

2018

Task Complexity, and Operators' Capabilities as Predictor of Human Error: Modeling Framework and an Example of Application

Maria Chiara Leva

Technological University Dublin, maria.leva@tudublin.ie

Alberto Caimo

Technological University Dublin, alberto.caimo@tudublin.ie

R. Duane

Technological University Dublin

See next page for additional authors

Follow this and additional works at: <https://arrow.tudublin.ie/engschmanconn>



Part of the [Manufacturing Commons](#), and the [Risk Analysis Commons](#)

Recommended Citation

Leva, M.C., Caimo, A. & Duane, R. (2018). Task Complexity and Operators' Capabilities as Predictor of Human Error: Modeling Framework and an Example of Application. *28th International European Safety and Reliability Conference, ESREL 2018, Trondheim, Norway, 17 June to 21 June.*

This Conference Paper is brought to you for free and open access by the School of Manufacturing and Design Engineering at ARROW@TU Dublin. It has been accepted for inclusion in Conference Papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](#)

Authors

Maria Chiara Leva, Alberto Caimo, R. Duane, Micaela Demichela, and Lorenzo Comberti

Task complexity, and operators' capabilities as predictor of human error: Modeling framework and an example of application

M.C. Leva, A. Caimo & R. Duane
Dublin Institute of Technology, Ireland

M. Demichela & L. Comberti
Politecnico di Torino, Italy

ABSTRACT: This paper presents the initial framework adopted to assess human error in assembly tasks at a large manufacturing company in Ireland. The model to characterize and predict human error presented in this paper is linked conceptually to the model introduced by Rasch (1980), where the probability of a specified outcome is modelled as a logistic function of the difference between the person capacity and item difficulty. The model needs to be modified to take into account an outcome that is not dichotomous and feed into the interaction between two macro factors: (a) Task complexity: that summarises all factors contributing to physical and mental workload requirements for execution of a given operative task & (b) Human capability: that considered the skills, training and experience of the people facing the tasks, representing a synthesis of their physical and cognitive abilities to verify whether or not they are matching the task requirements. Task complexity can be evaluated as a mathematical construct considering the compound effects of Mental Workload Demands and Physical Workload Demands associated to an operator task. Similarly, operator capability can be estimated on the basis of the operators' set of cognitive capabilities and physical conditions. A linear regression model was used to fit a dataset collected in R. The estimation of task complexity and operator skills was used to estimate human performance in a Poisson regression model. The preliminary results suggest that both elements are significant in predicting error occurrence.

1 INTRODUCTION

1.1 *Scope of work and background*

This paper presents the initial framework adopted to assess human error in assembly tasks at a large manufacturing company in Ireland [1].

The aim of this study was to carry out an observational, empirical study on the existing human errors in the dispatching department, find a way to model the issue and if possible propose approaches to reduce and eliminate errors and variations in the end product. The company dispatches technology goods to national and international customers and the focus of the project was the assembly of goods for dispatch. Operators prepare the goods at workstations along conveyor lines, however at these conveyors inefficiencies and inaccuracies relating to human performance were identified. Two primary workstations were selected for inclusion in the dispatching unit based on their recorded error rates. Conditions vary and fluctuate at workstations, which may increase the probability of making mistakes, including the complexity and number of the activities, environmental conditions and the quality of the product. An understanding of both the

human nature (characteristics, feelings, and behavioural traits) and the impact of the features of the workstation on human nature (typology of activities, working load, anxiety induced, environmental factors etc.) was required to holistically determine the performance shaping factors for the workstations under examination. The focus is on the role of operator's capability to complete tasks and the means to reduce human errors whilst retraining product quality. Changes were proposed for the assembly lines at the dispatching stations, including changes in the procedures and training to employ an understanding of human performance and improvements to safety, with an overall beneficial impact on both productivity and quality.

The researcher conducted a task analysis of the critical activities completed by operators when packing out the variety of product units at two primary workstations. Questionnaires were prepared examining the skills requirements, skills rating of operators, mental workload requirements, physical workload requirements, perceived task complexity and motivation. Finally, the implementation of an applied model Task Execution Reliability Model (TERM) was used to identify the main fac-

tors affecting human performance for this settings. Three methods were used to inform the research:

1. Firstly, an examination of performance shaping factors in the literature to inform a set of specific questionnaires.
2. Secondly, the collection and analysis of the data from the questionnaires completed by operators, technicians, supervisors, group leaders and process engineers in the manufacturing facility familiar with the work undertaken at the workstations under examination.
3. Thirdly, focus group sessions were run discussing possible participatory redesign for process and procedures at the workstations
4. Finally the data from the questionnaire was also used to predict task complexity and error occurrences using two different types of regression models.

2 MODELLING HUMAN ERROR

2.1 Human error in manufacturing

Human nature can be shaped and driven by factors including individual characteristics, personal issues, physical and psychological conditions (Tooby & Cosmides, 1990). These factors interact with each other and may determine the output and productivity of the performance of the individual. Human performance is unavoidably susceptible to human error, as humans are not infallible and the occurrence of errors must be expected (Karl & Karl, 2012). Humans are often capable of recognising errors and rectifying such errors before any serious or critical consequences occur (Sheridan, 2008). With this in mind, human performance can be accepted and understood as the definitive product of the balance between task complexity and capability (Morgeson et al, 2010).

When the capabilities and limitations of humans are understood, incorporated and acknowledged, Harris (2006) argues that benefits can include increased efficiency and improved safety performance. Individual employee's competencies may be challenged by fatigue, stressors and unpredictability, whilst competencies may benefit from skills, training and a clear comprehension of the task (Miller and Parasuraman, 2007, Jo et al, 2012, Kostina et al, 2012). The capabilities of the operator and the physical skills required for the task must be taken into consideration when reviewing tasks and the errors associated with them (Harris, 2006). A balance between workload, both physical and mental, ought to be reached to reduce human errors among competent operators (Miller and Parasuraman, 2007).

2.2 The TERM model: Task execution reliability model

The model used is linked conceptually to the model introduced by Rasch (1980) to analyse correct or incorrect execution of a task as a function of the trade-off between (a) the respondent's abilities, attitudes or personality traits and (b) the item difficulty. In the Rasch model, the probability of a specified outcome (e.g. right/wrong results) is modelled as a logistic function of the difference between the person and item difficulty parameter.

The mathematical form of the model is provided in equation (1).

$$\Pr(X_{ni} = 1) = \frac{e^{\beta_n - \delta_i}}{1 + e^{\beta_n - \delta_i}} \quad (1)$$

Let X_{ni} be a dichotomous random variable with binary values where, for example, $X_{ni} = 1$ denotes a correct response and an $X_{ni} = 0$ an incorrect response to a given assessment item. In the Rasch model for dichotomous data, the probability of the outcome is given by:

where β_n the ability of person n and δ_i the difficulty of item i.

The model needs to be radically enhanced to take into account an assessment of performance that is not dichotomous and feed into the interaction between two macro factors:

- Task Complexity (TC): summarising all factors contributing to physical and mental workload requirements for execution of a given operative task.
- Human Capability (HC): summarising the skills, training and experience of the people facing the tasks, representing a synthesis of their physical and cognitive abilities to verify whether or not they match the task requirements.

Task complexity can be evaluated as a mathematical construct considering also the compound effects of two main factors: "Mental Workload Demands" (MW) and, where relevant, "Physical Workload Demands" (PW), both associated to an operator task. Recent sensorised EEG experimental studies have shown that the simultaneous executions of tasks, whether physical or cognitive, tends to increase cognitive demands for the human brain (Mijović, 2017).

Similarly then, operator capability should be estimated on the basis of the operators' set of cognitive capabilities and physical conditions. A regression model was used to fit a dataset collected in R. The model and the preliminary results are discussed in chapter 3 of the present paper.

3 THE CASE STUDY: SUMMARY OF THE DATA COLLECTED AND THE TERM MODEL AS APPLIED

3.1 *The case study and the data collection plan*

The setting and focus of this study is a large electronic manufacturing facility in the south of Ireland, which prepares and distributes technology goods to both national and global customers. In the dispatching unit of the facility, operators are provided with work stations and conveyors to prepare the products for dispatch and shipment through pack out procedures. The aim of this study was to carry out an observational, empirical study on the existing human errors in the dispatching department of the facility in a subsequent phase the study also lead to the identification of suitable approaches to reduce and/or eliminate such errors.

Two primary workstations were the focus of the assessment of the project, namely the conveyor line and another packaging workstation called the POD cell. To examine these workstations, an overview of the existing error rate at the conveyor line was required to be used as a benchmark against other workstations in the facility and to identify any possible improvements. As a means of comparison, the error rates for nine control workstations from within the manufacturing facility were acquired to facilitate data analysis and interpretation.

Error rates for both the control and non-control workstations were calculated in the same manner. Records were filtered from 1st December 2016 to 31st March 2017 for all workstations to retrieve the information for the calculations. This four months timeframe was deemed adequate due to the large number of products passing through the workstations. We considered only errors classified as stemming from a human related cause.

The human error rates were calculated using the following formula:

$$\text{Number of Human Errors/the opportunity for error} \quad (2)$$

where the number of human errors were the errors recorded or captured due to a human cause

While the opportunities for error were the total output at the workstation i.e. number of processed units

For the pod and the conveyor, to attain the number of human errors, data relating to MWDs (missing, wrong or damaged) goods was collected within the four month period from 01/12/16 to 31/03/17. The MWDs originate from customer complaints or returned goods following disparities

from the sales orders or damaged goods. MWDs can be slow information to capture, due to the possible time lapse between the shipment of an order, the start of use of the product by the customer and the identification of an error. MWDs may be reported some months after a product was shipped, however due to the nature of the timeframe selected, it was deemed appropriate that by the completion of the project, the number of MWDs recorded for that time frame would be sufficient. The opportunity for error was derived from the total output at the workstations within the four months period from the beginning of December 2016 to the end of March 2017.

For the control workstations, the numbers of human errors were retrieved from an online software platform within the four months period outlined above. The platform is used to record both the total output at the workstations and the number of errors recorded. The platform records errors with varying root causes through a classification system, many of which are not of a human nature. Twenty-seven classifications were deemed suitable for inclusion for the human errors recorded.

In the control workstations, when an error has occurred, the operator or technician is forced to input an error report at the time of the error occurring detailing the source of the error i.e. human, equipment, technical. The process cannot continue until an error report has been submitted. Due to this, the error reports recorded in the system can be regarded as representative of the total number of errors occurring during the timeframe. When an error is recorded, users are prompted to categorise the error under a variety of descriptions. The categories can include aspects of technology or equipment failure, and not all were relevant for inclusion in the error rate calculation.

3.2 *The observation and questionnaire protocol used for the wider case study*

Members of staff who work closely with the workstations involved in the project and the control workstations were invited to complete questionnaires to assess their opinions relating to:

- The importance of skills at different workstations
- Skills rating of individual operators
- Job satisfaction/motivation
- Mental workload requirements
- Physical workload requirements
- Perceived task complexity

Two questionnaires were prepared with one for supervisors, group leaders and process engineers, and a second questionnaire for operators and

technicians. Questionnaires were broken up in this fashion in order to capture observable variables from the supervisors/management and the individual subjective opinions of the operators. There was a difference in the type and volume of questions in the questionnaires, as the supervisor/group leader questionnaires asked two different things:

- Asked supervisors role participants to rate the skills of operators under their supervision
- Asked all participants to rate the skills requirement to complete work at the workstations

The questionnaires were completed by the employees of all eleven workstations and their supervisors leading to a total of 149 employees completing the questionnaire (100% response rate).

Participants were asked to rate their answers on a 10-point Likert Scale, with one meaning low and ten meaning high. Questionnaires were used to measure the mental and physical workload, worker skills, job satisfaction (motivation) and the perceived task complexity for operators, supervisors, group leaders and process engineers. As different duties and tasks require certain skills (e.g. manual skills, memory), practical training and underpinning knowledge, the questionnaire was designed to capture information relating to the following areas:

Mental Workload Requirements

- Need to cope with pace
- Variance of product
- Recognition requirements
- Load due to quality of coordination
- Requirement for training/experience
- Requirements for human machine interface (HMI)

Physical Workload Requirements

- Ergonomic score (REBA Assessment)
- Dexterity requirements/manual skills

- Adherence to procedure
- Reliance on automation

Job Satisfaction/Motivation

- Motivation e.g. satisfaction, meaningfulness

Worker Skills

- Memory
- Decision-making
- Recognition
- Coordination/communication—teamwork
- Coping with pace
- Experience
- Dexterity/manual skills
- Physical resilience
- Adherence to procedure

Perceived Task Complexity

- How mentally demanding are the tasks
- How physically demanding are the tasks
- How complex is this task

The error rate for all eleven workstations has been calculated and is outlined in [Table 1](#)

Data collection involved a rich integration of data from many sources, acquired observationally or through documented information. There were four primary sources of data:

1. The questionnaires outlined above. The data collected would facilitate the assessment of the relationship between the task complexity (mental workload requirements, physical workload requirements) and the worker capability (cognitive skills, physical skills).
2. Focus groups were conducted to understand the process and procedures at the workstations and aspects of the workstations that would benefit from redesign.
3. Key Performance Indicators (KPI's) were gathered for information relating to:

Table 1. Error rate dataset collected for each workstation.

| Workstation No | No of human errors | Opportunity for errors i.e. total output | No of operators | Error per 1000pc |
|----------------|--------------------|--|-----------------|------------------|
| 1 Pod | 0 | 747 | 19 | 0.01 |
| 2 Conveyor | 14 | 8,913 | 19 | 1.5 |
| 3 Control A | 3 | 12,055 | 19 | 0.2 |
| 4 Control B | 1 | 1,359 | 19 | 0.7 |
| 5 Control C | 44 | 221 | 2 | 203.6 |
| 6 Control D | 93 | 221 | 2 | 425 |
| 7 Control E | 28 | 3,971 | 5 | 7 |
| 8 Control F | 81 | 3,971 | 7 | 20.3 |
| 9 Control G | 368 | 5,402 | 5 | 68.1 |
| 10 Control H | 107 | 5,402 | 0.019 | 19.8 |
| 11 Control I | 133 | 5,402 | 0.0246 | 24.6 |

- The actual time at the workstation (productivity KPI)
 - The number of quality issues due to human error (quality KPI)
4. Error rates for the workstations were formulated to provide insight into the rate of human error and its resulting quality effects on the workstation end products.
Videos/Pictures

In order to capture and assess information regarding the routine activities and work patterns of staff in the facility, video recordings and photographs were taken as an observational method of data collection. The videos were used to:

- Measure the amount of time the entire task took to complete
- Measure the amount of time an aspect of the task took to complete e.g. closing with sellotape
- Compare the procedure completed to the actual projected procedure for the completion of actions
- Task analysis using Video TimerPro software to break down the tasks required of the operator to complete
- The photographs and video recordings were used to:
- Provide a basis for the Ergonomic Risk Assessment method used
- Compare comparable tasks completed at alternate work stations

For the first part of the regression model, an assessment of task complexity was conducted. The data gathered was evaluated on the basis of Task Complexity with a linear regression model. In order to complete this evaluation, a task complexity index was applied, namely:

Task Complexity index = a (Memory req.) + b (recognition req) + c (coordination req.) + d (cope with pace req) + e (Experience req) + f (Resilience req.) + g (adherence to procedure req.)

The Correlation matrix obtained for the element used for the regression to evaluate task complexity obtained in the statistical software R are shown in Figure 1.

Figure 2 reports the preliminary results of the linear regression model used to predict task complexity in R.

The model indicates that the parameters used to estimate task complexity in the linear regression are quite significant. They predict task complexity with a Standard error of 0.2991 on 36 degrees of freedom. The adjusted R squared obtained is 0.93996 and the F statistics on 36 Degrees of freedom is 96.52, with a p value of 2.2 e-16. Therefore

| | M | D. | R | C | C | .. | E | Dx | P | A |
|------------------------|-------------|----|---|---|---|----|---|----|---|-----|
| Memory.Capacity | 1 | | | | | | | | | |
| Decision.Making | , 1 | | | | | | | | | |
| Recognition | . . 1 | | | | | | | | | |
| Coord.Comm | , , 1 | | | | | | | | | |
| Cope.with.Pace | | | | | 1 | | | | | |
| Experience | + , . . 1 | | | | | | | | | |
| Dexterity | | | | | | | | | 1 | |
| Physical.Resilience | 1 | | | | | | | | | |
| Adherence_to_procedure | | | | | | | | | | . 1 |

Figure 1. Correlation matrix evaluated for the element used for the regression to evaluate task complexity.

| Coefficients: | | | | |
|---------------------------------|----------|------------|---------|--------------|
| | Estimate | Std. Error | t value | Pr(> t) |
| (Intercept) | -0.62930 | 0.36179 | -1.739 | 0.090509 . |
| taskreq\$Memory.Capacity | 0.19195 | 0.04300 | 4.464 | 7.63e-05 *** |
| taskreq\$Recognition | 0.18100 | 0.02609 | 6.938 | 3.96e-08 *** |
| taskreq\$Coord.Comm | 0.10559 | 0.03672 | 2.876 | 0.006733 ** |
| taskreq\$Cope.with.Pace | 0.14421 | 0.04006 | 3.600 | 0.000951 *** |
| taskreq\$Experience | 0.13699 | 0.04973 | 2.754 | 0.009161 ** |
| taskreq\$Physical.Resilience | 0.07111 | 0.04113 | 1.729 | 0.092341 . |
| taskreq\$Adherence_to_procedure | 0.19745 | 0.03367 | 5.865 | 1.05e-06 *** |

Figure 2. Results obtained from R to evaluate the relevance for the coefficient used to estimate task complexity.

the linear regression model to estimate task complexity seems to deliver significant results.

For the second part of the model, an estimation of the error occurrence of each workstation considering task complexity and operator capability was conducted. The use of the Rasch model with the dataset gathered was not possible as for the Rasch model the output needed to be a binary success or failure for each individual task. This was a type of data which was not able to be collected. Due to this, a generalised linear regression with a Poisson model, which was still based on the assumption that Human Performance can be represented as directly dependent from two macro-factors of task complexity and human capability, was used (see formula 3).

$$\lambda_i = e^{\beta_0 + \beta_1 x_1 - \beta_2 x_2 + \epsilon_i} = e^{\eta_i} \rightarrow \log \lambda_i = \eta_i \quad (3)$$

where λ_i is the amount of error recorded, x_1 is task complexity and x_2 is operator skill level/capacity. The results obtained in R suggest that both elements are significant in predicting error occurrence, as shown in Figure 3.

The likelihood ratio test results confirmed the meaningfulness of the significance for the parameter chosen for estimating the error rate with this model, as shown in Figure 4.

However the limited data set and that the estimates of skill rating were gathered done using a subjective rating. Therefore the model could be

Table 2. Summary of data collected and revised for each workstation used in the regression model.

| Id workstation | Average skills recorded | Task complexity | Errors_on_10000 parts |
|----------------|-------------------------|-----------------|-----------------------|
| 1 Pod | 6.45 | 7.4 | 1 |
| 2 Conveyor | 6.45 | 7.28 | 15 |
| 3 Control A | 6.45 | 6.8 | 2 |
| 4 Control B | 6.45 | 6.8 | 7 |
| 5 Control C | 7.18 | 8 | 2036 |
| 6 Control D | 7.23 | 9 | 4250 |
| 7 Control E | 7.09 | 7.57 | 70 |
| 8 Control F | 7.86 | 7.33 | 203 |
| 9 Control G | 5.33 | 6.33 | 681 |
| 10 Control H | 6.27 | 6.4 | 190 |
| 11 Control I | 7.83 | 6.88 | 246 |

```
Call:
glm(formula = errorrate$errors_on_10000parts ~ errorrate$average.skills +
     (errorrate$task.complexity), family = poisson, data = errorrate)

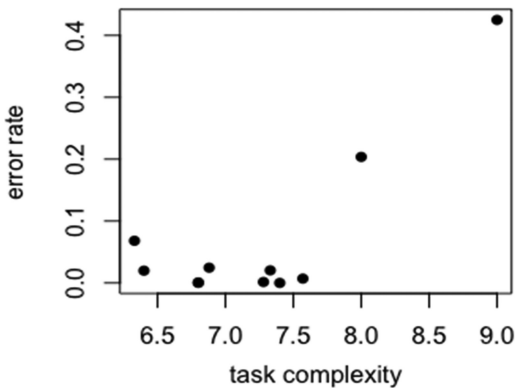
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-31.533  -21.480   -1.975    8.522   31.609

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.41163    0.19002  -12.69  <2e-16 ***
errorrate$average.skills -0.47110    0.03457  -13.63  <2e-16 ***
errorrate$task.complexity  1.57751    0.01710   92.24  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 3. Results of the analysis run in R for the generalised Poisson linear model.

```
Model 1: errorrate$error.rate ~ errorrate$average.skills + errorrate$task.complexity
Model 2: errorrate$error.rate ~ 1
#Df LogLik Df Chisq Pr(>Chisq)
1  4 33.667
2  2 28.272 -2 10.79  0.004539 **
```

Figure 4. Results of the analysis run in R for the likelihood ratio test for the generalised Poisson linear model.



the incidence of human related errors across the workstations examined.

The results may be limited by the four month time frame for which human errors were considered. However results shown that task complexity can be significantly predicted starting from the variables observed in the case study.

The TERM model used (the Poisson generalised linear regression) also suggests that both task complexity and operator's skill are valid predictors of error occurrence in a workstation. It is maybe also possible that while task complexity increases a corresponding linear increase in worker skills and capability is not able to sufficiently compensate for the increased complexity.

REFERENCES

- Harris, D. (2006). The influence of human factors on operational efficiency. *Aircraft Engineering and Aerospace Technology*, 78(1), pp. 20–25.
- Jo, S., Myung, R. and Yoon, D. (2012). Quantitative prediction of mental workload with the ACT-R cognitive architecture. *International Journal of Industrial Ergonomics*, 42(4), pp. 359–370.
- Karl, R. and Karl, M. (2012). Adverse Events: Root Causes and Latent Factors. *Surgical Clinics of North America*, 92(1), pp. 89–100.
- Kostina, M., Karaulova, T., Sahnó, J. and Maleki, M. (2012). Reliability estimation for manufacturing processes. *Journal of Achievements in Materials and Manufacturing Engineering*.
- Leva, M.C., Ciarapica Alunni, C., Demichela, M. & Allemandi, G. (2016) Addressing human performance in automotive industry: identifying main drivers of human reliability. Irish Ergonomic Society Conference 2016.
- Mijovic B. 2017. How hard is walking. blog <https://blog.mbraintrain.com/@boggisha> (last accessed 20/09/2017).
- Miller, C. and Parasuraman, R. (2007). *Designing for Flexible Interaction Between Humans and Automation: Delegation Interfaces for Supervisory Control*. *Human Factors*, 49(1), pp. 57–75.
- Morgeson, F., DeRue, D. and Karam, E. (2010). Leadership in Teams: A Functional Approach to Understanding Leadership Structures and Processes. *Journal of Management*, 36(1), pp. 5–39.
- Rasch G. (1980) *Probabilistic Model for some Intelligence and Attainment Tests*. University of Chicago Press. Chicago.
- Sheridan, T. (2008). Risk, Human Error, and System Resilience: Fundamental Ideas. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3).
- Tooby, J. and Cosmides, L. (1990). On the Universality of Human Nature and the Uniqueness of the Individual: The Role of Genetics and Adaptation. *Journal of Personality*, 58(1), pp. 17–67.