The Effect of the Routine/Non-Routine Nature of Tasks on Serious Accident Rates

Bernard McGarvey[=] [*] and Bruce Beck^

**Bulk Pharmaceutical Manufacturing, ^Department of Health and Safety*
Eli Lilly and Company
Clinton Laboratories
Clinton, IN 47842

Abstract

An accident model and a model of human cognition were combined to create a hypothesis about a major source of causation for accidents at a manufacturing site. This led to a definition of a routine/non-routine task or situation as well as some details around the different types of non-routine situations that can occur. The hypothesis is that situations or tasks that are non-routine but where the people involved cannot recognize the non-routine nature of the task, lead to a significant number of accidents.
A survey was performed to gather data to see if there is any support for this hypothesis. The data was analyzed and it was shown that a large portion of the accidents do occur in non-routine tasks or situations but that the people involved considered them to be routine. The most common reason for a situation/task being non-routine was non-standard conditions and infrequent tasks.
Finally, some recommendations are made that may reduce the number of accidents occurring under such non-routine situations or tasks. In particular, a special training course was created and delivered to all plant operations people to raise their awareness of the non-routine issue. As the notion of routine/non-routine tasks was explained to people, it was very easy for them to relate to activities not only at work but perhaps more importantly, to activities outside work. Many people gave us examples of non-work related tasks that had lead to accidents or near misses when they were performing such non-routine activities.

Introduction

An important metric for any organization is the rate of OSHA recordable accidents[1], measured as the number of serious accidents per year per 100 employees. The industry standard would indicate that a serious accident rate of less than 1 per 100 employees represents outstanding performance (even though the real goal is zero serious accidents). While we would like to have no accidents at all, when a serious accident does

---

[=] Author
[*] Corresponding Author
[1] Also referred to as serious accidents.

occur, we have a duty to learn as much as we can from the accident. This learning can be used to reduce the probability of that particular accident, as well as other related accidents, from happening again. Therefore it is very important to analyze the accidents that have occurred from the perspective of event based learning. There are two basic approaches to doing this,

a) We can consider that each accident is a special cause event and analyze each accident individually. From a regulatory perspective, we have an obligation to do this and we will learn something from this approach.

b) We can step back somewhat and look at all the accidents from a common cause perspective. This is the approach favored by Deming[2]. In this approach, we look for trends in the data that might uncover some systemic issues. Solving such systemic issues would then lead to significant improvement and significant learning.

In this report, a common cause analysis as described in (b), is presented. Originally, this approach was developed because of some limited observations about the causes of accidents. People involved in accidents had reported that when the accident had occurred, they had been doing a routine task but after some questioning, it began to look as if there were some aspects to the tasks that were non-routine. The common cause analysis allowed the validity of these observations to be determined. More recently, the analysis has been recast in light of an accident model and some understanding of human behavior. Therefore, the report will be structured by first discussing the accident model and the human behavior model. Then it will be shown how these models lead to a hypothesis about the causes of accidents, how this hypothesis was subsequently validated with data and finally, what conclusions and recommendations have been made based on the hypothesis validation.

The Accident Model of Houston(1971)

For the analysis presented here, we wanted a model that focuses on the time sequence of events that leads up to an accident. Houston(1971) has described such a model[2]. The overall model is shown in Fig. (1). This model considers that an accident occurs in three phases,

> a) The Induction Phase - this is where the sequence of events sets up a situation which is conducive for an accident to occur. At the end of this phase, something has occurred that is outside the normal range of behaviors for the system in question. Some people refer to this as an incident. But this incident has not yet become an accident.
>
> b) The Development Phase - in this phase, the consequences of the incident are determined. For example, the incident might be a minor or a major accident. If an accident does occur, it happens in this phase.

[2] Houston, D. E. L., "New Approached to the Safety Problem", 1971, I.Chem.E. Symposium Series No. 34.

c) The Resolution Phase - in this phase, the final impact of the accident is realized. For example, how much time will an injured worker spend off work, what will the full consequences be, what improvements will be made as a result of the accident investigation. The resolution phase contains a feedback loop that leads back to the induction phase.

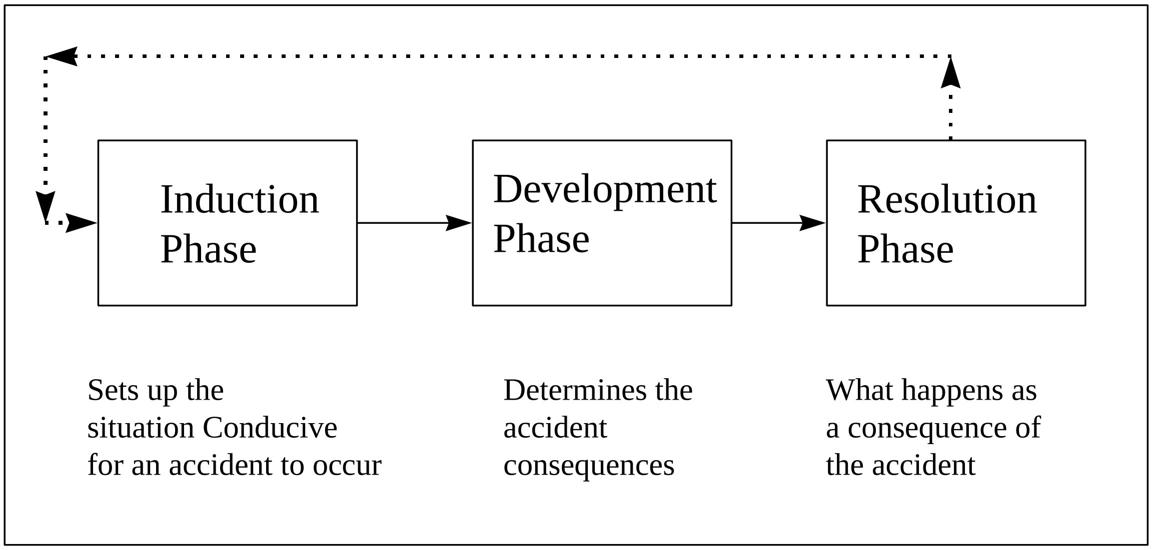


Figure (1): The Overall Accident Model

One of the critical differences between the induction and development phases that Houston points out relates to where the initiative lies in each phase. In the induction phase, the initiative is with the people involved in the situation (this could be operators, engineers, etc…). In the development phase, this initiative passes to the situation rather than stays with the people involved. The people involved lose control as the sequence moves from the induction phase to the development phase. Thus it is preferable to intervene in the accident process in the induction phase. This model places great emphasis on what is going on just prior to the accident happening. It focuses on the fact that it is not what happened yesterday, the day before or the many other times this task was performed that will cause an accident. Similarly, it is not what is going to happen tomorrow that will cause an accident now. It is what is going to happen in the next short period of time as this task is done that will cause the accident.

This leads to the following principle:

Principle of Accident Causality

It is the state of a system just prior to an accident which determines whether an accident will or will not occur.

This may seem like an obvious and trite observation. However, if one accepts it, it has some important consequences. First, it means that extrapolating previous performance may be a poor indicator of what will happen the next time a task is performed. Second, a person's assessment of the current situation is crucial in determining whether or not there are unacceptable levels of risk in the task that is about to be performed. There are, however, issues with how people assess situations. Assessing situations is a cognitive process and, in order to understand these issues, one must look at a model of how human cognition works.

The Human Cognition Model of Reason(1990)

We use Reason's (1990)[3] model of human cognition. This model recognizes that there are two cognitive subsystems at work in humans, the automatic subsystem and the intentional subsystem.

a) **The automatic subsystem** is characterized by many parallel operations. This parallelism means that the automatic subsystem is very fast. As an example, when we speak, we actually generate several competing activities that are carried out in a parallel fashion. We have to correlate speech syntax, vocal chords and mouth movements together. This subsystem is also used when we type keystrokes very fast at a keyboard. Sometimes, these parallel activities overlap and we create errors, such as when we hit several keys at the same time or out of order to make a typing error.

The automatic subsystem works by making use of schemata - these are organized collections of information and response patterns. These schemata are kept in long-term memory and there seems to be a large capacity to keep such schemata. The automatic subsystem works by finding schemata for a particular situation. It does this in two ways. First, it uses pattern matching. If a schema can be found that resembles the current situation, then that schema is used as a model for the situation and the automatic subsystem triggers responses that are appropriate for that schema. Second, if such a matching schemata cannot be found, frequency gambling is used. Here, the schema that has been used most frequently under similar conditions in the past is used. Most of the time, this frequency gambling will produce an appropriate schema. There is a statistical chance, however, that such an approach will result in an error because the chosen schema is inappropriate. There is something about the situation that the automatic

[3] Reason, James, 1990, "Human Error", Cambridge University Press, ISBN 0-521-31419-4

subsystem does not recognize and so it chooses an incorrect schema. Therefore we end up doing something we have done many times before, rather than what we should do. Comedians use the automatic subsystem effectively when they tell jokes. Typically, a joke starts off by describing a situation that is familiar to an audience and then at the end, the comedian adds the "punch line" which completely undermines the schema that the audience has created. People interpret this as funny. As an example, Jones(1995)[4] gives an example of such a joke in the context that we have described.

In summary, the automatic subsystem is a very fast, easy to activate subsystem that uses experiences to decide what should be done in the current situation. It is simple, uses very little brain resources and is not very demanding or stressful on the individual. However, it is exactly those features that make it easy to use that cause it to be unreliable in certain cases, such as when subtle but important changes have occurred in a system. It makes us depend on historical performance when such historical performance may be inappropriate.

b) **The attentional subsystem** has powerful logical capabilities and is approximately equal to consciousness. Thus, using this subsystem will be more reliable than using the automatic subsystem, assuming that it is used properly. However, using the attentional subsystem properly is difficult because this system is slow, stressful and difficult to maintain for more than brief periods of time. The resources that can be assigned to the attentional subsystem are very limited. For example, the number of things we can keep reliably in short-term memory is about seven. This limitation of resources can lead to errors in memory and logical analysis. An example of how this affects logical thinking is when you ask someone to consider the implications of an event occurring - say a valve failing. Generally, people are good at thinking of the direct implications - the flow stops. However, if there are less obvious implications, they may be missed (even though when they are pointed out to people, they will see them as obvious). The presence of the attentional subsystem also explains why people cannot participate in such efforts as process hazards reviews for more than a few hours without feeling mentally exhausted[5].

In applying this cognition/human error model to safety, we must first understand that there are three ways that we can deal with hazards:

a) Remove the hazard completely.

b) Place a barrier around the hazard.

c) Institute a procedure to allow personnel to manage their tasks with the hazard present and train people on the procedure.

[4] Jones, M. D., "The Thinker's Toolkit", Random House, 1995, ISBN 0-8129-2601-3.
[5] This type of mental exhaustion is referred to as *cognitive strain*.

Very often, it is not practical to remove a hazard even though this is the preferred approach. For example, using a flammable organic solvent always has a flammability hazard. Therefore, we tend to use the other two approaches for dealing with hazards. This strategy places more emphasis on the people involved in the task being able to recognize

a) that the barrier may not be in place as it was intended to be or

b) that the situation has changed such that the current procedure is not adequate.

We can now combine the accident model and the human cognition model with the above observation to form a hypothesis about why accidents occur. Our strategy for dealing with hazards places a lot of emphasis on hazard recognition. The accident model says that being able to recognize hazards just prior to doing a task is very important. The human cognition model says that in assessing a situation, people are more likely to use the automatic subsystem. This subsystem has a high probability of missing any subtle differences between the current situation and the schemata for that situation that a person has stored away. This means that we are likely to see a lot of accidents where the situations (or tasks) involved are non-routine in some way but the people involved think the situations (or tasks) are routine. This is the hypothesis that we would like to test. The remainder of this report is concerned with the testing of this hypothesis.

Taxonomy for Routine/Non-Routine Tasks

One of the initial concerns was that we needed to develop a robust definition of what we mean by a non-routine task. One of the issues that was found in the original studies was that operators considered tasks as routine that we considered non-routine. A group was created to formulate such a robust definition and the following definition was developed. Four attributes of a task were identified that would be used for this definition

a) **Comfort Level** with situation or task by an individual

A High Comfort Level would mean that:
- Individual did not require direction from procedure or other written instruction.
- Individual did not require direction from others for performing task - performed task on own volition.
- Individual has a high level of confidence in his/her ability to perform task.

A Low Comfort Level would mean that:
- Individual required written instructions for task.
- Individual sought out input from others before taking action.
- Individual is tentative or unsure of independent action.

b) **Capability/ Experience** demonstrated by individual regarding situation or task

High Capability:
Recent experience in performing task.
Demonstrated capability in executing task successfully and appropriately

Low Capability:
Inexperienced in performing task
Inability to perform task in past without error.

c) **Frequency** that the individual(s) have done the tasks before

High Frequency:
Individual performs task on some cyclic or repetitive basis
Individual performs task as part of primary loop responsibilities.

Low Frequency:
Individual performs task on when needed basis in a non-cyclic manner.
Individual does not normally perform task as part of normal job responsibilities

d) **Standard Conditions** exist for situation/task or they have been compromised

Exist:
SOP conditions exist for task or activity.
Individual's primary responsibilities and specific task match.

Compromised:
Situation or conditions are not normal for task.
Unusual factors not normally present are introduced to situation or task.
Individuals involved do not perform task as part of normal responsibilities.

If any one of the four attributes indicates a non-routine situation, then the situation should be considered non-routine. It is interesting that in the airline industry, a similar concept exists known as error chains[6]. Here, eleven links have been identified which could lead to a higher risk of an accident. These links are equivalent to the attributes we have listed above.

In order for the taxonomy to be useful, it must provide a robust definition of a non-routine task or situation. The survey conducted enabled this robustness to be measured. The details are presented in the data analysis section.

Data Collection

[6] See the website http://www.natcavoice.org/av/f97/humanerr.htm

Using the taxonomy developed, a team now selected at random a sample of 35 serious accidents from 1997. The team members independently assessed each accident with respect to the non-routine definition developed above by reviewing the accident report. Since some of the team members were directly involved in the accident reviews, they also used this knowledge as part of their assessment. Each team member assessed the routine/non-routine nature of the situation/tasks involved in the accident as well as which of the four attributes were present that caused the task or situation to be deemed as non-routine. **It should be noted that for all of the accidents in the data set, the question of whether or not the situation or task had been routine or not had been asked of the personnel involved in the accident and in all such cases, the respondents had answered routine.**

Data Analysis

The first analysis that was done with the raw data is to look at the robustness of the non-routine definition. Table(1) shows the data organized so that the robustness can be measured. For each accident, the table indicates whether each of the seven team members rated the accident as having occurred under a non-routine situation or not. The "% Agreement" column measures the extent of the agreement for each accident. If all seven team members agreed that it was non-routine, then the % agreement is 100%. If 2 team members said it was non-routine and 5 said routine, then the agreement is 5/7 = 71.4%. This is the same value as 5 team members agreeing that it was non-routine and 2 members agreeing that the accident was routine.

Table(1) Robustness of the Non-Routine Definition

Number	1	2	3	4	5	6	7	Total	% Agreement
1	0	0	1	0	0	0	0	1	85.7%
2	1	1	1	1	1	1	1	7	100.0%
3	0	0	1	0	0	0	0	1	85.7%
4	1	1	1	1	1	0	1	6	85.7%
5	1	1	1	1	1	1	1	7	100.0%
6	1	1	1	1	1	1	1	7	100.0%
7	1	1	1	1	1	1	1	7	100.0%
8	0	1	1	1	1	1	1	6	85.7%
9	1	1	1	1	1	0	1	6	85.7%
10	0	0	1	1	1	0	0	3	57.1%
11	1	1	1	0	1	1	1	6	85.7%
12	1	1	1	1	1	1	1	7	100.0%
13	1	1	1	1	1	1	0	6	85.7%
14	1	1	1	1	1	1	1	7	100.0%
15	1	1	1	1	0	1	1	6	85.7%
16	1	1	1	1	1	1	0	6	85.7%
17	1	1	1	1	1	1	1	7	100.0%
18	1	0	1	0	1	0	0	3	57.1%
19	1	1	1	1	1	1	1	7	100.0%
20	0	1	0	0	0	0	0	1	85.7%
21	0	0	0	0	0	0	0	0	100.0%
22	0	0	1	1	0	0	0	2	71.4%
23	1	1	1	0	1	1	1	6	85.7%
24	1	1	1	1	1	1	1	7	100.0%
25	1	1	1	1	1	1	1	7	100.0%
26	0	0	0	0	0	0	0	0	100.0%
27	0	1	0	1	1	1	1	5	71.4%
28	1	1	1	1	1	1	1	7	100.0%
29	0	0	0	0	1	0	0	1	85.7%
30	1	0	1	0	0	0	0	2	71.4%
31	0	0	0	0	0	1	0	1	85.7%
32	0	0	0	0	0	0	0	0	100.0%
33	1	1	1	1	1	1	1	7	100.0%
34	1	1	1	1	1	1	1	7	100.0%
35	0	1	1	0	1	1	1	5	71.4%

Given that there are seven people in the assessment, there are eight possible ways the assessments can be made and these are shown in Table(2). It can be seen that there are only four distinct possibilities for the % Agreement variable, 100%, 85.7%, 71.4% and 57.1%.

Table(2) Possible Values for the % Agreement Variable

Number Assessing the Accident as Routine	Number Assessing the Accident as Non-Routine	% Assessment
7	0	100%
6	1	85.7%
5	2	71.4%
4	3	57.1%
3	4	57.1%
2	5	71.4%
1	6	85.7%
0	7	100%

The average value for the % Agreement variable over the 35 accidents is 89%. In Appendix (A), it is shown that this average value is statistically significant. In other words, the chances of getting an average value of 89% given that the participants chose routine or non-routine in a random fashion with an equal chance of choosing one or the other, is very small. Thus, it can be concluded that the current definitions of routine and non-routine are robust definitions.

Next, the total number of routine assessments can be compared with the number of non-routine assessments. There were 165 non-routine assessments and 80 routine assessments. This means that the analysis indicates that the % of accidents that are occurring under non-routine conditions is 165/(165+80) = 67.3%. The range for the seven individual assessments is between 60% and 80%. Therefore the analysis has estimated that just over two thirds of accidents are occurring during a non-routine task or situation but the people involved do not realize that it is non-routine. In fact, the actual percentage may be more, as some people made the non-routine assessments based on knowledge they had about the accident above the data gleaned by each participant from the accident reports. Prior to the analysis, people would have said that the % of accidents under non-routine situations was very low or even zero.

Breaking down the data further, we can ask what percentages of the non-routine accidents are due to non-standard conditions, frequency, comfort level and capability. The results are[7]

[7] As some of the non-routine accidents have more than one of the attributes associated with it, the total here is greater than 100%.

Standard Conditions	80.6%
Frequency	38.2%
Comfort Level	6.7%
Capability	14.5%

This clearly indicates that the majority of non-routine tasks or situations is due to non-standard conditions or low frequency.

Finally, 111 out of the 165 non-routine accident assessments (67%) had only one non-routine attribute associated with them. This indicates that it is enough for only one of the attributes to be non-routine for the risk of an accident to be significantly increased.

Conclusions and Recommendations

The data analysis validates our hypothesis about the impact of non-routine situations and tasks on accident rate. A significant portion of the accidents occur when the situation or task is done under non-routine conditions but the people involved do not recognize the situation or task as non-routine. The reason for this is that people depend a lot on the automatic subsystem of their cognitive system to assess the state of a situation. This means that getting people more tuned in to recognizing these non-routine situations is a key to improvement.

Based on this validated hypothesis, three countermeasures were developed to improve the accident rate.

a. The accident report form was modified to incorporate a section to help people recognize the significance of non-routine situations or tasks. This also helps keep a focus on this aspect of accident causation as the accident investigation is performed. As an example, during such an accident investigation, the person involved was asked if there has been anything unusual about the circumstances around which the accident had occurred. Initially, he replied no, that he had carried out the task in the usual manner. However, upon further investigation, the person was asked to retrace his actions in detail. In doing so, he realized that he had changed his method slightly (but still within the bounds of the procedure) and it was this slight change that had lead to the accident. He had not realized that this slight change could lead to a significant increase in the risk of an accident.

b. A special training course was developed to heighten people's awareness of the issues surrounding non-routine activities. The point was emphasized that people need to take the time to ensure that they are aware of all the subtle changes that might make the task or situation non-routine. This course was delivered to all people on the site.

c. In some cases, routine/non-routine decision trees were developed. This was done to help people make a good decision as to whether the particular situation or task was non-routine.

d. Job aids were developed to help remind people on the job of the decisions involved in determining the non-routine nature of a situation or task.

These interventions lead to the following results

a. The language of routine/non-routine is becoming more normal both at the site involved and also at other sites. Language is always one of the first things that must be established when creating a significant change in the way people think.

b. Stories are being created. For example, a maintenance technician tells about having to do some work on a pump. As normal, he used a crane to lift up the pump once he had it disconnected. However, **because he was more aware of the risks of non-routine activities**, he took a bit more time to assess the situation before he started the work. He realized that the pump had been changed to a bigger size since the last time he had worked on it. He did not know whether this would make a difference or not but rather than take an unnecessary risk, he called for others to help him evaluate the situation.

c. Data analysis supports that the interventions have improved the accident rate. To show this, accidents were classified into chronic accidents and acute accidents. Chronic accidents are those that happen over a long period of time such as ergonomic related injuries. Acute accidents are those that occur quickly at the time of the accident, such as burns or lacerations. If the interventions were successful, then we would expect the percentage of acute accidents to go down after the interventions were put in place. The reason for this is that we only target acute accidents with our interventions. Prior to the interventions, the percentage of acute accidents was 66.7%. After our interventions, this percentage was reduced to 43.2%.

d. The routine/non-routine thinking was applied to areas other than safety. For example, in the utilities area of the site, they developed a routine/non-routine decision tree. This decision tree was applied to a major job on a large pump motor. It was found that the current procedure was in error because of a small change that had been made to the pump some time earlier. Had the utilities area used the current procedure, it would have lead to major damage to the motor upon restart. Worse, the area was under pressure to supply the market with product and the downtime caused by the damaged motor would have lead to costly supply problems!

In summary, the interventions lead to the intended improvements in safety and had some additional unintended improvements in other areas.

Acknowledgements

The authors would like to thank all the members of the teams involved in creating the routine/non-routine definitions and analyzing the accident reports.

Appendix (A): Statistical Analysis of % Assessment Data

In order to assess the statistical significance of the actual % assessment value obtained, we must first calculate the probability distribution of the % assessment value if the routine/non-routine assessments had been made purely by random choice with an equal probability of choosing routine or non-routine. Consider the assessment of one accident. The assessment of each person is like a binomial trial with seven trials, each trial representing the choice of a person[8]. For a trial with r routine assessments, the probability is given by the binomial distribution,

$$P(r) = C_r^n \left(\frac{1}{2}\right)^{n-r} \left(\frac{1}{2}\right)^r = \frac{n*(n-1)*(n-2)....(n-r+1)}{1*2*3.......r} \left(\frac{1}{2}\right)^n \qquad (A.1)$$

where P(r) is the probability of getting r choices of routine, C^n_r is the number of combinations of r objects from n objects and n is the total number of trials, which in the present case is 7. Looking back at Table (2), it can be seen that there are four possible values for the % assessment random variable and each of these four values can result from two possible outcomes from the trials. For example, a value of 100% can result from either getting all choices as routine and all choices as non-routine.

Table(A.1) Probabilities for the % Assessment Variable

Number Assessing the Accident as Routine	Number Assessing the Accident as Non-Routine	% Assessment	Probability
7	0	100%	0.0078
6	1	85.7%	0.0547
5	2	71.4%	0.1641
4	3	57.1%	0.2734
3	4	57.1%	0.2734
2	5	71.4%	0.1641
1	6	85.7%	0.0547
0	7	100%	0.0078

Thus the probability distribution for the % Assessment variable, which we will denote A, is

P(A =100%) = 2*0.0078 = 0.0156
P(A=85.7%) = 2*0.0547 = 0.10938
P(A=71.4%) = 2*0.1641 = 0.328
P(A=57.1%) = 2*0.2734 = 0.5469

The average value for A, μ_A is calculated from the formula

[8] This is like a coin toss with an equal chance of heads or tails.

$$\mu_A = \sum P\left(A = A_I\right) * A_I \tag{A.2}$$

where the summation in Eqn. (A.2) is taken over all possible values of A, denoted A_I. Substituting in the values for the probabilities from table (A.1) gives the mean value as 65.6%. The standard deviation of A, σ_A, is given by the formula

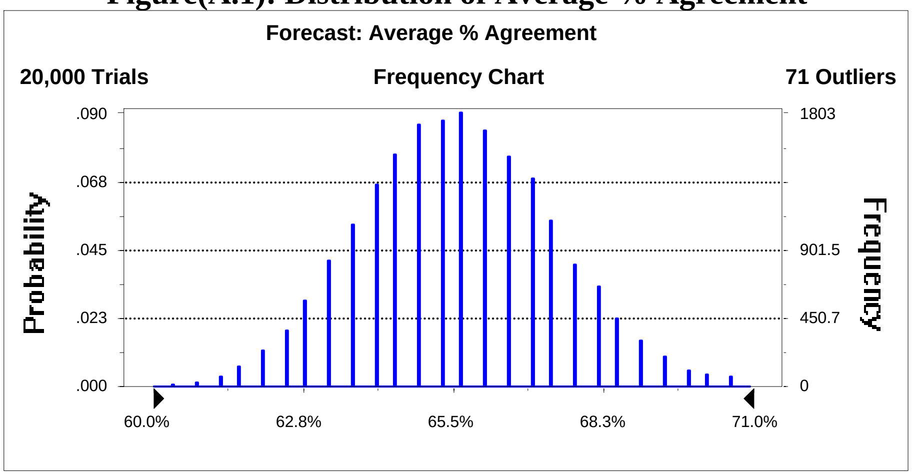
$$\sigma_A^2 = \sum P\left(A = A_I\right) * \left(\mu_A - A_I\right)^2 \tag{A.3}$$

where again the summation in eqn. (A.3) is taken over all possible values of A. Using the values for table (A.1), eqn. (A.3) gives the value for σ_A as 10.6%.

The average value of the % assessments over the 35 accidents is a random variable that is the average of 35 independent random variables with probability distributions like the variable A. This random variable will have the same mean as A, but will have a standard deviation that is reduced by a factor of the square root of 35. Thus the random variable representing the average % assessment has a mean of 65.6% and a standard deviation of 1.8%. Considering that normal variation will be within 3 standard deviations of the mean, the normal range for this random variable, if the routine/non-routine choice was made purely by a random choice, is 60.2% to 71.0%. The value obtained via the survey is 89%, which is clearly well outside this range. In fact it is (89.0-65.6)/1.8 = 13 standard deviations from the mean.

As an independent check on the calculations presented above, a Monte Carlo simulation was created of the routine/non-routine assessment by random choice. The resultant distribution for the average % Assessment is shown in Fig. (A.1) below.

Figure(A.1): Distribution of Average % Agreement



It can be seen that this distribution is consistent with the calculated range of 60.2% to 71.0%. This again supports the conclusion that the definition we have created for routine/non-routine is a robust definition.