AC 2011-2173: EVALUATION OF RISK IN EARLY DESIGN’S USABILITY IN FAILURE ANALYSIS INSTRUCTION

Katie Grantham, Missouri University of Science & Technology

Ryan Arlitt, Missouri University of Science and Technology

©American Society for Engineering Education, 2011
When engineers retire, they take their expert knowledge with them. Preservation of this expert knowledge in a usable form is beneficial for the advancement of any engineering field. Risk in Early Design (RED) is one method for preserving expert risk analysis knowledge. The purpose of this paper is to examine the usability of RED when incorporated with a hybrid problem-based and just-in-time inductive teaching method for failure analysis instruction. This test was conducted in a sophomore level lab class at a college in the Midwest in the fall of 2010. The lab was designed to assist in teaching mechanics of materials, and was composed of approximately 200 students. A questionnaire was used to determine the usability and perception of RED. The questionnaire helped to identify areas where the application was hindering student performance. Function selection, in-application instructions, and risk report download and interpretation were identified as areas with poor usability. Initial improvements to the interface were made based upon feedback from the questionnaire.
1. INTRODUCTION

The goal of this research project is to test usability of the Risk in Early Design (RED) application when used as an expert knowledge source for tasks previously thought to require engineering experience. As technology progresses, it is critical that educational efforts focus on preparing students to build on the new developments, rather than continuously teaching them to “reinvent the wheel.” The teaching of new technology is not limited to the integration of novel hardware and software into the engineering curriculum. It is also important to teach the next generation of engineers decision-making skills that build upon the current level of expertise in the workforce. Therefore, it is imperative that new technology also be used to prepare the engineers of tomorrow to analyze and understand engineering systems by conveying the knowledge associated with years of corporate experience during their undergraduate studies.

The teaching strategy presented in this paper is a hybrid problem-based and just-in-time inductive teaching method. The cornerstone for the method is the existence of a knowledgebase of “engineering experience.” In this case, the Risk in Early Design (RED) knowledgebase was developed as part of a risk assessment project that leveraged historical failure data in electromechanical systems to predict and prevent such failures in the design of new electromechanical systems [1]. Student satisfaction with RED’s usability was measured in a case study designed to leverage RED as a teaching tool. This evaluation took place in the 2010 fall semester at a university in the Midwest.
2. SCOPE

The National Academy of Forensic Engineers (NAFE) [2] defines forensic engineering as “the application of the art and science of engineering in matters which are in, or may possibly relate to, the jurisprudence system, inclusive of alternative dispute resolution.” NAFE also asserts that “the practice of licensed Professional Engineers as Forensic Engineers is important for the protection of the public health, safety and welfare.” Preliminary interviews of engineering firms have demonstrated the need for safety engineers in industry. Currently, there is only one forensic engineering program at the graduate level [3]; none exist at the bachelor’s degree level [4]. Statistics compiled by the American Society for Engineering Education (ASEE) for the 2005-06 academic year indicate that engineering graduation and enrollment rates at U.S. universities are not reflecting the country’s increasing demand for engineering talent [5]. One reason for the gap may be that the traditional rigid engineering curriculum has not adapted to the diverse needs of a quickly changing technological world, such as the advances in the forensic engineering profession.

This paper presents research that was conducted to investigate the use of knowledgebase guided teaching strategies to enable courses with “engineering experience” as a prerequisite to be taught at the undergraduate level. This research will contribute to the formation of an undergraduate forensic engineering program that will leverage the industrial need and media popularity of forensics. This paper’s contribution to the creation of a forensics program is in the formation and verification of reverse failure analysis coursework through improvements to the RED application.
3. **RISK IN EARLY DESIGN (RED)**

Risk in Early Design (RED) is a probabilistic risk assessment method that leverages historical failure data to provide failure data based upon the functions that a system must perform. This is accomplished using a series of matrices that contain historical data on component function and failures, along with an algorithm that presents failure modes, likelihoods, and severities for user selected functions. RED uses simple heuristics and mathematics to communicate cataloged historical product-specific risks as early as the conceptual design phase. Given the functions of a design, RED outputs potential risks based on historical failure data [6].

3.1 **RED DATABASE POPULATION**

The RED database draws information from three sources in order to store failure information: functional models, bills of materials, and failure reports. Failure reports provide documented cases of system failures, cataloged using a failure mode taxonomy [7] in order to standardize the terminology used in the database. Bills of materials provide components found in those systems. The components are cataloged using a component basis [8]. Functional models, consisting of material, energy, and signal flows connecting function blocks, provide the functionality of the failed systems. Similarly to the other two elements used in RED database population, functional models follow a functional basis terminology. The functional basis terminology standardizes the language to describe product function, leading to meaningful and repeatable function representations [9].

Functions and components are drawn from these sources to populate the function-component (EC) matrix. This matrix shows which components have historically
accomplished which functions, using a 1 to denote a relationship and a 0 to denote no relationship. For example, function “A” in the EC matrix Figure 3.1 has been accomplished by components 2, 4, 5, and 7. The component-failure (CF) matrix shows how often each component has failed by each failure mode. In the CF matrix shown in Figure 3.1, component 1 has failed by failure mode “b” twice, failure mode “c” four times, and failure mode “e” once. The EC and CF matrices are multiplied to produce the function-failure (EF) matrix, which shows how often each function has failed by each failure mode. For example, function “A” in the EF matrix in Figure 3.1 has failed by failure mode “a” ten times. The teaching strategy presented in this research uses an existing RED database of electromechanical failures to support RED operations.

Figure 3.1. RED Database Population Sources and Matrices
3.2 RED RISK ANALYSIS PROCESS

The results from prior work [10] on developing the RED method have yielded a process for identifying and assessing risk during the conceptual design phase. This risk identification method was tested in the university’s mechanics of materials lab to determine if it can successfully provide “engineering experience” from which the students can draw on to initiate their failure investigations and classifications. The steps for using RED to guide a failure analysis investigation, shown in Figure 3.2, are: (1) generate the functional model of the failed part, (2) select the relevant functions from the historical failure database, and (3) perform risk calculations. The results displayed on the fever chart and the related risk report present students with a ranking of failures that occurred in similar components. In the example in Figure 3.2, the fever chart shows the number of failures that have occurred in the database for the selected functions at each likelihood and consequence pair. Here, five risks have occurred at a consequence of one and a likelihood of two, one risk has occurred at a consequence of four and a likelihood of one, and two risks have occurred at a consequence of three and a likelihood of four. The students used type of information to guide their investigations.
Figure 3.2. RED Process for Failure Investigation Guidance
4. **RED AS A TEACHING AID**

4.1 **TEACHING STRATEGY**

The problem-based teaching method, as its name implies, confronts students with a poorly defined, real world problem. Students work in teams to identify learning needs and develop solutions to the problem [11]. Problem-based learning has been shown to positively affect knowledge retention and skill development [12].

Just-in-time teaching typically consists of web-based preliminary exercises that the instructor uses to adjust lessons just before class based on student responses. Online enrichment pages and stand-alone instructional material support the in-class lesson. Just-in-time teaching promotes increased study outside of class and increased student-instructor interaction during class [13]. Just-in-time teaching has been assessed in physics instruction using the Force Concept Inventory, and has shown normalized student gains between 35% and 40% [11].

Both problem-based and just-in-time teaching are inductive teaching methods highlighted by Prince and Felder [11]. The authors describe inductive teaching as any teaching method that presents students with specific information that creates a need for more general facts or principles. Often this is accomplished by tasking the students with interpreting some specific data that requires these more general principles. This is highlighted as directly opposing the traditionally used deductive teaching, in which instructors present general principles and then show examples to reinforce them. The authors state that people are most strongly motivated to learn when they perceive a need to know, and that inductive teaching and learning are preferable methods of achieving this effect.
The teaching method applied in the experiment utilizes failed components, such as a bolt from a bridge, as an enabler for problem-based teaching. The students are presented with the problem of determining how the component failed, creating a need to know more general principles about failure analysis. The information that the students gain from RED is obtained just-in-time to help them analyze these failed components. In this sense, this teaching method does not conform with traditional just-in-time teaching. Whereas traditional just-in-time teaching relies on the instructor to adjust the learning material based on preliminary student feedback, in this case guidance in learning these more general failure analysis principles is provided by RED. Upon completing the lab, students should have learned general failure analysis principles based on their experiences with the specific component analyzed.

Additionally, the mechanics of materials lab course where this method was tested currently utilizes enrichment materials on its website in the form of related information that shows the materials’ real-world relevance.

4.2 RED AND FAILURE ANALYSIS INSTRUCTION

For an example of how RED would typically be used, consider the situation of students in a problem-based learning exercise who were presented with a failed shaft and tasked with identifying the failure mode. Having extremely limited “engineering experience” from which to initiate their investigation, the students would use the RED method. First, the students would identify the functions of the shaft, and produce a functional model similar to the one found in Figure 4.1.
Next the students would enter its functions into the RED software. Sample output of the software is shown in Table 4.1. The results show that the functions transfer mechanical energy, secure solid, export mechanical energy, and import mechanical energy are most at risk of failure due to high cycle fatigue. Continuing down the report toward functions with lower severity, the solid and mechanical energy flows are also at risk due to brittle fracture and stress corrosion. These results indicate that the first course of action taken by the students would be to determine if the physical characteristics of the failed part and failure environment match with the most common type of failures.
provided. Continuing with this example, if the shaft experienced a significant amount of cycles and there was a physical break in the component, then the students could focus their analysis on determining if the failure was caused by high cycle fatigue. If it does not meet the criteria for high cycle fatigue, students would move down to brittle fracture and then stress corrosion. In this case, the shaft failed by brittle fracture. This teaching strategy will be assessed, and if found successful will promote more use of similar concepts to be incorporated into undergraduate curricula.

Table 4.1. Truncated RED Results for Shaft

<table>
<thead>
<tr>
<th>Severity</th>
<th>Function</th>
<th>Failure Mode</th>
<th>Likelihood</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Transfer Mechanical Energy</td>
<td>High Cycle Fatigue</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>Secure Solid</td>
<td>High Cycle Fatigue</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>Export Mechanical Energy</td>
<td>High Cycle Fatigue</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>Import Mechanical Energy</td>
<td>High Cycle Fatigue</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>Export Solid</td>
<td>High Cycle Fatigue</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>Import Solid</td>
<td>High Cycle Fatigue</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>Transfer Mechanical Energy</td>
<td>Brittle Fracture</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>Secure Solid</td>
<td>Brittle Fracture</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Med</td>
<td>Export Solid</td>
<td>Stress Corrosion</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Med</td>
<td>Export Mechanical Energy</td>
<td>Stress Corrosion</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
In the context of this research, RED is presented to students as a black box. Students were provided with a functional model of their failed component, such as the one in Figure 4.1, thus removing the need for the students to be familiar with functional modeling to perform the exercise. This allowed a greater sample size of students who were able to generate RED output of potential failure modes of a failed component. Prior to performing the experiment, functional models were generated for all of the components that would be used in the lab.
5. EXPERIMENTAL DESIGN

The Midwest university’s mechanics of materials lab is designed to assist in teaching mechanics of materials, in which students learn about topics such as material properties, strain testing, and testing machines [14]. Students gain hands-on experience in the lab to reinforce learning of lecture topics.

In the Failure and Fully Plastic Action Lab, students “look at the definition of failure, failure theories, and real-life examples of failed components.” Additionally, students “investigate failed components, estimate what caused the failure, and propose a remedy” [15]. These aspects of the lab make it a good fit for testing RED as a teaching method.

Evaluation of RED’s usability was performed as part of a larger experiment within the Failure and Fully Plastic Action Lab [15] in the university’s mechanics of materials lab class. The experiment was designed to fit within the existing structure of the class. At the beginning of the semester, students in each section formed groups of their own choosing. These groups were typically three to four students in size. The ten lab sections were divided into an experimental group and a control group. Three sections met on Monday, three on Tuesday, two on Wednesday, and two on Thursday. The Tuesday and Thursday sections were selected as the experimental group, because one of the instructors in three of those five sections had experience with RED. This was done to mitigate the risk of any unforeseen issues with the RED deployment that might prevent students from using it. The experimental group contained a total of 101 students divided into 34 groups, and the control group contained a total of 96 students divided into 33 groups. The
experimental group used the RED tool in addition to performing the lab, and the control group performed the lab without the tool.

Prior to performing the lab, each group selected a failed component to analyze from the pool of 17 available components. The students were each issued a failure mode taxonomy handout and a preliminary assessment form requesting that the student determine the failure mode of the selected failed component. The failure mode taxonomy provides the failure modes, along with a “primary identifier” and a definition of the failure mode, in order to aid failure mode identification. The primary identifier is the highest level of classification in the failure mode taxonomy, and helps to narrow one’s focus to the appropriate failure mode. For instance, the primary identifier “Corrosion (Material deterioration due to chemical or electrochemical interaction with the environment)” contains twelve corrosion failure modes [7].

After completing preliminary assessments, students performed the lab. Lab activities included detailed observations of the failed component. Outside of class, the students in the experimental group ran a RED analysis on their failed item and saved the risk report to aid them in answering lab questions. These students were required to submit the risk report with their lab report to ensure that they performed the RED analysis.

All students answered questions regarding the failure and its prevention using a post-lab failure assessment form. Post-lab assessments, lab reports, and a questionnaire regarding RED were gathered digitally using an online tool. For this study, student perception of RED’s usefulness and usability were gathered from the experimental group using a questionnaire.
6. RESULTS AND DISCUSSION

A survey was designed to measure the usability of the RED tool implementation, student perception of their own performance in the case study, and the usefulness of RED in the case study. The survey consisted of 13 questions on a Likert scale and two open-ended questions. The survey was deployed through the Blackboard web-based course management system after students completed the lab. Blackboard’s capabilities include allowing students to download and turn in assignments and surveys online. Students were incentivized to complete the survey with bonus points, and there were 80 respondents out of a possible 101 in the experimental group.

Questionnaires were selected because they can be used to collect a large amount of data using few resources. Questions pertaining to the system’s usability included questions targeted to specific areas of usability as well as open ended questions designed to uncover problems that may have been missed by tool evaluators. Questions dealing with specific areas of usability were framed after a set of Likert scale and open-ended questions designed to assess the usability of a software system, provided by Dix et al. [16]. Six of the Likert scale questions asked students to rank their level of agreement with how well the RED application addressed specific areas of usability, such as feedback, ease of navigation, and ease of access. These usability questions are seen below.

Please answer the following questions based on the following ratings:

1 - I strongly disagree
2 - I disagree
3 - I neither agree nor disagree
4 - I agree
5 - I strongly agree

1. The RED application tells me what to do at each step in the risk identification process.
   1 2 3 4 5

2. It is easy to recover from mistakes I make while using the RED application.
   1 2 3 4 5

3. It is easy to get help within the RED application when needed.
   1 2 3 4 5

4. The RED application always gives me feedback to tell me what it is doing.
   1 2 3 4 5

5. It is easy to navigate through the RED application.
   1 2 3 4 5

6. The RED application was easy to access.
   1 2 3 4 5

Open-Ended Questions

1. What did you dislike about the RED application? Please suggest improvements.
2. What did you like about the RED application?

By comparing these responses for each of these usability aspects, a prioritization for addressing each was obtained. The open-ended questions asked students for likes and dislikes about the RED application, in order to uncover unanticipated problems with the usability and with RED in general. Table 7.1 shows the means of those responses, ranked from highest to lowest level of agreement. The ranking in this table provides a guide as to which aspects of the RED software possess the lowest degree of usability. Usability aspects that received lower mean scores may reflect lower levels of satisfaction with that aspect of the usability. Based on these mean scores for each response, the survey suggests the following order of importance for usability improvements: provide feedback, provide help and guidance within the application, improve navigation, improve error recovery, and improve accessibility.
The remaining questions were designed to assess the student perception of RED’s helpfulness and their own performance in the exercise. The responses to these questions are summarized in Table 6.2, and provide a baseline for comparison when improvements are made to the instruction technique used in the case study. Responses to these questions indicate that students were confident in their assessments, while confidence in RED’s ability to aid in failure assessment was less pronounced. After improvements are made to this teaching strategy in a future semester, this survey will be administered again to determine whether the improvements were successful.
Two open-ended questions regarding the students’ likes and dislikes about RED were asked in order to identify unanticipated usability problems that were not otherwise addressed by the survey. Responses to those questions were clustered into categories with responses having similar themes. After those categories were formed, they were named based on the theme associated with the cluster. Students who took the survey but did not respond to the open-ended question were placed in the “No Response” cluster. Multi-part
responses that fit into multiple categories were counted once in each of those categories. For example, consider the following response to the question about dislikes:

“The data received is slightly difficult to sift through. Possibly organize the data in a manner that will ease in finding what exactly one is looking for. Make selecting multiple functions easier to do.”

This response contains two themes. First, the student indicates that they had difficulty using the RED report. Second, the student indicates difficulty with the user interface. This response was split into two responses and placed into groups with similar responses. When all clusters were formed, these two clusters were named “Report Clarity” and “Interface Clarity” respectively.

Student “likes,” seen in Figure 6.1, clustered around three main categories. In order of frequency, students commonly liked RED’s ease of use, thought it was useful in the exercise, and liked the large amount of information provided. In general, students felt that the instructions and procedures involved in producing the RED output were easy to understand. Additionally, many students indicated that RED was useful in determining the failure mode of the component. Similarly, students liked the large quantity of information provided by the application.
Student “dislikes,” seen in Figure 6.2, also clustered around three main categories. Interface clarity, meaning the student had issues with performing the desired tasks due to the human interface, was mentioned the most. Report clarity, meaning that students had issues understanding the risk report, was also mentioned frequently. The report clarity cluster included difficulties choosing the correct type of report to download, difficulties formatting that report into a readable one, and difficulties interpreting what the results meant. A significant group of students also stated that RED was not useful in determining the failure mode of their failed component. This could be attributed to difficulties interpreting the report or student confidence in their initial answer. Several students also mentioned having access difficulties and problems understanding the functional model.
The disparity between having a high ease of use and poor interface clarity might be explained by the tutorial provided with the RED application. While students felt that RED was easy to use, it was likely due to the step-by-step instructions provided in the tutorial. The disparity between students who thought that RED was useful and those who did not could be explained by a perception that RED report interpretation does not require a human-in-the-loop. In order to be useful in this context, RED needs a human to select a failure mode that fits the specific case.

Based on the survey data, several improvements were identified that can increase the usability of the RED tool. A map graphic of where the user is in the RED process,
accompanied by instructions and provided on every page of the application, should prevent users from getting lost or stuck by providing feedback and navigation assistance. A welcome page with a basic overview and instructions on how to use the application, as well as an easily accessible link to the RED tutorial, should improve the amount of help and guidance available. Retaining function selection after the user submits would allow the user to make changes more easily if a mistake is identified, improving error recovery. Finally, students identified the function selection interface as difficult to use. Changing the scroll box to a different interface would reduce the time required to search for and double check function selections.
7. INTERFACE IMPROVEMENTS

Several improvements, which address some of the usability issues uncovered in the case study, have been made to the RED interface. These improvements are still in progress, but for now include clarifications to the heuristic selections and a greatly improved function selection interface.

7.1 HEURISTIC CLARIFICATION

A lack of instructions within the RED application was identified as one area for improvement. As a step toward providing better guidance, clarifications have been added to the heuristic selection step of the RED process as seen in Figure 7.1. This will indicate to the user the reason for making this selection as well as giving a better understanding of how their choices will affect the risk calculations.

Choose your Product and Design Choices.

Human Centric, System Level
Returns average likelihoods and high consequence.

Human Centric, Subsystem Level
Provides the most caution, returns high likelihoods and consequences.

Unmanned, System Level
Provides the least caution, returns average likelihoods and consequences.

Unmanned, Subsystem Level
Returns high likelihoods and average consequences.

Figure 7.1. Heuristic Clarifications
7.2 FUNCTION SELECTION INTERFACE

Figure 7.2 depicts the function selection interface tested in the case study. This interface requires users to scroll through a list of every function that appears in the RED database and individually select functions relevant to the user’s system. Survey responses revealed user issues verifying that all of the desired functions were selected. Users also experienced difficulty determining if their desired function appeared in the RED database. A new interface was developed that addresses these issues.

Choose the Functions to Use.

| Available Radioactive Energy |  |
| Change Control Signal |  |
| Change Control Signal To Control Signal |  |
| Change Electrical Energy |  |
| Change Electromagnetic Energy |  |
| Change Gas |  |
| Change Hydraulic Energy |  |
| Change Liquid |  |
| Change Mechanical Energy |  |
| Change Pneumatic Energy |  |

Choices:  

Submit

Figure 7.2. Function Selection Interface Tested in Case Study

Figure 7.3 depicts the updated function selection interface. Upon first encountering the screen, the “Available functions” box on the left displays every function that has appeared in the RED database. Users can scroll through this list, select desired functions, and press the right arrow button to add to the list of “Chosen functions.” Users also have
the option of typing the function into text box above the function list. This action dynamically filters the list of available functions to reflect what the user has typed. The user can press enter to add the first function on the list, or choose a function from filtered list. Upon selecting all desired functions, the user can review the list of chosen functions on the right before clicking the button to generate a report. This increased clarity will decrease the effort required by users to select all desired functions and verify that function selection before generating a risk report. The redundant methods for accomplishing the same task may improve user speed and accuracy.

Choose the Functions to Use.

<table>
<thead>
<tr>
<th>Available Functions</th>
<th>Chosen Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Change Mechanical Energy}$</td>
<td>$\text{Change Mechanical Energy}$</td>
</tr>
<tr>
<td>$\text{Convert Human Energy To Mechanical Energy}$</td>
<td>$\text{Convert Human Energy To Mechanical Energy}$</td>
</tr>
<tr>
<td>$\text{Import Human Energy}$</td>
<td>$\text{Import Human Energy}$</td>
</tr>
</tbody>
</table>

Choose all
Hold control (or command on Mac) to select multiple functions. Hold shift to select areas.

Generate Report

Figure 7.3. Updated Function Selection Interface
8. CONCLUSIONS

This research seeks to take a step toward the verification and validation of the Risk in Early Design tool and methodology, and to improve RED’s utility as an educational tool. Based on survey data, areas for improvement and important usability aspects of the RED tool were identified. Initial improvements were made to the RED interface, and further improvements will increase the utility of RED. After these improvements are made, repeating the case study and survey will allow verification of the interface changes. Positive changes to the usability will promote better learning by reducing the barrier between the tool’s interface and the information that the tool is conveying. These improvements should increase RED’s utility as an expert knowledge preservation device.
9. FUTURE WORK

Future work for this research can be summarized in three parts. First, improvements to the RED application’s usability will allow the application to more effectively accomplish its goals through improved user interaction. Second, improvements to RED’s risk reports will enable users to more easily interpret risks. Third, work will be done to analyze the effectiveness of the hybrid problem-based just-in-time teaching method based on student failure assessments gathered in the experiment. These improvements and analyses will enable a future case study to assess the benefits of an improved RED interface.
10. REFERENCES


9. Hirtz et al. A Functional Basis for Engineering Design: Reconciling and Evolving Previous Efforts. NIST Functional Note 1447


