOEs per reactor/year was below 5.0. For the last five years (2011–2015), it was reduced to 1.5.

3.2 Root Cause Analysis

The root causes of the 805 LOEs were identified. Figure 3 presents the contribution of human-related, equipment-related, and management-related root causes to LOEs. There were 50% LOEs relevant to human factors and 48% relevant to equipment factors for all NPPs, see Fig. 3(1). The LOEs from conventional and advanced NPPs were separated. As shown in Fig. 3(2), 52% and 46% of LOEs in conventional NPPs were related to human factors and equipment factors, respectively. For advanced NPPs, 44% and 52% of LOEs were related to human factors and equipment factors, respectively. The percentage of human-related LOEs in advanced NPPs (44%) was lower than that in

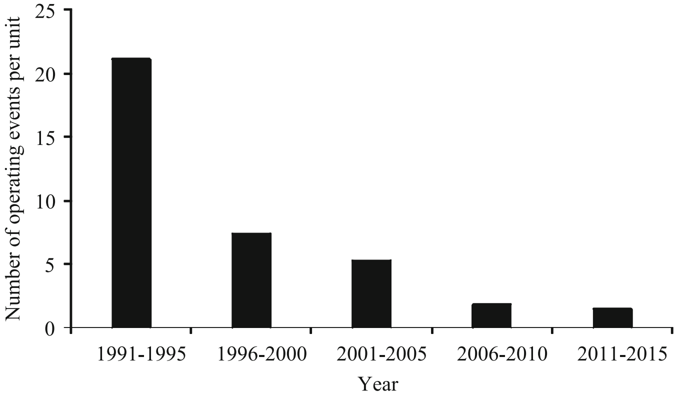


Fig. 2. Number of operating events per reactor/year during 1991–2015.

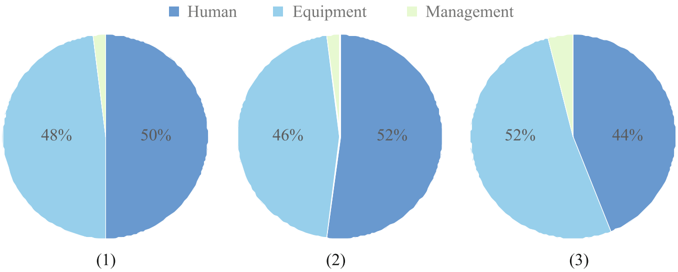


Fig. 3. Number of operating events per reactor/year during 1991–2015: (1) all NPPs, (2) conventional NPPs, and (3) advanced NPPs.

conventional NPPs (52%). Note that it does not mean that the likelihood of operation errors in advanced NPPs is lower than that in conventional NPPs, because that human-related LOEs were not just caused by operation errors and the likelihood of operation errors and other errors were not directly compared between conventional and advanced NPPs.

The human causal factors in all NPPs, conventional NPPs, and advanced NPPs were further subdivided into specific categories. Their relative importance was ranked in terms of the percentage of LOEs attributed to these categories (see Table 1). The ranked top 2 categories of human casual factors in all NPPs were Personal Work Practices and Written Procedures and Documents. They covered more than 50% of human related LOEs. There were 36% human-related LOEs belonging to the category of Personal Work Practices (e.g., independent checking not used, or improper equipment used) and 27% belonging to the category of Written Procedures and Documents (e.g., document unavailable or incorrect). One difference between conventional and advanced NPPs was the ranked third category. In conventional NPPs, it was Verbal Communication (relevant to 9% of human-related LOEs); however, it was Personal Factors (relevant to 14% of human-related LOEs).

Table 1. Human-related root causes in all NPPs, conventional NPPs, and advanced NPPs

Root causes	All NPPs	Conventional NPPs	Advanced NPPs
Personal work practices	36%	37%	34%
Written procedures and documents	27%	27%	27%
Personal factors	11%	9%	14%
Verbal communication	9%	11%	6%
Qualification and training	6%	7%	3%
Supervisory methods	4%	3%	7%
Work organisation	3%	2%	5%
Man-machine interface	3%	2%	4%
Work scheduling	1%	2%	0%

Below are several findings related to human causal factors:

(1) The number of human-related LOEs is gradually decreased. However, human is still a major contributor to LOEs compared with other factors. About half of operating events are related to human factors.
(2) The main human-related causal factors were poor work practices, not following required procedures, not verifying operation content, not familiar with system working, not adequately monitoring, lack of expertise, inadequate communication, and lack of communication between operators.
(3) The latent failures of systems and components, caused by the human-related causal factors in maintenance and test at the initial operational stage, finally triggered active system failures.

3.3 Experience and Lessons

As mentioned before, experience feedback is critically important for continuously improving NPP safety. Several lessons are learned from investigating, summarizing, classifying, and analyzing the LOEs occurred in Chinese NPPs.

(1) Lack of risk analysis

- Risk analysis of the operations unspecified by procedures should be conducted and recorded as written instruction.
- The potential human errors and their worst consequences for any tasks and operations should be considered as hazards. Specific preparedness should be made against these hazards.
- Several human error prevention tools are strongly required, including “questioning attitude” and “pre-work meeting”.

(2) Lack of work preparation

- Preparation plans in maintenance, test, and examination should be completed.
- Human, tools, equipment, procedures should be prepared, checked, and verified in advance.

- Operators should use the "pre-work meeting" as a human error prevention tool and verify the work preparation.
- Operators should use the "work handover" as a human error prevention tool.

(3) Lack of communication and coordination

- Teamwork training and skills should be enhanced.
- Communication and verification across different sites and departments should be performed before the work.
- Equipment status, operation position, and operation steps and sequence should be communicated in the "pre-work meeting".
- Operators should use the "three-way communication" as a human error prevention tool in any situations including face-to-face communication and wireless communication with interphone, telephone, etc.
- Coordination should be conducted in critical tasks and missions. Specific work groups should be built to perform the coordination work.

(4) Unqualified operation by contractors

- Contractor training and awareness should be improved.
- Directors for any missions should supervise the contractors and control the critical points in the work processes.

4 Conclusions

The importance of international operating experience feedback is increasingly recognized. This paper shares the information of operating events at Chinese NPPs during 1991–2015. A total of 805 LOEs were collected and analyzed. The number of LOEs per unit declined gradually. As found in the existing literature, human and equipment were the two contributors to the occurrence of LOEs in Chinese nuclear industry. The percentage of human-related LOEs in advanced NPPs (44%) was lower than that in conventional NPPs (52%). The ranked top 2 human causal factors in all NPPs were Personal Work Practices and Written Procedures and Documents. The lessons were given about lack of risk analysis, lack of work preparation, lack of communication and coordination, and unqualified operation.

Acknowledgments. This study was supported by the National Natural Science Foundation of China (Project no. 71371104, 71601139, and 71671102).

References

1. Rasmussen, J.: Human error and the problem of causality in analysis of accidents. Philos. T. Roy. Soc. B. **327**, 449–462 (1990)
2. Kirwan, B.: The validation of three human reliability quantification techniques - THERP, HEART and JHEDI: Part I - technique descriptions and validation issues. Appl. Ergon. **27**, 359–373 (1996)

3. Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)
4. Bento, J.P.: Collection, analysis and classification of human performance problems at the swedish nuclear power plants. In: IAEA (ed.) Human Error Classification and Data Collection: Report of a Technical Committee Meeting, pp. 83–94. IAEA, Vienna (1989)
5. Subramanya, K.C.: Human Errors - human caused, environment caused. In: IAEA (ed.) Human Error Classification and Data Collection: Report of a Technical Committee Meeting, pp. 61–70. IAEA, Vienna (1989)
6. Hirotsu, Y., Suzuki, K., Kojima, M., Takano, K.: Multivariate analysis of human error incidents occurring at nuclear power plants: several occurrence patterns of observed human errors. Cogn. Technol. Work **3**, 82–91 (2001)
7. Korea Institute of Nuclear Safety (KINS). http://opis.kins.re.kr
8. World Nuclear Association (WNA). http://www.world-nuclear.org/information-library/country-profiles/countries-a-f/china-nuclear-power.aspx
9. Huang, W., Zhang, L.: Cause analysis and preventives for human error events in Daya Bay NPP. Nucl. Power Eng. **19**, 64–68 (1998). (in Chinese)
10. Liu, P., Xing, Y., Xi, H.: Accident sequence precursor analysis of Daya Bay and Ling'ao nuclear power plants. Atom. Engery Sci. Tec. **27**, 1857–1861 (2013). (in Chinese)
11. Zou, Y., Zhang, L., Dai, L., Li, P., Qing, T.: Human reliability analysis for digitized nuclear power plants: case study on LingAo II NPP. Nucl. Eng. Tec. **49**, 335–341 (2017)
12. Wang, Z., Li, Z.: Summary Reports of Operating Events in Chinese Nuclear Power Plants. China Environmental Science Press, Beijing (2012). (in Chinese)
13. National Nuclear Safety Administration (NNSA): Reporting System of Nuclear Power Plants Utilities. NNSA, Beijing, China (1995). (in Chinese)
14. National Nuclear Safety Administration (NNSA): Operation Safety Requirements of Nuclear Power Plant. NNSA, Beijing, China (2004). (in Chinese)
15. National Energy Administration (NEA): Event Coding of Nuclear Power Plant. NNSA, Beijing, China (2010). (in Chinese)

Description of Diagnosis Process: A Review of Existing Measures and a New Approach

Xi Lyu(✉) and Zhizhong Li

Department of Industrial Engineering, Tsinghua University,
100084 Beijing, China
lvxl2@mails.tsinghua.edu.cn, zzli@tsinghua.edu.cn

Abstract. Human diagnosis receives increasing concern in many industries. Researchers need to properly describe the process of diagnosis before trying to analyze or improve human diagnostic performance. This study reviews the existing ways to describe the process of diagnosis and summarizes them in terms of three sub-processes of diagnosis, i.e. hypothesis generation, information seeking, and information integration. Then a new approach is proposed, drawing ideas from information entropy and fuzzy signal detection theory. The proposed approach serves to describe information seeking and information integration with more precision.

Keywords: Human factors · Diagnosis · Information seeking · Information integration

1 Introduction

Recent advances in human-technology systems have been introducing growing automation to various industries, such as automatic driving, digitalized main control rooms in nuclear power plants, assistant applications on smartphones, etc. As a result, the role of humans in these systems have been shifting from accomplishing manual operations to performing kinds of cognitive activities [1–3]. As long as automation works fine, the systems get jobs done with lessened human effort. Nevertheless, once automation malfunctions plus failing to fix itself, the systems have to rely upon humans to notice something is wrong, to diagnose what goes wrong and to bring things back to the normal. Diagnosis is the process of understanding, awareness, or assessment of the abnormal situation [4], and serves to setting the premise for correct responses, for example, initiating operating procedures in nuclear power plants [5].

Under various situations, individuals carry out diagnosis quite differently [6]. When the situations are familiar, experienced diagnosticians are able to match patterns of symptoms with those stored in their memory, hence arriving at conclusions by direct recognition [6–9]. When the problematic situations are within boundaries of expectation and well prepared for, people can perform diagnosis following the guidance of established rules, procedures, or checklists [6, 10, 11]. While confronted with novel situations, people have to go through an iterative hypothetic-deductive process until finally reaching a conclusion [6, 12, 13]. In this kind of circumstances, people first generate hypotheses about possible causes of abnormality from initially obtained

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_34

information. Then they seek more information, and integrate the information at hand to verify the hypotheses. If the verification is satisfactory, mission complete; otherwise, another iteration starts by seeking new information or generating new hypotheses. Here we can distinguish three sub-processes of diagnosis, namely, 'hypothesis generation', 'information seeking' (or cue acquisition), and 'information integration' (or cue interpretation, likelihood judgment) [7, 12–15]. Although the most identifiable in the hypothetic-deductive manner of diagnosis, these sub-processes prevail in other types of diagnosis as well. For example, in the pattern-recognition diagnosis, the 'pattern' is the result of information seeking, and 'recognition' itself is a way of information integration. In the procedure-guided diagnosis, rules, procedures, and checklists are themselves well-formatted ways to help people carry out these sub-processes. From this point of view, we can overcome the diversity and describe any process of diagnosis in a homogenous way, i.e. by describing its sub-processes separately and as a whole.

Diagnosis has long been intensively studied [4, 16, 17], but this subject is far from being resolved. Research on diagnosis covers a wide variety of contexts, for example, medicine [8, 13, 17], nuclear power plants [5, 18, 19], electronics [20–23], process plants [24, 25], manufacturing [26–28], aviation [2], ships [29], abstract games [30–32], etc. Even though there are arguments about some universal, context-free template of diagnosis across domains [13, 31], abilities and strategies to accomplish diagnostic tasks are quite context-specific [10], even for different tasks of the same field [7, 13, 15]. There are growing emphasis on diagnosis in many human-technology systems undergoing technologic progresses. When applying the conclusions of previous studies to new contexts, we need to examine the generalizability very carefully. Meanwhile, there remains some controversy concerning some topics about diagnosis, for example, whether principle knowledge helps diagnosis [11, 33], how is information seeking and information integration related [13, 34].

Since diagnosis takes various forms across many domains, and relevant literature is not conclusive, there is an enduring need for studying diagnosis under certain contexts. In order to investigate the factors of interest about diagnosis properly, we first have to describe the process of diagnosis properly. In the following sections, the authors will examine existing ways of describing the process of diagnosis from the literature and then propose a new approach. This study is supposed to help people better describe and understand the process of diagnosis in both existing and emerging contexts.

2 Existing Ways to Describe the Process of Diagnosis

Researchers used various variables to describe the process of diagnosis. In this study, the authors chose to categorize these variables as whether they describe hypothesis generation, information seeking, information integration, or the overall process of diagnosis. Hypothesis generation refers to the process that people come up with possible explanations of the confronting situation. Information seeking refers to the process that people search and acquire kinds of data relevant to the diagnostic task. Information integration refers to the process that people interpret the data at hand to make possibility judgments about the generated hypotheses. Variables are classified into each category if they represent the specific characteristics of the corresponding process.

This classification is in line with Duncan & Gray [35]'s and Henneman & Rouse [36]'s notions of 'differentiating between *product* measures and *process* measures of human performance'. The current study takes a step further by decomposing product measures into three collections, each corresponding to one of the three sub-processes. In many occasions, variables describing the sub-processes are necessary. At some times, overall measures are not sufficient to discriminate group differences, while at other times researchers need to know how the conclusions are obtained other than what conclusions are obtained. Correctness or accuracy ranks as the ultimate indicator of human performance on diagnostic tasks. However, less-than-optimal diagnosis can result in many ways, specifically, failing to generate correct hypotheses [13, 14, 24, 37, 38], failing to acquire sufficient information [13, 15, 24], or failing to appropriately integrating acquired information and hence to give adequate likelihood judgments [13, 37, 39–41]. A single measure of accuracy is not enough to describe diagnostic successes or failures thoroughly. Imagine a situation where one participant has made a correct diagnosis, but detailed analysis revealed that he/she did not acquire enough necessary information, and that this judgment was simply a rash guess. This kind of failures can stay unnoticed for long if researchers only care about final successes. Moreover, measures describing sub-processes can better pinpoint what part of diagnosis goes wrong, and hence design specific countermeasures. Supposing, researchers observe that their participants tend to fail on some diagnostic task. The researchers find that this is because the participants misinterpret the data, though provided sufficient necessary data. Then the researchers are supposed to design aid that serve to help their participants make better use of the data at hand, other than collect more data.

Besides this grand classification, the authors further grouped the variables into certain types based on what they essentially measure. Table 1 provides a summary of types of variable to describe the process of diagnosis used in 29 studies from the literature. A brief examination of Table 1 reveals that quite a proportion of studies focused on describing the overall process of diagnosis and the sub-process of information seeking, while the sub-processes of hypothesis generation and information integration received much less attention. This imbalance makes some sense in that researchers can observe the information seeking sub-process and the overall process of diagnosis from participants' extrinsic behaviors, while the sub-processes of hypothesis generation and information integration are inside the participants' minds, hence not cognizable unless explicitly measured.

Describing the Overall Process. Most studies used success rates (or counts) and completion times as 'flag' measures of diagnostic performance, like many other jobs. However, there are a few occasions when these two types of indicators are less adequate. Due to task difficulties close to extremes, small sample sizes, or even deliberate designs, diagnostic results turned out nearly all-correct [21, 28, 42] or nearly all-wrong [24]. Some studies reported tasks that end until success [22, 33] or even without strict success criterion [5, 20, 23]. For those diagnostic tasks where successes could result from trial-and-error, some researchers used 'first diagnosis successes' as an alternative [35, 42]. Others used performance ratings instead [20, 47], or focused on human errors [5, 22, 28]. Another problem is that the possibility of a speed-accuracy trade-off (SATO) cannot be eliminated beforehand [12, 16, 36]. Some studies in Table 1

Table 1. Summary of types of variables to describe the process of diagnosis.

Process of diagnosis	Type of variable	Literature that used this type of variables
Overall	Successes/Failures (counts, ratios, types)	[5, 11, 13, 19, 22, 24–30, 33, 43–46, 48]
	Time (lengths, counts)	[2, 5, 10, 11, 19, 27, 28, 30, 31, 33, 36, 43–48]
	Quality (ratings, scores)	[20, 33, 36, 47]
	Confidence	[2]
	Process tracing	[22–24, 26, 31]
	Cost	[36]
	Compositions (additions, products)	[35, 36]
Hypothesis generation	Number	[13, 19, 46]
	Evolution	[13]
	Time point	[13, 24]
	Correct hypothesis formulation	[13, 19, 24]
Information seeking	Actions in total (counts, ratios)	[2, 13, 19, 21, 25, 32, 42–48]
	Actions in terms of nature (counts, ratios)	[24, 31, 33, 42, 47]
	Actions in terms of timeline (counts, ratios)	[21, 24, 26, 42]
	Actions in terms of usefulness (counts, ratios)	[2, 13, 21, 31, 35, 36, 42, 43, 47]
	Strategies/Routines	[20, 22, 23, 30, 44]
	Thoroughness	[13, 35]
	Efficiency	[13, 43]
	Cost of actions	[36]
	Time (total, average)	[36]
Information integration	Number	[20]
	Accuracy	[13]
	Error types	[13]
	Time point	[24]
	Prematurity	[35, 36]

evaluated the changes of accuracy and time simultaneously, while others evaluated one while holding the other constant [22, 33]. However, if a study reports only one of the two, its conclusions should be examined carefully to see if the changes are at expense of the other. Except for success and time indicators, another popular type of variables is that resulting from process tracing. This type includes sequences [23, 24] or counts [26, 31] of coded events, global strategies (reasoning by recognition or by elimination) [22], etc. Although these variables are essentially process measures rather than product measures, they incorporated events from all three sub-processes. Therefore, they are classified as describing the overall process of diagnosis and serve to provide new angles of view.

Describing Hypothesis Generation. Hypothesis generation can be evaluated based on records about what hypothesis is generated at what time. Researchers usually obtained these records by analyzing participants' vocal protocols [13, 24], or sometimes by participants' voluntary reports [19]. Elstein et al. [13] demonstrated the evolution of hypotheses by counting the numbers of hypotheses active one-quarter, half-way, and at the end of diagnostic tasks. Although the total number of generated hypotheses were sometimes of interest [13, 19, 46], the most important issues are about the 'right answer' hypothesis: whether it is formulated [13, 24] and when it is formulated [19].

Describing Information Seeking. The simplest way to describe information seeking is to count participants' information seeking actions (checks, views, readings, etc.). However, this type of variables only provide a rough view of the full scene. We cannot even determine whether it is good to taking more information seeking actions or less, since more actions may give more information, but at the cost of spending more time, very much resembling a SATO. So many studies made further classifications. Information seeking actions can be different in nature, for example, executed on different parts of the system [31, 33, 42, 47]. If certain time-points are of great importance, such as the first diagnosis [42], identifying initial symptoms [26], or checking a critical component [21], relevant information seeking actions can be classified along the timeline. Duncan & Gray [34] classified information seeking actions as 'redundant' if the actions did not help in eliminating improbable faults, or 'extra' if the actions were taken after sufficient information for diagnosis had been acquired. In similar manners, many other studies [2, 31, 36, 43] classified information seeking actions according to their usefulness. Besides the efforts to classify information seeking actions one by one, some variables described information seeking more globally, such as the strategies or routines to regulate series of actions. Brinkman [30] distinguished between the tracing-back strategy and the hypothesis-and-test strategy and compared their proportions. Rasmussen & Jensen [23] analyzed diagnostic routines in terms of topographic search, functional search, search by evaluation of a fault and check measurements. Reed & Johnson [22] observed four local-level strategies: compare-and-conquer, heuristic path following, stateless analysis and endpoint analysis. At last, the performance of information seeking in total can be evaluated by thoroughness (percentage of cues acquired) [13, 35] and efficiency (the degree to which critical cues were acquired) [13, 43].

Describing Information Integration. In general, human judgment is evaluated against results from a 'standard' model, such as empirically established relations [49], or Bayesian posteriors [50]. For diagnostic tasks, the 'standard' models should give their results based on the data that actually acquired by humans. As an instance, Duncan & Gray [35] and Henneman & Rouse [36] recorded premature diagnoses that the participants made before they had sufficient information. On the other hand, researchers can directly compare the criteria, for example, the weights that the participants assigned to each cue with those assigned by subject-matter experts [13].

In summary, there exist many approaches to describe the overall process and each sub-process of diagnosis, though with different emphases. Researches generally chose the variables suitable for their study settings. However, it is far from concluding that an exhaustive list of variables relevant to diagnosis in the literature could suffice to describe any diagnostic process. Consider that we plan to evaluate information seeking actions by usefulness more in detail. We can make a distinction between useful, redundant, and extra actions according to Duncan & Gray [34]. Then what about the differences between two useful actions? Reasonably, we can argue that an action eliminating two faults is better than another action eliminating only one. How can we formalize this argument? On the other hand, we can identify premature diagnoses [34, 36], but how can we distinguish between two premature diagnoses? Furthermore, can we describe over-matured diagnoses, which are stated with less strength than that provided by the information at hand? In next section, the authors attempt to resolve the questions above by proposing a new approach to describe the process of diagnosis.

3 A New Approach

In this section, the authors intend to propose a new approach of describing the process of diagnosis in fine details. This approach borrows ideas from Shannon's concept of information entropy [51] and fuzzy signal detection theory [52]. This approach is supposed to be the most feasible under its standard paradigm, while it can still prove useful in other circumstances.

Consider a diagnostic problem where there are in total n possible hypotheses about the cause of abnormality. Before a participant receiving a certain piece of information, the objective probabilities of these hypotheses are $\{p_i\}_{i=1,\ldots,n}$, and the entropy is $-\sum_{i=1}^{n} p_i \log_2(1/p_i)$. After the participant receiving a piece of information s, the probabilities rationally (by causal or Bayesian reasoning) become $\{p_i|s\}_{i=1,\ldots,n}$ and the entropy becomes $-\sum_{i=1}^{n}(p_i|s)\log_2(1/(p_i|s))$. We define the value of information s as the difference between the entropies, i.e.

$$\mathrm{IV}(s) = \left[-\sum\nolimits_{i=1}^{n}(p_i|s)\log_2\left(\frac{1}{(p_i|s)}\right)\right] - \left[-\sum\nolimits_{i=1}^{n} p_i \log_2\left(\frac{1}{p_i}\right)\right]. \tag{1}$$

Thus, the usefulness of each information seeking action can be evaluated by the value of the information obtained, given that the probabilities before and after receiving this information are known. One can easily verify that the redundant and extra actions in [34] are of zero information value, and is able to compare useful actions by how much information value they provide.

At some point of time, the participant is asked to make a set of judgments on the likelihoods of all hypotheses, and the results are $\{l_i\}_{i=1,\ldots,n}$ (scaled from 0 to 1). Meanwhile, the 'rational' probabilities of all hypotheses, after taking all the information that the participant has acquired, are $\{q_i\}_{i=1,\ldots,n}$. In terms of fuzzy signal detection theory, we can treat $\{q_i\}_{i=1,\ldots,n}$ as signals and $\{l_i\}_{i=1,\ldots,n}$ as responses. Resembling [52], we can have the outcomes of hit, miss, false alarm and correct rejection as

$$\text{Hit}, \mathrm{H_i} = \min(\mathrm{q_i}, \mathrm{l_i}); \tag{2}$$

$$\text{Miss}, \mathrm{M_i} = \max(\mathrm{q_i} - \mathrm{l_i}, 0); \tag{3}$$

$$\text{False alarm}, \mathrm{FA_i} = \max(\mathrm{l_i} - \mathrm{q_i}, 0); \tag{4}$$

$$\text{Correct rejection}, \mathrm{CR_i} = \min(1 - \mathrm{q_i}, 1 - \mathrm{l_i}). \tag{5}$$

Also we can have hit rate (HR), miss rate (MR), false alarm rate (FAR) and correct rejection rate (CRR) as

$$\text{Hit rate}, \mathrm{HR} = \sum\nolimits_{i=1}^{n} \mathrm{H_i} / \sum\nolimits_{i=1}^{n} \mathrm{q_i}; \tag{6}$$

$$\text{Miss rate}, \mathrm{MR} = \sum\nolimits_{i=1}^{n} \mathrm{M_i} / \sum\nolimits_{i=1}^{n} \mathrm{q_i}; \tag{7}$$

$$\text{False alarm rate}, \mathrm{FAR} = \sum\nolimits_{i=1}^{n} \mathrm{FA_i} / \sum\nolimits_{i=1}^{n} (1 - \mathrm{q_i}); \tag{8}$$

$$\text{Correct rejection rate}, \mathrm{CRR} = \sum\nolimits_{i=1}^{n} \mathrm{CR_i} / \sum\nolimits_{i=1}^{n} (1 - \mathrm{q_i}). \tag{9}$$

Since HR + MR = 1 and FAR + CRR = 1, we can describe the performance of the current judgment by HR and FAR. One can verify that premature diagnoses [34] tend to have high value of FAR.

In this paradigm, we describe information seeking by evaluating the value of information obtained in each action, and describe information integration by evaluating HR and FAR derived from the results of likelihood judgments and 'rational' probabilities considering all acquired information. In addition, we view the set of likelihood judgment made first and early in the process as hypothesis generation. Thus, we also describe hypothesis generation in terms of HR (or MR if more preferred) and FAR.

This approach enables us to describe the process of diagnosis with fine-grained scalars. As a result, we can compare information seeking actions or judgments by referring to reasonable quantities. We may even be able to reveal some changes in human performance that we cannot realize when described by rough classifications or subjective ratings. If properly customized, this approach can help to compare results of studies on various diagnostic tasks based on a common benchmark.

There are certainly several methodological drawbacks about this approach. First, the provisions of 'rational' probabilities are effort demanding even impossible. If we can identify clear relations between hypotheses and acquirable information, we can obtain the 'rational' probabilities by causal or Bayesian reasoning. This may be the case for some lab experiments. However, for many diagnostic tasks under complex or dynamic contexts, 'rational' probabilities are hardly available. Second, we must set up the environment so that all information-seeking actions are clearly observable, consequently introducing intrusiveness. Third, although not always adding up to unity [38], likelihood judgments on different hypotheses are certainly correlated, undermining the validity of HR and FAR. Fourth, asking participants to judge all hypotheses is not essentially the hypothesis generation in common sense and is likely to disturb the latter.

Nevertheless, the authors never attempt the proposed approach to fit all types of diagnostic tasks. Just as those variables discussed in the previous section, the authors recommend using this approach when appropriate, especially when there is a need to describe the sub-processes of information seeking or information integration more in detail.

4 Conclusion

In various industries, humans are facing challenges from diagnostic tasks, which can be carried out in a diverse range of ways. There are unceasing attempts to analyze and improve human diagnosis in many contexts of interest. None of these attempts can avoid the necessity of properly describing the process of diagnosis. With this notion in mind, this article reviewed existing ways of describing the process of diagnosis from the literature and proposed a new approach. The results of literature review indicate that the cited studies utilized many types of variables with different emphases on separate sub-processes and the overall process of diagnosis. The proposed approach is supposed to describe information seeking and information integration with well-defined and precise quantities. After all, one should choose variables that are most suitable to describe the diagnostic tasks of interest.

Acknowledgements. This study was supported by the National Natural Science Foundation of China (Project no. 71371104).

References

1. Sheridan, T.B.: Telerobotics, Automation, and Human Supervisory Control. MIT Press, Cambridge (1992)
2. Mosier, K.L., Sethi, N., Mccauley, S., Khoo, L., Orasanu, J.M.: What you don't know can hurt you: factors impacting diagnosis in the automated cockpit. Hum. Factors **49**(2), 300–310 (2007)
3. Ha, J.S., Seong, P.H.: A review of studies on operator's information searching behavior for human factors studies in NPP MCRS. Nucl. Eng. Technol. **41**(3), 247–270 (2009)
4. Rasmussen, J., Rouse, W.B.: Human Detection and Diagnosis of System Failures. Springer, New York (1981)
5. Kim, J.H., Seong, P.H.: The effect of information types on diagnostic strategies in the information aid. Reliab. Eng. Syst. Saf. **92**(2), 171–186 (2007)
6. Rasmussen, J.: Diagnostic reasoning in action. IEEE Trans. Syst. Man Cybern. **23**(4), 981–992 (1993)
7. Elstein, A.S.: Thinking about diagnostic thinking: a 30-year perspective. Adv. Health Sci. Educ. **14**(1), 7–18 (2009)
8. Coderre, S., Mandin, H., Harasym, P.H., Fick, G.H.: Diagnostic reasoning strategies and diagnostic success. Med. Educ. **37**(8), 695–703 (2003)
9. Gruppen, L.D., Frohna, A.Z.: Clinical reasoning. In: International Handbook of Research in Medical Education, pp. 205–230. Springer, Netherlands (2002)

10. Morris, N.M., Rouse, W.B.: The effects of type of knowledge upon human problem solving in a process control task. IEEE Trans. Syst. Man Cybern. **SMC-15**(6), 698–707 (1985)
11. Shefherd, A., Marshall, E.C., Turner, A., Duncan, K.D.: Diagnosis of plant failures from a control panel: a comparison of three training methods. Ergonomics **20**(4), 347–361 (1977)
12. Wickens, C.D., Hollands, J.G., Banbury, S., Parasuraman, R.: Engineering Psychology and Human Performance, 4th edn. Pearson Education, New Jersey (2013)
13. Elstein, A.S., Shulman, L.S., Sprafka, S.A.: Medical Problem Solving: An Analysis of Clinical Reasoning. Harvard University Press (1978)
14. Weber, E.U., Böckenholt, U., Hilton, D.J., Wallace, B.: Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. J. Exp. Psychol. Learn. Mem. Cogn. **19**(5), 1151–1164 (1993)
15. Elstein, A.S., Shulman, L.S., Sprafka, S.A.: Medical problem solving: A ten-year retrospective. Eval. Health Prof. **13**(1), 5–36 (1990)
16. Morris, N.M., Rouse, W.B.: Review and evaluation of empirical research on troubleshooting. Hum. Factors **27**(5), 503–530 (1985)
17. Norman, G.: Research in clinical reasoning: past history and current trends. Med. Educ. **39** (4), 418–427 (2005)
18. Kim, J., Kim, J., Park, J., Jang, S.C., Shin, Y.C.: Some empirical insights on diagnostic performance of the operating crew in a computer-based advanced control room. Hum. Factors Ergon. Manuf. Serv. Ind. **21**(4), 379–396 (2011)
19. Chen, K.J., Li, Z.Z.: How does information congruence influence diagnosis performance? Ergonomics **58**(6), 924–934 (2015)
20. Schraagen, J.M.C., Schaafstal, A.M.: Training of systematic diagnosis: a case study in electronics troubleshooting. Le Travail Humain **59**(1), 5–21 (1996)
21. Besnard, D., Bastien-Toniazzo, M.: Expert error in trouble-shooting: An exploratory study in electronics. Int. J. Hum. Comput. Stud. **50**(5), 391–405 (1999)
22. Reed, N.E., Johnson, P.E.: Analysis of expert reasoning in hardware diagnosis. Int. J. Man-Mach. Stud. **38**(2), 251–280 (1993)
23. Rasmussen, J., Jensen, A.: Mental procedures in real-life tasks: a case of electronic trouble shooting. Ergonomics **17**(3), 293–307 (1974)
24. Patrick, J.: Analysing operators' diagnostic reasoning during multiple events. Ergonomics **42** (3), 493–515 (1999)
25. Munley, G.A., Patrick, J.: Training and transfer of a structural fault-finding strategy. Ergonomics **40**(1), 92–109 (1997)
26. Patrick, J., James, N., Friend, C.: A field study of training fault-finding. Le Travail Humain **59**(1), 23–44 (1996)
27. Hochholdinger, S., Schaper, N.: Training troubleshooting skills with an anchored instruction module in an authentic computer based simulation environment. J. Tech. Educ. **1**, 7–22 (2013)
28. Kluge, A., Termer, A.: Human-centered design (HCD) of a fault-finding application for mobile devices and its impact on the reduction of time in fault diagnosis in the manufacturing industry. Appl. Ergon. **59**, 170–181 (2017)
29. Raaijmakers, J.G.W., Verduyn, W.W.: Individual differences and the effects of an information aid in performance of a fault diagnosis task. Ergonomics **39**(7), 966–979 (1996)
30. Brinkman, J.A.: Verbal protocol accuracy in fault diagnosis. Ergonomics **36**(11), 1381–1397 (1993)
31. Ross, C.T., John, A.A.: The reduction of uncertainty and troubleshooting performance. Hum. Factors **39**(2), 254–267 (1997)
32. Rouse, W.B.: Human Problem Solving Performance in a Fault Diagnosis Task. IEEE Trans. Syst. Man Cybern. **8**(4), 258–271 (1978)

33. Ham, D.H., Yoon, W.C.: The training effects of principle knowledge on fault diagnosis performance. Hum. Factors Ergon. Manuf. **17**(3), 263–282 (2007)
34. Kulatunga-Moruzi, C., Brooks, L.R., Norman, G.R.: Using comprehensive feature lists to bias medical diagnosis. J. Exp. Psychol. Learn. Mem. Cogn. **30**(3), 563–572 (2004)
35. Duncan, K.D., Gray, M.J.: An evaluation of a fault-finding training course for refinery process operators. J. Occup. Sychol. **48**, 199–218 (1975)
36. Henneman, R.L., Rouse, W.B.: Measures of human problem solving performance in fault diagnosis tasks. IEEE Trans. Syst. Man Cybern. **SMC-14**(1), 258–271 (1984)
37. Dougherty, M.R.P., Hunter, J.E.: Hypothesis generation, probability judgment, and individual differences in working memory capacity. Acta Physiol. (Oxf) **113**, 263–282 (2003)
38. Thomas, R.P., Dougherty, M.R., Sprenger, A.M., Harbison, J.I.: Diagnostic hypothesis generation and human judgment. Psychol. Rev. **115**(1), 155–185 (2008)
39. Anderson, J.R.: Cognitive Psychology and Its Implications, 7th edn. Worth Publishers, New York (2009)
40. Bergus, G.R., Chapman, G.B., Levy, B.T., Ely, J.W., Oppliger, R.A.: Clinical diagnosis and the order of information. Med. Decis. Making **18**(4), 412–417 (1998)
41. Brehmer, B.: Models of Diagnostic Judgment. Human Detection and Diagnosis of System Failures, pp. 231–239. Springer, US (1981)
42. Cornell, F.G., Damrin, D.E., Saupe, J.L, Crowder, N.A.: Proficiency of Q-24 Rada Machanics: III. The Tab Test – A Group Test of Trouble-Shooting Proficiency. Air Force Personnel and Training Center Res. Bull., AFPTRC-TR-54-52 (1954)
43. Finch, C.R.: Troubleshooting instruction in vocational technical education via dynamic simulation. Pennsylvania Dep. Education Project, Rep. 10064, 19-1024 (1971)
44. Glass, A.A.: Problem solving techniques and troubleshooting simulations in training electronic repairmen. Ph.D. dissertation. Columbia University, New York (1967)
45. Patrick, J., Haines, B.: Training and transfer of faultfinding skill. Ergonomics **31**(2), 193–210 (1988)
46. Sanderson, P.M.: Knowledge acquisition and fault diagnosis: experiments with PLAULT. IEEE Trans. Syst. Man Cybern. **20**(1), 225–242 (1990)
47. Steinemann, J.H.: Comparison of performance on analogous simulated and actual troubleshooting tasks, U.S. Naval Personnel Research Activity, Mem. SRM 67-1 (1966)
48. Yoon, W.C., Hammer, J.M.: Deep-reasoning fault diagnosis: an aid and a model. IEEE Trans. Syst. Man Cybern. **18**(4), 659–676 (1988)
49. Dawes, R.M., Faust, D., Meehl, P.E.: Clinical versus actuarial judgment. Science **243**(4899), 1668–1674 (1989)
50. Kahneman, D., Tversky, A.: On the psychology of prediction. Psychol. Rev. **80**(4), 237–251 (1973)
51. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **5**(1), 3–55 (1948)
52. Parasuraman, R., Masalonis, A.J., Hancock, P.A.: Fuzzy signal detection theory: basic postulates and formulas for analyzing human and machine. Hum. Factors **42**(4), 636–659 (2000)

A Case Study of Human Reliability in Maquiladora Industry Manufacturing Process

Teresa Carrillo-Gutierrez[1(✉)], Rosa María Reyes Martínez[2],
Jorge de la Riva Rodríguez[2], Jaime Sanchez-Leal[2],
and Aidé Aracely Maldonado Macias[3]

[1] Facultad de Cs. Químicas e Ingeniería,
Universidad Autónoma de Baja California,
Calzada Universidad 14418, Tijuana, Baja California, Mexico
tcarrillo@uabc.edu.mx, d13111803@itcj.edu.mx
[2] División de Estudios de Posgrado e Investigación,
Tecnológico Nacional de México, Instituto Tecnológico de Ciudad Juárez,
Av. Tecnológico 1340, El Crucero, Cd. Juárez, Chihuahua, Mexico
rosyreyes2001@yahoo.com, {jriva, jsanchez}@itcj.edu.mx
[3] Departamento de Ingeniería Industrial y Manufactura,
Universidad Autónoma de Cd. Juárez,
Del Charo Av. 450 N, 32310 Ciudad Juárez, Mexico
amaldona@uacj.mx

Abstract. This document presents an advance study of the sociotechnical factors that influence the human errors that affect the manufacture processes in the maquiladora industry. The purpose of the research is to determine the root causes of human errors from an approach of the human reliability in sociotechnical systems. The scientific methodology used is located in cognitive anthropology, by applying methods of analysis from a cultural consensus. It was used on data collection, an instrumented and structured interview known as free listings. The methodological approach of the study corresponds to a mixed methods. The study was designed transversal, exploratory and descriptive. The interviewed population was supervisors, technicians and engineers, of the two main modalities existing in the companies production area and those from quality area. With this research it is possible to make a contribution for a better classification of root causes in human errors affecting now the manufacture processes in the maquiladora industry in Tijuana, Mexico. With this knowledge may propose preventive and corrective actions in continuous improvement projects inside manufacture processes.

Keywords: Human reliability · Error analysis · Free listings · Cultural consensus

© Springer International Publishing AG 2018
R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3_35

1 Introduction

The purpose of this work of research is to determine the root cause of human errors from an approach of human reliability in socio-technical systems. First off, analytical methodologies and measurement of human reliability in the manufacturing processes published in the specialized literature and, afterwards, a study advance is presented of human reability performed in the manufacturing industry in Tijuana.

According to [1], the manufacturing industry has become a characteristic element of the north border region of Mexico due to the accelerated changes that its presence has produced in terms of population growth, filial industries or providers, commerce and services. The term "maquiladora" (factory) is a derivative of "maquila", word of Arabic origin which in its first uses related to the activity of grinding, which refers to the portion of grain that corresponds to the grinder in exchange for its service, the meaning of the term evolved until designating "any particular activity in an industrial process - for example, assembling or packaging - done for a part that is not of the original manufacturer" [2]. Based on this definition, factory refers to a wide industry made up of a great variety of goods and services. Modern factories carry out product assembly operations which, once processed, create products which are re-exported to the US and other countries. In Baja California, there are 906 registered establishments, according to the National Council of the Manufacturing Industry and Export Manufacturing [3] and National Institute of Statistic and Geography (INEGI) [4].

2 Literature Review

2.1 The Cultural Consensus Theory

Cognitive anthropology provides a useful theoretical orientation in the study of culture. This perspective explains that culture is composed of a structure of schemed interconnection, shared knowledge that constructs meanings, representations of social reality, direct behaviors and easiness of behavior interpretation [5–7].

Cognitive anthropology is the study of the relationship between society and human thinking. Cognitive anthropologists study how people that belong to social groups think in regard to the objects and events that make up their world, from physical objects to abstract events [8]. They have developed methodological focuses to measure and compare the shared cultural models, which have as an objective to reduce to a minimum the *etic* focus and maximize the *emic* aspects of investigation and, for that matter, utilize mixed techniques (qualitative and quantitative) in the compilation of data [9].

Culture is the set of behaviors and beliefs, learned and shared. The Cultural Consensus Theory (CCT) is a collection of analytical techniques that can be used to estimate cultural beliefs (which are normative of a group) and the degree in which individuals learn and report it. The CCT estimates culturally correct answers to a series of questions (group beliefs) and simultaneously estimate the knowledge of each informer or the degree in which the answers are shared [10].

Rommey and collaborators [11] establish the use of the boss as a central idea of their theory according to or consesus between the informers, to make inference regarding their differential competition in knowledge of the shared information, which constitutes the culture. Scientists suppose that correspondence between the answers of two informers is a function of the measure in that each one correlates with the truth.

Cultural consensus can be seen as a theory and a methodology [10]. It is a theory in the sense that the idea formalizes, that the culturally correct information can be established from the individuals' shared knowledge. It's a methodology because it can be applied in different topics and circumstances. An advantage of the concensus model is that it incorporates the derivatives and concepts of well-established theories and is based on rigorous postulations [11].

From theoretical and methodological perspectives, the development of the consensus is based on the following three postulations [5]:

1. Common truth. It's the existence of a fixed key answer that is applicable to all informants, which is to say, that a correct answer exists for each aspect. This just means that it is assumed that all the informants come from a common culture and if the common truth is the same for all the informants in the sample, the postulation of common truth is fulfilled.
2. Local independence. Suppose that the answers of each informant are given independently from those of the other informants, in other words, an answer is not influenced by the knowledge or presence of another one.
3. Item homogeneity. Each informant has a fixed cultural competition regarding all the questions. This is a postulated fort which says that these are all of the same level of difficulty; this postulation refers to guarantee that the questions are extracted of a coherent domain.

2.2 Analysis of Cultural Domain and the Mixed Methods of Investigation

The analysis of cultural domain begins with open interviews, traditional and semi-structured, and some of them more specialized, like the free list. The objective of analysis is to generate the terms which individuals use to speak in respect to a cultural domination in particular. Once the set of terms that make up a domain have been generated, the similarities and differences of meaning between the terms are analyzed using a variety of interview techniques [5].

The free list technique has been widely used by outstanding investigators in the field of Cognitive Anthropology [8, 12]. The listings can identify elements in *emic* category and acumulate data fast and easy. The listings are a well-established ethnographic method that are based on three postulations: the first makes reference to some terms that stand out more, better known, important or familiar and when people do free listings they tend to mention it on first term. Second, individuals that possess higher knowledge in regard to the subject are listed down more than those that possess less knowledge. The third postulation indicates that the terms that are mostly mentioned indicate the most outstanding elements of the topic [13, 14]. Free listings are the general technique of major utilization for isolating and defining a cultural domain.

The draw for piles without restrictions is perhaps the best example of qualitative and quantitative data. In this technique the participants simply indicate how the terms are grouped over the foundations of similar meaning and separated from other groups over the difference of meaning [5]. It is an investigation technique to classify the elements of a cultural domain. The Taxonomies or tree structures can be obtained in an interview with informers that classify the elements in groups and later divide them into smaller groups [13].

2.3 The Human Reliability

In the work of Garcia [15] the following definitions are introduced: Operational Reliability is one of the modern strategies which generate great benefits to whose who apply it. It is based on statistical analysis and conditional analysis, guided to maintain availability and reliability in the equipment with active participation of company personnel. Operational Reliability has the industrial capacity implicit; in other words, it includes the processes, technology and people to fulfill its function or the purpose that is awaited of it within its design limits and under a specific operational context.

It is important to remark out that in an Operational Reliability to system the analysis of its four operational fronts are necessary: Human Reliability, Process Trust, Equipment Reliability and Design Reliability, over which must be acten upon if continous improvement is desired on the long-term. Any fact isolated from improvement can bring benefits, but without considering the other factors, its advantages are limited and diluted and they progress to be only the result of one project and not of an organizational change [15].

Also, it is considered that Human Reliability is a discipline which is part of the field of systems reliability in the measure that the man is considered an integrating part of a system. It is considered that the human component is of much higher complexity than any other component and, thus, the applicable techniques to the study of human reliability or, complementarily, of human error are specific and integrate psychological and organizational aspects to the habitual mathematical techniques.

Now, a large amount of human trust definitions exist. Arquer and Nogareda [16] define it as “the body of knowledge that refer to prediction, analysis and reduction of human error, focusing on the role of the person on the design operations, maintenance, use and gestation of a sociotechnical system”. For that matter, human reliability has human error as a study objective.

Namely, the dominating definition of human error is planted by Reason [17] who defines it as “a generic term which accompanies all those occasions in which a sequence of physical or mental activities fails in reaching its desired results and when these failures can’t be attributed to the intervention of an opportunity”. Also, human error is a complex construct which has received constant attention between the studious of human factors in the dynamic and complex systems [5]. Human error is also defined as the behavior of people that exceed the tolerance limit defined for a particular task [18]. In general, investigators define it as the cause of an action, like something that was done wrong, or like it came out wrong [19]. The study of human error has also been approached from three different focuses: Engineering, Congnitive Psychology and Cognitive Ergonomics [20].

3 Method

A literature revision is presented regarding human error in the context of human reliability in the industry and service processes. Results of a structured questionnaire are also presented directed towards the quality area personnel to identify which are the human errors that occur with the most frequency and create defects in its product, like knowing its possible causes. A convenience sample was carried out due to the goals were of probe and not generalization. Forty six enterprises answered the questionnaire. A second instrument was applied called questionnaire of free listings was applied in three firms to learn the opinion of the experts of the area of quality, production, engineering and maintenance.

4 Results

4.1 Human Error

Understanding the human factor error is to understand a complexity of interrelated elements [21], among which are:

1. The mental processes as the reception and identification of information, decision making and its relationships with superior mental functions such as perception, attention, memory, intelligence, etc.
2. The organizational factors as management mode and the role of supervisors, among others.
3. Physiological factors, such as mental and physical illnesses, the deterioration of the visual and auditive systems, aging, etc.
4. Personal factors such as the hardly avoidable extra-laboral problems.
5. The transient states of anxiety, fatigue, etc.
6. The level of routine and monotony of the task.
7. The questionable necessity of "deviation" from the norms for the range of imposed objectives for the task. With regard to human error, Sebastian [21] points out the following aspects:
    - Potentiality: an action does not have to result in a degradation of the system's performance or in wanted effects to be considered as an error; it is enough for those mistakes to be triggered in the future.
    - Actors: human error is committed by: operators, designers, supervisors, directors, maintenance personnel.
    - Consequences: error is defined in terms of unwanted effects about effectiveness, security or performance.
    - Origin: The effects are "activated" by inappropriate and unwanted demeanors.

For Garcia [15], when the interaction between people and production systems is considered, human errors can be classified in four categories:

1. Anthropometric factors: they are related to the size and the physical resistance of the operator who will perform a task when the operator can't physically adjust to the

conditions of the system or equipment. Their errors do not constitute the cause of the problem, the majority of the causes are the effect of a system failure, which requires a modification or redesigning.

2. Sensory factors: They relate with the expertise with which people use their senses to see what is going on in their environment. They have to do with aspects like good visibility or noise level, which require to mitigate a corrective action.
3. Physiological factors: They refer to environmental tensions which affect human performance, since they generate fatigue. To reduce them, changes in the organizational clime must be made or in the processes to be done.
4. Psychological factors: They refer to the internal aspects which have root in the psychic of people. They can generate intentional errors or unintentional and in the majority of cases they require specialized treatment.

4.2 Human Error Analysis

In a general and approximative way, Sebastian [21] classifies the perspectives of the human error analysis within the following four runnings: (a) Explanations outside of the person, (b) Explanations inside of the person, (c) Explanations from the person-machine interaction, (d) Explanations from the person-context relationship.

According to Faig [22], to obtain a correct measurement of the system reliability, the contribution from human error must be considered. The analyses of the systems design, from procedures and posterior accident reports show that human error can cause an immediate accident or can play an important role in the development of unwanted events. Without the incorporation of human error probabilites, the results are incomplete and often valued incorrectly. For the estimation of human error probability, human error must be understood along with the variables which determine it. The modeling of human error is complex in such way that assumptions, mechanisms and approximations that are used for the behavior model can't be utilized for all the behavior models of human activities [22].

The utilized models have distinct limitations, mainly due to the following: (A) Human behavior is a complex subject that can not be described as a simple component. Human behavior can be affected by diverse factors, such as social, environmental, psychological and diverse physical factors, which are difficult to model and quantify. (B) Human actions can't be considered to have states of binary failure/success, like the case of component failures. On the other hand, human interactions can't be analyzed by the human trust analysis and cover all the spectrum of human interactions. (C) The major problem of the human trust analysis is in the lack of human behavior data.

On the improvement on the reliability of industrial systems, the key point is in the man-machine interaction, but it's obvious that it turns out much more complicated from that of man than the machine. This is a difficulty which has given rise to many lines of mutlidisciplinary investigations, especially in those industrial sectors where the impact of possible human errors is stronger, which are nuclear energy, aviation and chemical industry [23].

Table 1. Some studies carried out in the field of production, quality and safety.

Context	Methodological approach	Focus on study of human error	Source
Production	Quantitative	Engineering	Fan-Jang (2000) [24]
	Quantitative	Engineering	Kumar et al. (2007) [25]
	Quantitative	Engineering	Myszewski (2010) [26]
	Quantitative	Engineering	Miralles et al. (2011) [27]
	Qualitative-Quantitative	Cognitive Ergonomics	Viña et al. (2011) [28]
	Quantitative	Engineering	Aju kumar & Gandhi (2011) [29]
	Quantitative	Engineering	Rigolin & Quartucci (2013) [30]
	Qualitative-Quantitative	Cognitive Psychology	Báez et al. (2013) [31]
Quality	Quantitative	Engineering	Sylla & Drury (1995) [32]
	Qualitative	Cognitive Psychology	Collazo (2008) [18]
	Quantitative	Engineering	Paun et al. (2011) [33]
	Quantitative	Engineering	Le et al. (2012) [34]
	Qualitative	Ergonomics	Sharma (2012) [35]
	Quantitative	Engineering	Pons et al. (2013) [36]
	Qualitative-Quantitative	Engineering	Power & Fox (2014) [37]
Security	Qualitative-Quantitative	Cognitive Ergonomics	Reyes-Martínez et al. (2012) [38]
			Reyes-Martínez et al. (2014) [39]
	Qualitative	Cognitive Ergonomics	Stave & Törner (2007) [40]
	Quantitative	Engineering	Peng-cheng et al. (2010) [41]
	Quantitative	Engineering	Fujita & Hollnagel (2004) [42]

On Table 1, the studies which have identified in the literature of human errors which contribute to the defects or failures in the processes or products in the context of the manufacturing industry in the production scope, quality scope and security.

4.3 Results from the Questionnaires

This study is in progress and, to this day, forty six enterprises of different industrial branches have been interviewed, all of them located in the city of Tijuana, Baja California, Mexico. Out of all the enterprises, nineteen are from the branch of medical products, twelve of the electronic products branch, four of plastics, three of aerospace branch, three of motor branch, three of furniture and two of metal-mechanic branch.

In the first open question, it is questioned how quality defects are detected in their enterprise. All the interviewed enterprises, all forty six, answer that the inspection methods are the main way in which they detect quality defects. The second question is: how do enterprises reduce incidence of the defects caused by human error? The most frequent answers are: PokaYoke, training, documentation and process flow followed by corrective and preventive actions (CAPA). In the third question, statistical techniques, inspection and 8D are the most frequent answers to the following question: What are the techniques which your enterprise utilizes to detect the causes of human errors and its effect on the quality of the product? Which confirms the engineering focus of enterprise trust analysis which responds the questionnaire.

For the fourth answer, enterprises are offered a listing of ten and seven factors of which are requested to be marked whichever could be the causes of human errors which affect the quality of the product in its enterprise. The factor which corresponds to the experience is obtained, also called the learning curve, in first place of frequency, following training. At descending frequency overload of work is continued and then communication. Afterwards, usage of inadequate or badly-shaped tools is followed. Bad detection of errors on pilot runs which escape to the production line, characteristics of the task, in other words, difficulty to perform the production operations and fatigue are found in equal to frequency.

On the fifth question, the importance which human error presents in the incidence of the product defects in its enterprise is questioned. Thirty-two interviewed enterprises answer that it is important between regular and very important. On question six, enterprises are offered a list of human errors and it is requested for them to mark down all which they consider to be able to provoke defects in the manufacturing process of their products. The most frequent errors are caused by procedure omissions, lack of experience from the operator along with distractions followed by inadequeate training.

The seventh open question is: Does your enterprise use a methodology to analyze and evaluate the human error in the incidence of the defects of the product in your enterprise? 70% of the enterprises answer affirmatively. All of them point out that the focus of the used methodology is engineering, in other words, mainly employ statistical techniques. Also, an approximation about possible causes of quality defects in the product was included in the draft survey as question eight. According to the opinion of the participants, show a predominance of the machine factor followed by the worker training factor. In question nine, enterprises are asked the following: In your enterprise, which of the following actions currently apply to eliminate or reduce human errors which cause quality defects on the product? The answer with most frequency was PokaYoke followed by Automization in congruence with the result of the second question.

The second questionnaire, known as free listings, has been applied in three enterprises, two of the medical field and another of the plastics field. It was twenty people in total who answered the questionnaire. Currently, the answers of the free listings are found in analysis.

5 Conclusions

The majority of the interviewed enterprises detect defects in the quality of the product through engineering methods of inspection. PokaYoke, preventive and corrective actions (CAPA), documentation and flow of process, as well as training are the most utilized strategies to reduce the incidence of defects caused by human error. The three most utilized techniques to detect causes of human errors and its effect on the quality of the product are the statistical techniques, inspection and FMEA. The results of the second questionnaire knows as free listings is found in analysis.

References

1. Douglas, L., Hansen, T.: Los orígenes de la industria maquiladora en México. Comercio Exterior **53**(11), 1045–1056 (2003)
2. Angulo, C.: Foreign lnvestment and the Maquiladora Export lndustry. Direct Foreign lnvestment. Banamex, México (1990)
3. Consejo Nacional de la Industria Maquiladora y Manufacturera de Exportación: Economic Report. Economic Analysis-Banco (2016)
4. Instituto Nacional de Estadística y Geografía (INEGI) (2016). http://www.inegi.org.mx
5. Reyes, R.M.: Desarrollo de una taxonomía de la causalidad del error humano para accidentes con lesión en manos en la industria arnesera. Tesis doctoral, Universidad de Guadalajara (2011)
6. Collins, C.C., Dressler, W.W.: Cultural consensus and cultural diversity: a mixed methods investigation of human service providers' models of domestic violence. J. Mixed Methods Res. **2**(4), 362–387 (2008)
7. D'Andrade, R.G.: The cultural part of cognition. Cogn. Sci. **5**, 179–195 (1987)
8. D'Andrade, R.G.: The Development of Cognitive Anthropology. Cambridge University Press, Cambridge (1995)
9. Tashakkori, A., Creswell, J.W.: A new era for mixed methods. J. Mixed Methods Res. **1**(1), 3–7 (2007)
10. Weller, S.R.: Cultural consensus theory: applications and frequently asked questions. Field Methods **19**(4), 339–368 (2007)
11. Romney, A.K., Weller, S.C., Batchelder, W.H.: Culture as consensus: a theory of culture and informant accuracy. Am. Anthropol. **88**, 313–338 (1986)
12. Schrauf, R.W., Sanchez, J.: Using freelisting to identify, assess, and characterize age differences in shared cultural domains. J. Gerontol. **63b**(6), S385–S393 (2008)
13. Weller, S.C., Romney, A.K.: Systematic Data Collection. Sage, Newbury Park, Beverly Hills, London, New Delhi (1988)
14. Quinlan, M.: Considerations for collecting freelists in the field: examples from ethnobotany. Field Methods **17**(3), 219–234 (2005)
15. García, O.: La Confiabilidad humana en la gestión de mantenimiento. Foro Taller sobre Innovación en el Sector Industrial de Procesos por Análisis de la Seguridad y el Control de Riesgos. Medellín, Colombia (2006)
16. Arquer, M.I., Nogareda, C.: NTP 360: Fiabilidad humana: Conceptos básicos. Instituto Nacional de Seguridad e Higiene en el Trabajo. Recuperado el 12 de mayo de 2016 del sitio Web del Ministerio de Trabajo y Asuntos Sociales de España (1994). http://www.insht.es/
17. Reason, J.T.: Human Error. Cambridge University Press, Cambridge (1990)

18. Collazo, G.: Error Humano: C4c06. Modelo de Determinación de Raíz Causa, Puerto Rico (2008)
19. Casares-Li, R., Rodríguez-Hernández, A., Viña-Brito, S.: Análisis de Errores Humanos mediante la Tecnología TErEH: experiencias en su aplicación. Rev. Ing. Ind. **37**, 49–58 (2016). Cuba
20. Cañas, J., Waerns, Y.: Ergonomía Cognitiva Aspectos psicológicos de la interacción de las personas con la tecnología de la información. Editorial Médica Panamericana, España (2001)
21. Sebastián, M.L.: Fallo humano: la quiebra de un paradigma. Apuntes de Psicología **27**(1), 21–51 (2009)
22. Faig, J.: NTP 619: Fiabilidad humana: evaluación simplificada del error humano (I). Instituto Nacional de Seguridad e Higiene en el Trabajo. Ministerio de Trabajo y Asuntos Sociales España (2003)
23. Ruiz-Moreno, J., Trujillo, H.: Modelos para la evaluación del error humano en estudios de fiabilidad de sistemas. Anales de Psicología **28**(3), 963–977 (2012)
24. Fan-Jang, Y., Sheue-Ling, H., Yu-Hao, H., Jinn-Sing, L.: Application of human error criticality analysis for improving the initiator assembly process. Int. J. Ind. Ergon. **26**, 87–99 (2000)
25. Kumar, R., Kumar, D., Kumar, P.: FM- a pragmatic tool to model, analyse and predict complex behavior of industrial systems. Eng. Comput. **24**(4), 319–346 (2007)
26. Myszewski, J.: Mathematical model of the occurrence of human error in manufacturing processes. Qual. Reliab. Eng. Int. **26**, 845–851 (2010)
27. Miralles, C., Holt, R., Marin-Garcia, J., Canos-Daros, L.: Universal design of workplaces through the use of Poka-Yokes: case study and implications. J. Ind. Eng. Manage. **4**(3), 436–452 (2011)
28. Viña, S., Rodríguez, A., Montero, R., Casares, R., Martínez, R.: El estudio del error humano en la industria biofamaceutica. Ação Ergonômica. Revista Brasileira de Ergonomía **6**(2), 31–37 (2011)
29. Aju kumar, V., Gandhi, O.: Quantification of human error in maintenance using graph theory and matrix approach. Qual. Reliab. Eng. **27**, 1145–1172 (2011)
30. Rigolin, M., Quartucci, C.: Application of human error theories for the process improvement of requirements engineering. Inf. Sci. **250**, 142–161 (2013)
31. Báez, Y., Rodríguez, M., De la Vega, E., Tlapa, D.: Factores que influyen en el error humano de los trabajadores en líneas de montaje manual. Información Tecnológica **24**(6), 67–78 (2013)
32. Sylla, C., Drury, C.: Signal detection for human error correction in quality control. Comput. Ind. **26**, 147–159 (1995)
33. Paun, A., Sergiu, D., Vladut, V., Gageanu, P.: Reducing the time consuming "coming back" in manufacturing process by using the anti-error systems. Ann. Fac. Eng. Hunedoara. Int. J. Eng. Tomo **IX**(3), 319–322 (2011)
34. Yang, L., Su, Q., Shen, L.: A novel method of analyzing quality defects due to human errors in engine assembly line. In: International Conference on Information Management, Innovation Management and Industrial Engineering (ICIII), vol. 3 (2012)
35. Sharma, R.: Conceptual framework for improving business performance with lean manufacturing and successful human factors interventions–a case study. Int. J. Qual. Res. **6**(3), 259–270 (2012)
36. Pons, R., Villa, M., Bermúdez, Y.: El análisis de fiabilidad humana en la mejora de procesos. Prospect **11**(2), 61–67 (2013)
37. Power, C., Fox, J.: Comparing the comprehensiveness of three expert inspection methodologies for detecting errors in interactive systems. Saf. Sci. **62**, 286–294 (2014)

38. Reyes, R.M., Prado, L.R., Aguilera, Á., Soltero, R.: Descripción de los conocimientos sobre factores humanos que causan accidentes en una industria arnesera mexicana. e-Gnosis **9**, 1–17 (2011)
39. Reyes-Martínez, R.M., Maldonado-Macías, A.A., de la O, R., de la Riva-Rodríguez, J.: Theoretical approach for human factors identification and classification system in accidents causality in manufacturing environment. In: Lean Manufacturing in the Developing World, chap. 18. Springer International Publishing Switzerland (2014)
40. Stave, C.: Törner.: exploring the organisational preconditions for occupational accidents in food industry: a qualitative approach. Saf. Sci. **45**, 355–371 (2007)
41. Peng-cheng, L., Guo-hua, C., Li-cao, D., Li, Z.: Fuzzy logic-based approach for identifying the risk importance of human error. Saf. Sci. **48**, 902–913 (2010)
42. Fujita, Y., Hollnagel, E.: Failures without errors: quantification of context in HRA. Reliab. Eng. Syst. Saf. **83**, 145–151 (2004)

Author Index

© Springer International Publishing AG 2018

R.L. Boring (ed.), *Advances in Human Error, Reliability, Resilience, and Performance*,
Advances in Intelligent Systems and Computing 589, DOI 10.1007/978-3-319-60645-3