Deficiencies in statistical analysis and the decision to launch the Challenger

To: Challenger 2 team
Sharon Ahlers (Engrc 335)

From: Aaron Sidford, 11/2/07

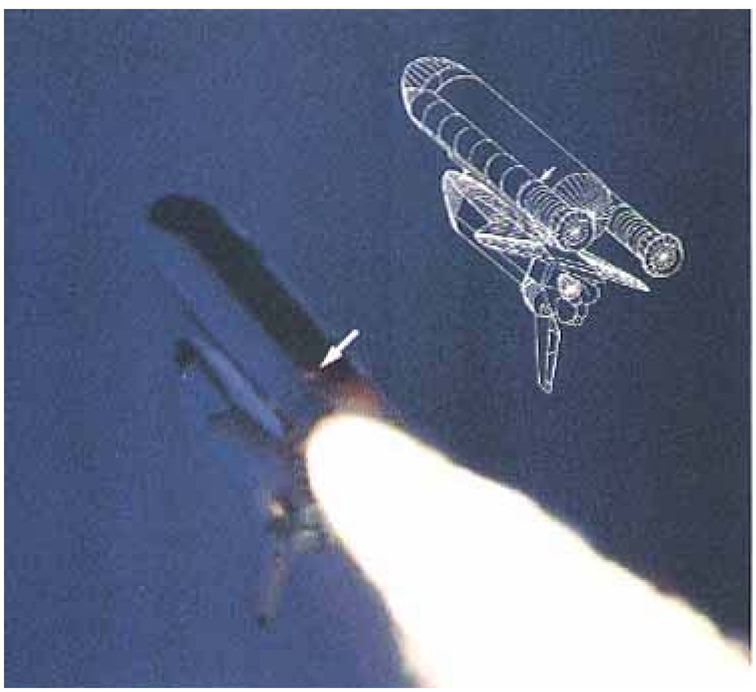


Table of contents

Executive summary

On January 28, 1986 at 11:38 AM Eastern Standard Time, the space shuttle Challenger was destroyed during launch and its crew of 7 lost there lives (Presidential Commission on the Space Shuttle Challenger Accident). There are many lessons to be learned from this tragedy. The accident illustrates the potentially disastrous consequences of poor organizational culture. It also shows how group think can have dire consequences and it shows the great need for proper organizational structure.

In Lighthall's (1991) article "Launching the Space Shuttle Challenger: Disciplinary Deficiencies in the Analysis of Engineering Data" a very different type of lesson is gleaned from the Challenger accident. He argues that a key factor in the Challenger accident was a deficiency in the training in the Morton Thiokol engineers. He argues that the engineers had the data to convince the managers to delay the launch of the Challenger but were unable to due to a lack of proficiency in statistics. On the day before launch engineers at Morton Thiokol recommended that NASA delay the launch of the Challenger due to the dangers posed by cold weather. Rather than follow this recommendation, the NASA managers demanded proof of the engineer's fears. Lighthall shows that the engineers needed to demonstrate a simple statistical correlation to the NASA manager but were unable to do so. Lighthall further illustrates that not only was such an analysis possible but that it could have been done 6 months before the Challengers launch. Lighthall uses this point to argue that there needs to be a greater emphasis on statistical analysis and inferential statistics in the engineering curriculum.

Lighthall does clearly show that better statistical analysis would have helped the engineer's case. However, he by no means shows that this is necessarily the key cause of the Challenger accident. He also fails to show that the true solution to accidents such as the Challenger is more statistical training for engineers. He doesn't demonstrate that the engineers lacked statistical training rather than time. He also does not demonstrate that statistics is the only discipline that could have saved the Challenger. Furthermore, he does not show that statistical analysis is really the job of engineers and that there isn't a better organizational fix to accidents such as the Challenger. Thus, Lighthall shows the power but not the necessity for advanced statistical training of engineers.

The enclosed individual case study report expands on all aforementioned points. For someone new to the Challenger case study, it is recommended that you read the glossary and introduction before reading the rest of the paper. Also, to aid your reading, note that glossary terms are italicized the first time they are used in the text. If you are more familiar with the case study and pressed for time you may skip the introduction and proceed right into the summary of the case sources section. Hopefully the report will provide some new information and insight to all.

Introduction

On January 28, 1986 at 11:38 AM Eastern Standard Time, the *space shuttle Challenger* was launched as part of *NASA Mission 51-L*. 73 seconds later there was an explosion and severe aerodynamic loads caused a complete structural breakup of the Challenger. The launch was a failure and tragically the shuttle's crew of 7 lost their lives (Presidential Commission on the Space Shuttle Challenger Accident).

Immediately after the tragedy, the Presidential Commission on the Space Shuttle Challenger Accident (also known as the Roger's Commission) was established to ascertain the cause of the accident and provide recommendations to avert such catastrophes in the future. They found that the immediate cause of the accident was the failure of a key structural component of the shuttle's rockets, its *O-rings*. They further noted that there were organizational problems with NASA that contributed to the catastrophe (Presidential Commission on the Space Shuttle Challenger Accident).

The commission also noted that the accident had not occurred in isolation. On the day before the Challenger's launch, engineers at *Morton Thiokol* proposed delaying the launch. Based on their experience with prior shuttle launches and their knowledge of the physics of the rockets they feared that cold temperatures predicted for the day of launch posed a significant threat. They correctly feared that the O-rings might fail due to these weather conditions. They raised these concerns to their managers and later that day there was a teleconference between these engineers, their managers, and the managers at NASA. The engineers failed to convince the managers to delay the launch and tragedy ensued (Tufte 1997).

The engineers had it right, so why couldn't they delay the flight and save the crew? There are many views on this issue. Engineering texts often point out this is simply what happens when "heroic engineers are ignored by villainous administrators" (Tufte 1997, 18). Management schools often point out that the failure was due to the potentially disastrous effects of group think and problems in the technical decision making process. Students of statistics often point out that it was due to poor statistical training of the engineers. Sociologists often point to problems with organization culture as the root of the problem. Others have argued that the problem was poor communication techniques illustrated by the engineers' presentation to their managers (Tufte 1997). The Challenger accident is a commonly used case study in a variety of fields, all of which often take very different views on the causes.

This report gives a detailed analysis of one very specific argument for why the engineers failed to save the Challenger, by analyzing "Launching the Space Shuttle Challenger: Disciplinary Deficiencies in the Analysis of Engineering Data." (Lighthall 1991) In this article, Lighthall argues that if those present at the teleconference had been more statistically oriented, the Challenger tragedy may have been averted. He both demonstrates this point and uses it to advocate more statistical training in the engineering curriculum. In this report, Lighthall's argument is summarized, key arguments against his analysis are given, and suggestions are given for how this information could be used to draw a more complete picture of failed meeting between Morton Thiokol engineers and NASA Managers.

Summary of the Case Sources

Rather than blame the engineers or the managers, Lighthall takes a more agnostic view and blames their training. In reference to the teleconference, he explores the question “could the flawed decision processes have resulted from an inadequate professional education that all participants shared?” (63). Lighthall answers this question in three phases. First he explores what the engineers wanted and needed to show in the teleconference with NASA. Second he explores how and why they failed to show it. Third, he shows that if the engineers were more statistically oriented they might have been able to show it. Finally, he concludes by summarizing his claims and recommends increasing the emphasis on statistical analysis and inferential statistics in the engineering curriculum.

What the Morton Thiokol engineers wanted and needed to show

The engineers were concerned that the cold weather conditions on the day before the Challenger’s launch could cause the O-rings to fail to seal. Their fears came from two sources. First, based on their knowledge of the physics of the rockets they knew that cold temperature caused the O-rings to harden. They thought that this reduced resiliency would hamper the ability of the O-rings to seal. Second, when compared to the temperature of previous shuttle launches, the predicted temperature of the Challenger launch was closest to the launch of *STS 51-C*. At the time STS 51-C was the coldest launch to date and had incurred more O-ring damage than any other flight. The *blow-by* found in the rockets of STS 51-C were far worse than on any other flight. The predicted temperature for the Challenger’s launch was even lower than the previous low set by STS 51-C and they feared that the O-ring damage on the Challenger would be even worse. It was based on these two facts and their general experience as rocket engineers that they proposed to NASA delaying the launch until weather conditions improved (Lighthall 1991).

To successfully convince NASA to delay launch the Morton Thiokol engineers needed to demonstrate that there was a negative correlation between O-ring temperature and failure of the O-ring to seal. The key measures used to quantify O-ring damage in previous flights was measures of *erosion* and reports on blow-by. Thus the engineers needed to convince the managers that there was a negative correlation between temperature and blow-by or a negative correlation between temperature and erosion. If they could have shown this correlation, they may have been able to convince the parties involved that the low temperatures predicted for the Challenger’s launch were below NASA’s risk tolerance and necessitated a launch delay (Lighthall 1991).

The failures in the presentation and discussion to delay launch

The engineers at Morton Thiokol needed to demonstrate the negative correlation between O-ring temperature and damage indicators, blow-by or erosion. However, they failed to do so. In their teleconference with NASA they sent over 13 slides of visual evidence. Of these 13 slides none presented a metric of O-ring damage in relation to temperature in prior flights. Six of the slides had no data about O-ring temperature, blow-by, or erosion at all. Of the remaining seven, none showed the relation between temperature and blow-by except for one which showed an experiment that was later discounted. Thus none of the slides showed the desired relationship between variables (Lighthall 1991).

Another problem with the charts presented was that none of them contained more than seven independent data points. There were 23 previous launches for which the engineers had data and they were not fully utilized. Tufte (1997) points out that drawing conclusion from limited data points is poor scientific practice and Lighthall points out that this lack of data prevented everyone in the teleconference from coming to real conclusions. By focusing their discussion to a limited number of data points, rather than discuss the correlation between temperature and O-ring damage, the managers were able to come up with a counterexample to the engineers' claims. Rather than focus on general data trends, the managers pointed out that there was also blow-by on the flight *STS 61-A*. This flight was launched in much warmer temperatures and thus the managers argued that there was little correlation between temperature and O-ring damage (Lighthall 1991).

The engineers knew what they wanted to demonstrate to the managers. The problem was not ignorance in this regard. Lighthall argues instead it was that they did not know how to effectively demonstrate the correlation between O-ring damage and temperature. Lighthall quotes a number of the people present at the teleconference between the engineers and managers to illustrate this inability. The engineers claimed that, with respect to the correlation, they "couldn't quantify it" (66) and "had no data to quantify it" (66). The people at the meeting claimed that they could not "get a correlation between O-ring erosion, blow-by, and temperature" (66). They claimed that they "didn't have any real data" (67). However, Lighthall shows that not only did the engineers have sufficient data to substantiate their claims, but that they had possessed this data for months.

How the engineers could have convinced the managers

Lighthall shows that the engineers did have sufficient data to establish a correlation between temperature and O-ring anomaly. Using the data that the engineers had available Lighthall gives 4 charts that he claims effectively give enough information to strongly suggest that the launch should be delayed. He even argues that the engineers had this data available for a significant amount of time before the launch (Lighthall 1991). These charts are given in Appendices A.1 through A.4 along with descriptions of how the engineers could have used them to convince the managers of the significant risks involved in launching the Challenger in such cold weather conditions.

Lighthall's conclusions

At the end of the paper, Lighthall summarizes his discussion into seven distinct conclusions. He claims the following:

(1) The engineers at Morton Thiokol and the managers at both Morton Thiokol and NASA could not quantify the suspected relationship between O-ring temperature and O-ring erosion.
(2) Everyone at the meeting was focusing on the specific relationship between O-ring failure and O-ring temperature (not relationships between temperature and hardness or other variables).
(3) The slides used by the Morton Thiokol engineers to convince the managers to delay the launch were irrelevant with respect to this causal relationship.
(4) The slides used by the Morton Thiokol engineers suggest that they had sufficient data to establish this causal relationship.
(5) The fact that the Morton Thiokol engineers had sufficient data but did not use it properly implies that they were lacking elementary methods of statistical analysis and inference.
(6) Proper statistical analysis of the data yield the following
    a. Erosion for O-ring *field-joints* and erosion for *nozzle-joints* had different causes although the erosion data was pooled together.[1]
    b. Erosion in the field joints was subject to temperature whereas erosion in the nozzle joints was subject to leak check pressure.
    c. The temperature of both flights that saw significant "blow-by" was atypical and arguably beyond safety margins.
    d. An unsafe degree of erosion in the challenger flight was predictable from the data.
(7) The data necessary to make a proper analysis of the flight erosion data was available six months before the Challenger launch.

Lighthall uses these conclusions to stress that there should be more emphasis on statistical analysis and inference in the engineering curriculum. With better statistical training, Lighthall hopes that engineers may be able to avert tragedies such as the Challenger in the future (Lighthall 1991).

[1] Lighthall focuses heavily on this point. It took a bit of work in his article to develop. Although it is an interesting point it is not particularly useful for the general research of the Challenger 2 team. For more information about this, read his article (or if you are in the Challenger 2 team, ask me).

Analysis of the Sources

Is the engineering curriculum really to blame?

Lighthall provides a well organized argument for how the Challenger accident could have been averted. He clearly provides a convincing argument that cold temperature and field-joint O-ring failure are correlated. He also argues well how (theoretically) this argument could have been given to NASA managers well before launch. However, his argument is by no means complete.

Lightall provides enough evidence to show that improved statistical training would have definitely helped the engineers at Morton Thiokol. However, he does not show that this training would have been enough to avert the Challenger tragedy or that this lack of statistical proficiency was the root cause of the challenger accident. In the next three sections these holes in Lightall's analysis are explored.

Did the engineers at Morton Thiokol lack training in statistics? Or did they lack time?

Clearly proper statistical analysis was not presented by the Morton Thiokol engineers to the NASA managers. However, was this analysis presented because the Morton Thiokol engineers did not have any idea how to perform it or because they lacked sufficient time to perform it? Lighthall does not say either way.[2] Lighthall provides numerous quotes of the Morton Thiokol engineers claiming that they could not quantify the relationship between cold temperature and blow-by. Other engineers claimed that they simply did not have the "hard data" (Lighthall 1991) necessary to establish the causal relationship between these two variables.

Thus, given more time, would the engineers have been capable of producing the desired statistical analysis? The answer to this question is unclear. Lighthall does explain that the engineers had sufficient data to perform the desired statistical analysis months before the Challenger's launch. However, Lighthall does not explain whether or not the engineers at Morton Thiokol had more pressing issues to deal with before the launch. The *field-joints* of the boosters were only at risk in cold whether. Even though the engineers had sufficient data to perform the analysis, they may not have thought there would be cold weather on the day of the Challenger's launch. They may have given priority to other work over studying the O-rings.

Furthermore, Lighthall does not even explain how long it would have taken a "statistically oriented engineer" (Lighthall 1991) to perform the desired statistical analysis on the O-rings. He points out how a statistically oriented engineer could have figured out how to do the analysis easily; he does not say how long the entire analysis process would take (or how long it took Lighthall). Yes, using modern technology, performing the analysis would be extremely quick and easy. However, in the hectic day

[2] Perhaps a reference studied by another Challenger 2 team member can answer this question. Tufte and Lightall do not address this but perhaps another source does. This is a place where good team research may tie our individual case study reports together.

before launch, this process might have taken a lot longer and the engineers may have had good reason to think it was not in their best interest. Lighthall does not explain either way.

Why statistics?[3]

Even if the arguments in the last section are invalid, that still doesn't necessarily mean that Lighthall's thesis is definitely correct. Suppose that the arguments in the last section are invalid. Suppose that Lighthall is completely right. Suppose that the engineers did lack statistical training (and not just time) and that better statistical training in the engineering curriculum could have prevented the Challenger accident. Even if we give Lighthall all these assumptions that does not necessarily mean that poor statistics is the root cause of the Challenger accident. Is statistics the only technical discipline that the engineers could have used to save the Challenger?

Lighthall makes a very good case for why statistics would have aided the challenger. However, we can make many cases for how increased training in different areas would have aided the Challenger. Perhaps, if the engineers had better training in the field of communication they would have been able to persuade the managers. Maybe more communication background would have allowed them to develop a more convincing presentation. Perhaps if the engineers had more training in material science they would have been better able to understand the properties of the O-rings and convey them to the managers. Maybe if the engineers simply had more aerospace engineering training in the first place they would have been able to design better O-rings. In light of this, why should the engineering curriculum focus on statistics rather than more communication, materials science, or aerospace engineering? Although there may be a convince argument to this end, Lighthall does not give one.

Whose job is statistical analysis really?

Yes, statistical training for the engineers would have helped their presentation, but is that really their job to know and do this analysis? In designing engineering curriculum there is always the idea that, "Yes, it would be nice to add course X, but then we're going to need to take away a course, and what course would that be?" There is always a fear that if you better train engineers in one field, say statistics, then perhaps their competency of their core technical discipline would suffer. The Challenger would not have been helped if the engineers at Morton Thiokol could establish correlation between temperature and O-ring failure, yet produced defective O-rings. Training the engineers in statistics may be a weak substitute for more diverse personnel and better organization structure, and Lighthall does not address this issue.

---

[3] The idea for the criticism in the section came partially from Tufte (1997). Tufte mentions that different schools of thought suggest that different knowledge might have saved the challenger. He put these views frankly by saying "if only the engineers and managers had the skills of field X, the argument implies, this terrible thing would not have happened." (Tufte 1997, 17). This got me thinking about whether or not Lighthall had actually given evidence to say there is something qualitatively more important about statistics than another discipline. I did not see any such argument and hence this section became what it is.

Conclusions

Lighthall provides a very convincing argument for how the engineers at Morton Thiokol could have given a much better presentation to the managers at NASA using a little statistical analysis. He shows that the engineers at Morton Thiokol wanted this analysis but were unable to produce it at the meeting with NASA. However, his argument that lack of statistical training in engineers is a root cause of the Challenger accident and that better statistical training is needed in engineers of the future is incomplete. Was the issue with the Morton Thiokol engineers' analysis a lack of training or a lack of time? Also was it really the engineers that need the statistical training and if so is there a better policy prescription than more statistical training? These are questions that Lighthall failed to answer. Hopefully, by collaborating with the other members of the Challenger 2 team we may be able to answer these questions and come up with a more complete picture of the fateful teleconference the day before the Challenger 2 accident.

Appendix A.1
Chart Presented In Lighthall 1991 (68)

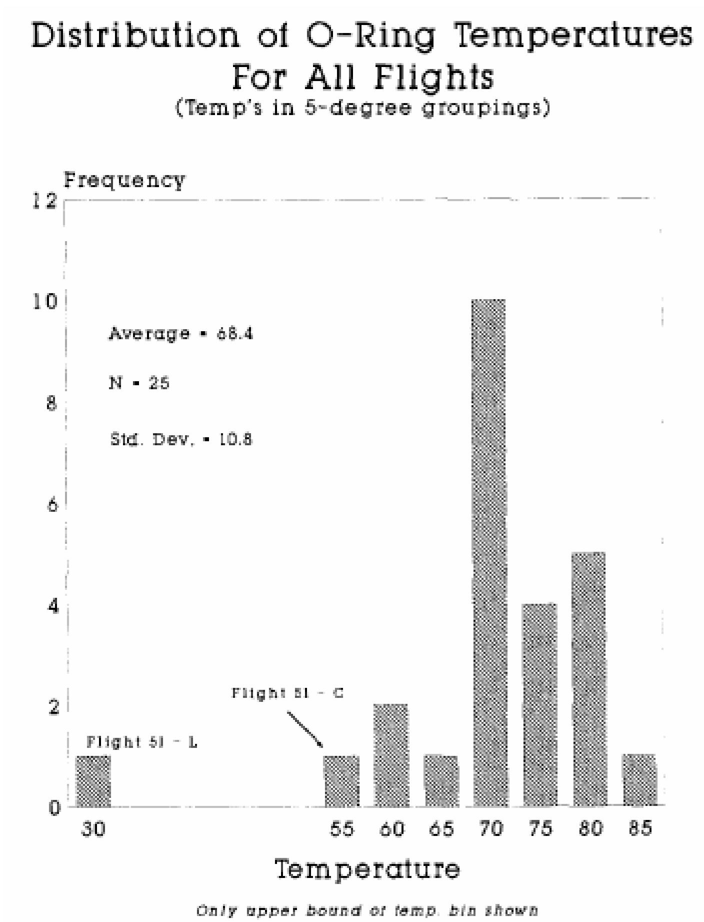
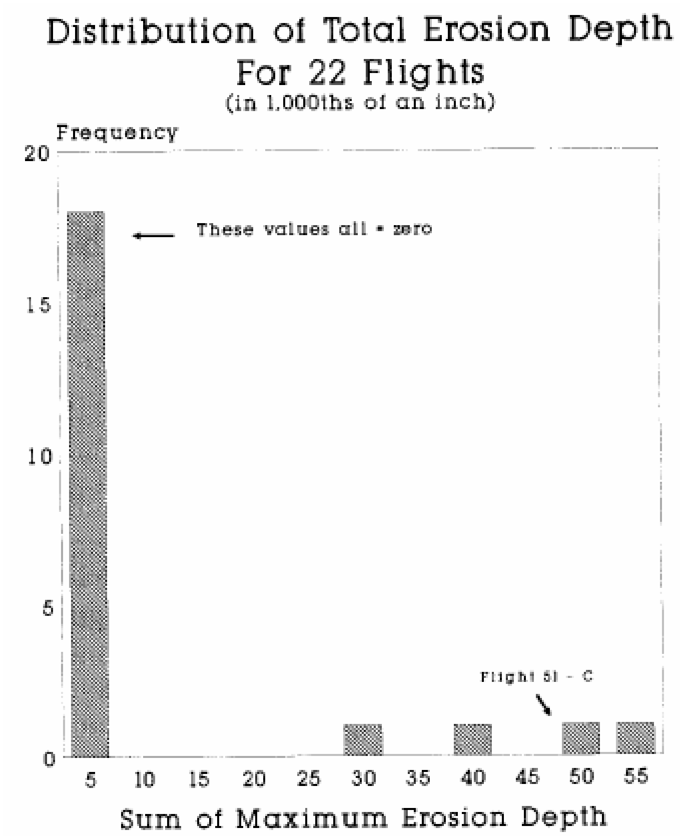


Fig. 1. O-Ring temperatures for all flights.

Figure 1 shows the distribution of the temperature during all previous shuttle launches. It clearly depicts that the Challenger launch was to occur at a temperature well beyond that of previous launches. The engineers could have argued that what would happen during a launch at this temperature is well outside of their knowledge base and is a significantly risky operation.

Appendix A.2
Chart Presented In Lighthall 1991 (69)



Only upper bound of erosion bin shown

Fig. 2. Erosion depth for 22 flights.

Figure 2 shows the distribution of erosion for different temperature launches. The engineers at Morton Thiokol could have used this chart to point out that very few of the previous launches had any significant erosion. Of the few that did, one of the ones with the most was STS 51-C, the shuttle launched in temperature conditions most similar to what the Challenger was to be launched in.

Appendix A.3

Chart Presented In Lighthall 1991 (70)

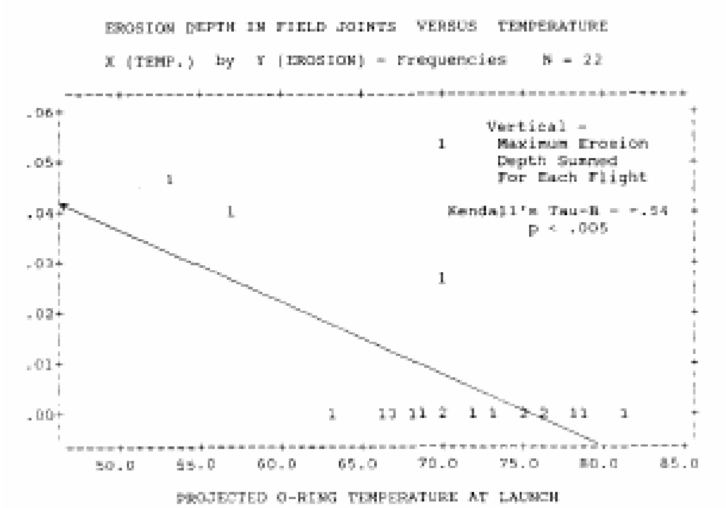


Fig. 3. Linear correlation between O-ring temperature and summed maximum depths (inches) of O-ring erosion for each flight for which complete data were available on January 27, 1986. Frequencies of observations at each bi-variate point in graph are indicated by number.

Figure 3 plots what the O-ring temperature at previous launches was predicted to have been against erosion found in the Shuttle's rockets. It also shows the results of statistical analysis that there is a negative correlation between temperature and erosion. The engineers at Morton Thiokol had this data available and could have used it to demonstrate the correlation between low temperature and O-ring damage they suspected.

Appendix A.4
Chart Presented In Lighthall 1991 (70)

TABLE I
TEMPERATURES, PREDICTED AND ACTUAL DEPTHS (INCHES) OF FIELD JOINT EROSION, AND UPPER 95% CONFIDENCE LIMITS OF PREDICTED DEPTHS OF EROSION FOR 22 SPACE SHUTTLE FLIGHTS

Flight No.	Temperature of O-rings	Actual Depth of Erosion (in thousandths)	Predicted Depth of Erosion (in thousandths)	Upper 95% Confidence Limit (in thousandths)
1	66	0	13	21
2	70	53	8	14
3	69	0	9	16
4	80	·	− 6	5
5	68	0	11	17
6	67	0	12	19
7	72	0	5	12
8	73	0	4	11
9	70	0	8	14
10	57	40	26	41
11	63	0	18	27
12	70	28	8	14
13	78	0	− 4	7
14	67	0	12	19
15 51-C	53	48	32	50
16	67	0	12	19
17	75	0	1	9
18	70	0	8	14
19	81	0	− 8	5
20	76	0	− 1	8
21	79	0	− 5	6
22	75	0	1	9
23[1]	76	0	− 1	8
25 51-L	29	·	66	107

[1]Data from Flight 24 are omitted because they would not have been available for analysis by the eve of the Challenger launch.

Figure 4 plots actual and predicted erosion for flights based on the analysis shown in Figure 3. The engineers could have used this chart to show that based on the analysis in Figure 4 the erosion that would occur in the Challenger is predicted to be much higher than any previous flight by at least a factor of two.

Glossary of technical terms

Blow-by – gas that passes over (blows-by) the O-rings during a space shuttle launch. Blow-by that results in heat but not damage to an O-ring was considered benign and the result of a brief gap in the O-ring seal, later filled within milliseconds. Blow-by that resulted in gray, or worse, black soot, between the primary and secondary O-rings was considered malignant and indicative of a dangerous potential failure of the O-rings to seal. Blow-by can cause the O-rings to become eroded making it even harder for the O-rings to seal later in the launch (Lighthall 1991, 64).

Erosion – wear of the O-rings during launch. Blow-by during launch can cause the O-rings to become eroded and make it more difficult for them to seal. During the first .17 seconds of launch there is blow-by and erosion on the O-rings at the same time the O-ring is trying to seal. Quoting one of Morton Thiokol's Managers, "it is a race between, will it erode more than the time allowed to have it seal" (Lighthall 1991, 64).

Field-joint – A place in the solid booster rockets joined. The Challenger accident was caused by a breach in a field-joint due to malfunctioning O-rings (Lighthall 1991, 64).

Morton Thiokol Corporation – the company that won the contract from NASA to develop the *Solid rocket motors* (Presidential Commission on the Space Shuttle Challenger Accident). Therefore it was this company and the engineers in it that were responsible for designing these rockets and the O-rings in them. (Tufte 1997, 17)

National Aeronautics and Space Administration (NASA) – the government agency responsible for space shuttle flights including the space shuttle Challenger (Tufte 1997, 17).

Nozzle-joint – Where the nozzle section the solid rocket boosters joined to the rest of the rocket. This joint was known to be have structurally different to the field-joints of the rocket. This joint was subjected to greater heat than the field-joints but different pressure than the field-joints (Lighthall, 1991).

O-rings – a rubber ring nearly 38 feet in circumference and ¼ inch thick used to seal field-joints (Tufte 1997, 16). In the field joints these O-rings were designed to be self-sealing. Pressure from the rocket was supposed to force the field joints into place and seal the joints closed. There was both a primary and a secondary O-ring for this purpose. If one O-ring failed to seal, the hope was that the other would (Lighthall 1991, 64). Cold weather on the day of the launch of the Challenger caused these O-rings to lose their resiliency and fail to seal. The result was an explosion during launch and the Challenger accident. (Tufte 1997, 16)

Solid rocket boosters – rockets used to generate 80% of the lift in the launch of a space shuttle such as the Challenger (Presidential Commission on the Space Shuttle Challenger Accident).

Solid rocket motors - one of the key components of the *solid rocket boosters* (Presidential Commission on the Space Shuttle Challenger Accident).

Space shuttle Challenger (STS 51-L) – a space shuttle used by NASA. On January 27, 1986 as part of NASA mission 51-L, there was an explosion during launch of this shuttle and its crew of 7 lost their lives in what we refer to as the "space shuttle Challenger accident.

Space Transportation System 51-C (STS 51-C) – A space shuttle launch on January 24, 1985 (nearly a year before the Challenger tragedy. The rockets were examined after launch and black soot was found in the O-rings. The engineers concluded that this meant that cold weather during the launch had caused the O-rings to harden, move more slowly, and fail to fully seal initially (causing gas to blow by and erode the O-rings). The temperature of the O-rings during this launch was estimated to be 53 °F (Lighthall 1991, 64).

Space Transportation System 61-A (STS 61-A) – A space shuttle launch before the Challenger. This launch also experienced sooted blow-by albeit much less severe than the blow-by seen in flight 51-C. However, this launch occurred at 75 °F. In the teleconferences where it was decided whether or not to delay the launch of the challenger, this flight was used by Nasa managers as a counterexample against the claim that cold temperature correlates with joint O-ring failure (Lighthall 1991, 66).

References

Lighthall, Frederick F. 1991. Launching the space shuttle Challenger: Disciplinary deficiencies in the analysis of engineering data. *IEEE Transactions on Engineering Management* 38 (1): 63-74.

Presidential Commission on the Space Shuttle Challenger Accident. 1986 (June 6). Washington, DC.

Tufte, Edward R. 1997. The decision to launch the space shuttle Challenger. In *Visual and statistical thinking: Displays of evidence for making decisions*. Cheshire, Conn: Graphics Press. 16-31.

The picture on the cover came from the Presidential Commission on the Space Shuttle Challenger Accident (1986).