CS 4001: Computing, Society & Professionalism

Munmun De Choudhury | Assistant Professor | School of Interactive Computing

Week 12: Visual Argument (Visual and Statistical Thinking by Tufte)

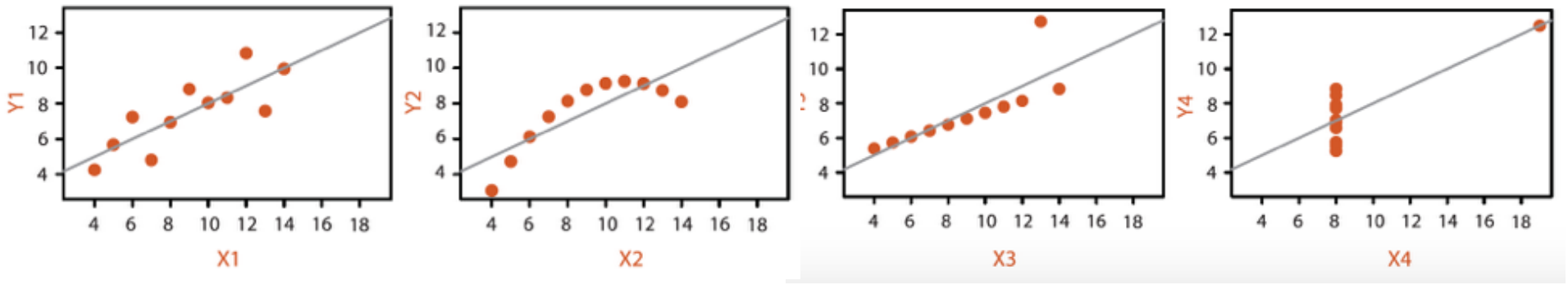
March 28, 2017

Why Visualize?

Munzner, 2014

Anscombe's Quartet: Raw Data

	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	



Why Visualize?

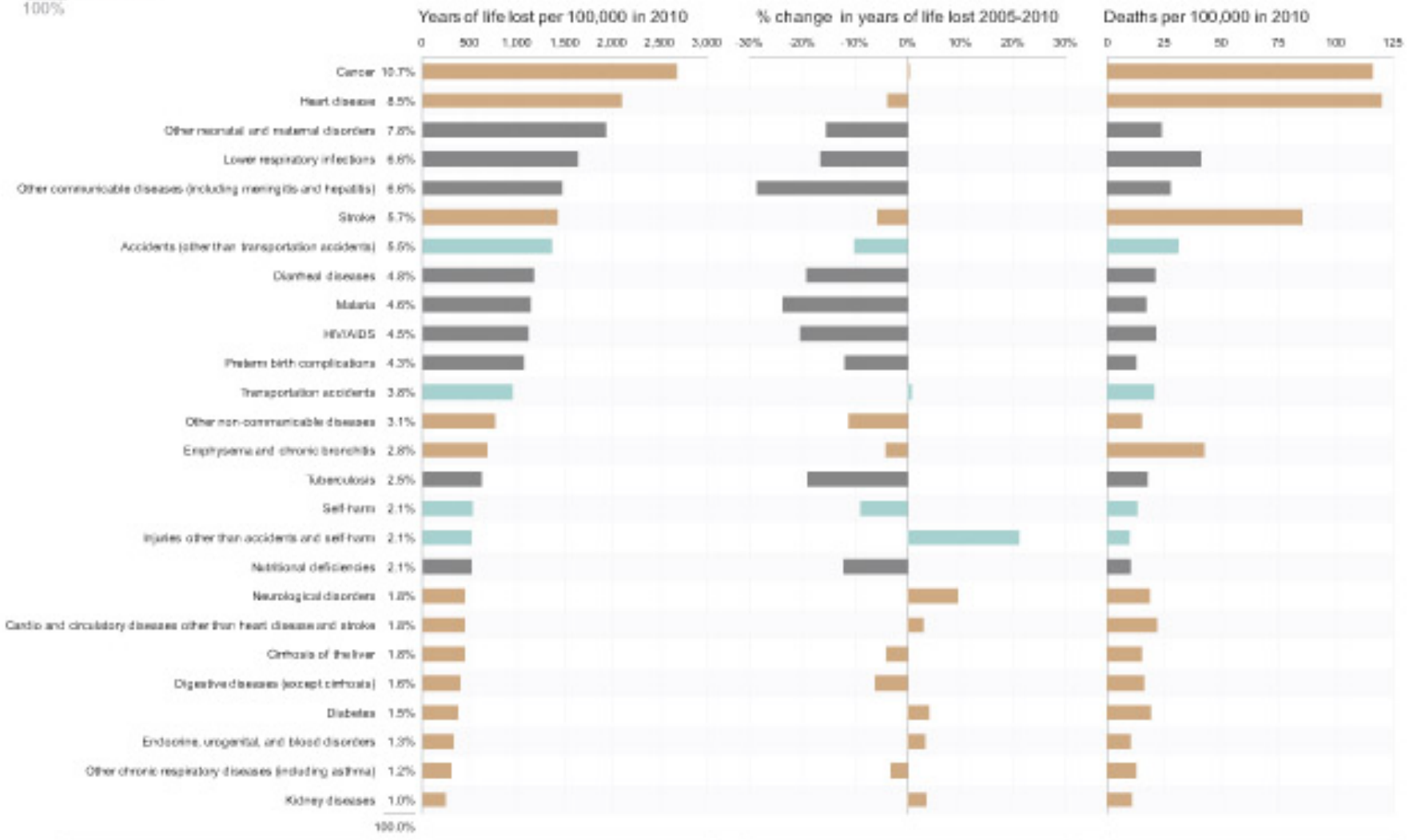
"Visualization is really about external cognition, that is, how resources outside the mind can be used to boost the cognitive capabilities of the mind"

— Stuart Card

Global Causes of Lost Life

44% Communicable, maternal, neonatal, and nutritional disorders
43% Non-communicable diseases
13% Injuries
100%

Comparing the number of deaths alone, as shown in the rightmost graph below, doesn't tell the entire story. Some causes of death have a greater effect on the young, which can be seen when comparing years of life lost in the leftmost graph.



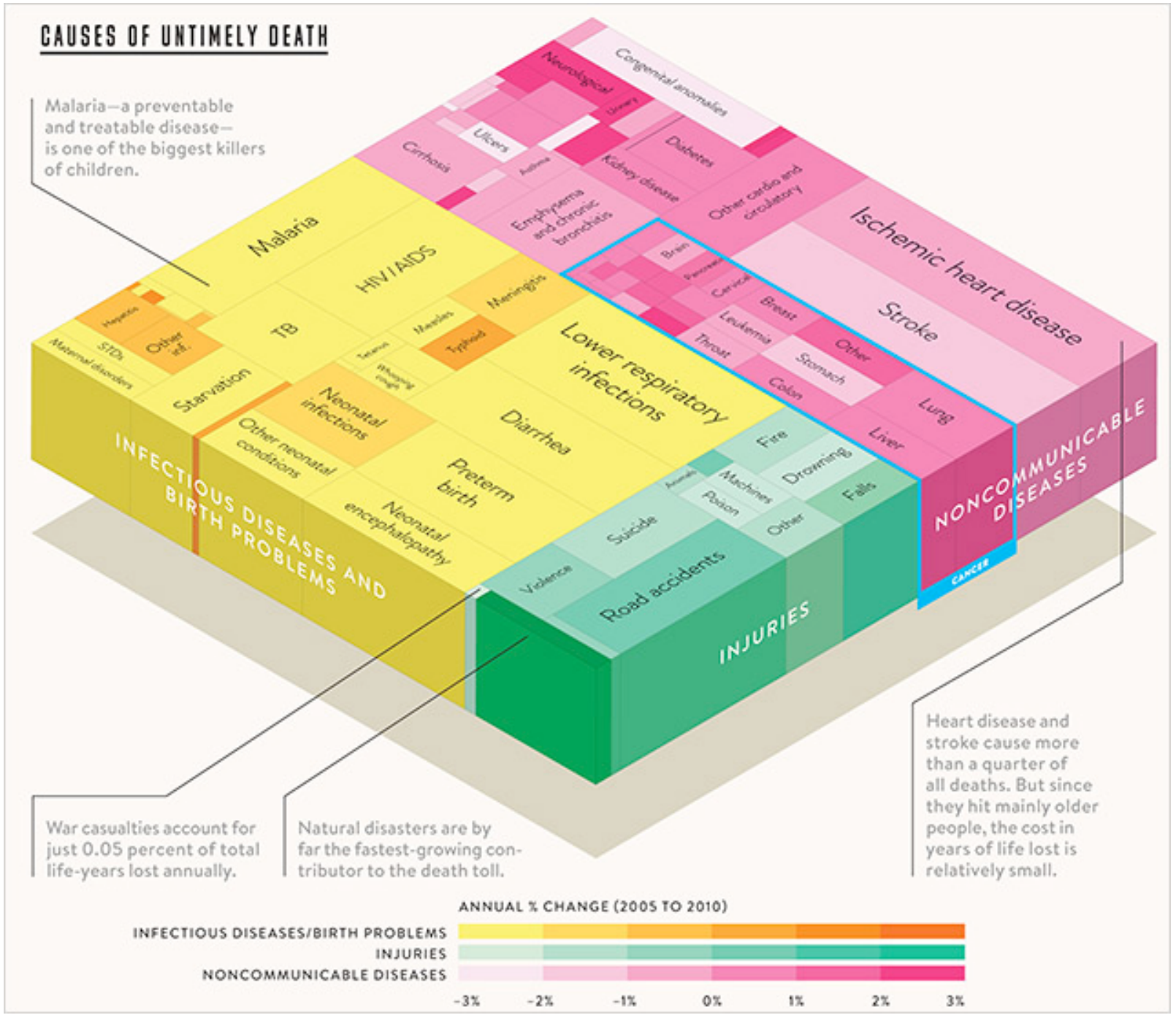
Some causes of death contribute disproportionately to years of life lost because of their effect on the young. For example, malaria, while not huge in the number of deaths, is much more significant in the number of years that are lost.

Two interesting changes reside in "Injuries other than accidents and self-harm." War, which accounted for only 0.06% of years of life lost, decreased since 2005 by 31.5% in years of life lost per 100,000 people. Natural disasters, which accounted for 0.65% of years of life lost, increased by 21.7% in years of life lost per 100,000.

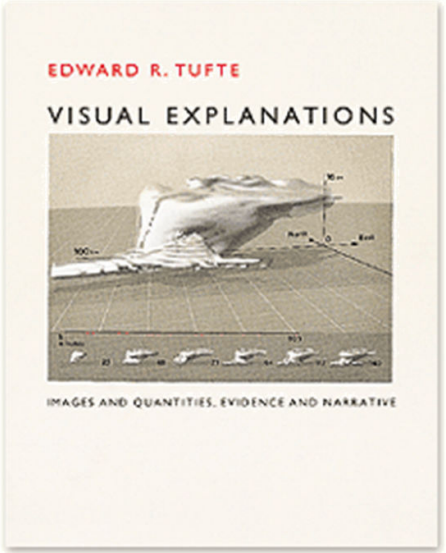
Communicable, maternal, neonatal, and nutritional disorders (the gray bars) are often easier to prevent through healthcare than other causes of death. This reveals itself in the graph above by the fact that all of these disorders have decreased during this five year period.

The five forms of cancer that cause the most deaths are trachea/bronchus/lung (2.9%), stomach (1.4%), liver (1.4%), colon/rectum (1.4%), and breast (0.8%).

All cardiovascular and circulatory diseases combined account for 30% of deaths.


CAUSES OF UNTIMELY DEATH
Malaria—a preventable and treatable disease—is one of the biggest killers of children.
Malaria
HIV/AIDS
TB
Hepatitis
Other inf.
STDs
Maternal disorders
Starvation
Meningitis
Measles
Tetanus
Whooping cough
Typhoid
Neonatal infections
Other neonatal conditions
Diarrhea
Lower respiratory infections
Preterm birth
Neonatal encephalopathy
INFECTIOUS DISEASES AND BIRTH PROBLEMS
Neurological
Congenital anomalies
Ulcers
Cirrhosis
Asthma
Kidney disease
Diabetes
Other cardio and circulatory
Emphysema and chronic bronchitis
Ischemic heart disease
Brain
Cervical
Breast
Leukemia
Throat
Other
Stomach
Colon
Liver
Lung
Stroke
NONCOMMUNICABLE DISEASES
CANCER
Fire
Drowning
Falls
Animals
Machines
Poison
Other
Suicide
Violence
Road accidents
INJURIES
War casualties account for just 0.05 percent of total life-years lost annually.
Natural disasters are by far the fastest-growing contributor to the death toll.
Heart disease and stroke cause more than a quarter of all deaths. But since they hit mainly older people, the cost in years of life lost is relatively small.
ANNUAL % CHANGE (2005 TO 2010)
INFECTIOUS DISEASES/BIRTH PROBLEMS
INJURIES
NONCOMMUNICABLE DISEASES
-3%
-2%
-1%
0%
1%
2%
3%

Edward R. Tufte's "Visual and Statistical Thinking: Displays of Evidence for Making Decisions"



"When we reason about quantitative evidence, certain methods for displaying and analyzing data are better than others. Superior methods are more likely to produce truthful, credible, and precise findings. The difference between an excellent analysis and a faulty one can sometimes have momentous consequences."

Poor displays often lead to invalid arguments and false conclusions.
Good displays help lead to valid arguments and true conclusions.

Two case studies with counter outcomes stemming from visual displays

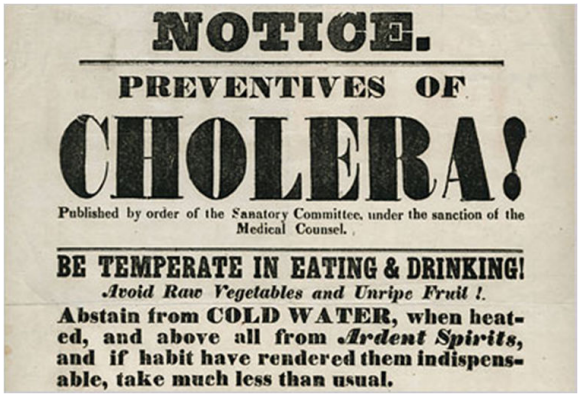
Case 1: John Snow intervenes in the London cholera epidemic of 1854

C

Cholera broke out in central London on August 31, 1854.
Cholera: severe watery diarrhea, vomiting, rapid dehydration
death can occur within hours of infection; fatality rate of 50%
killed millions in the 1800's in India, Russia, Europe, and N. America

C

Deficiencies in:
understanding of bacteria
technology
sanitary living conditions


NOTICE.

PREVENTIVES OF

CHOLERA!

Published by order of the Sanatory Committee, under the sanction of the Medical Counsel.

BE TEMPERATE IN EATING & DRINKING!

Avoid Raw Vegetables and Unripe Fruit !.

Abstain from COLD WATER, when heated, and above all from *Ardent Spirits*, and if habit have rendered them indispensable, take much less than usual.

Q

How is cholera transmitted?
How can we stop this cholera epidemic in central London?

H

Cholera is spread by: (1) breathing vapors of decaying matter or
(2) drinking contaminated water.

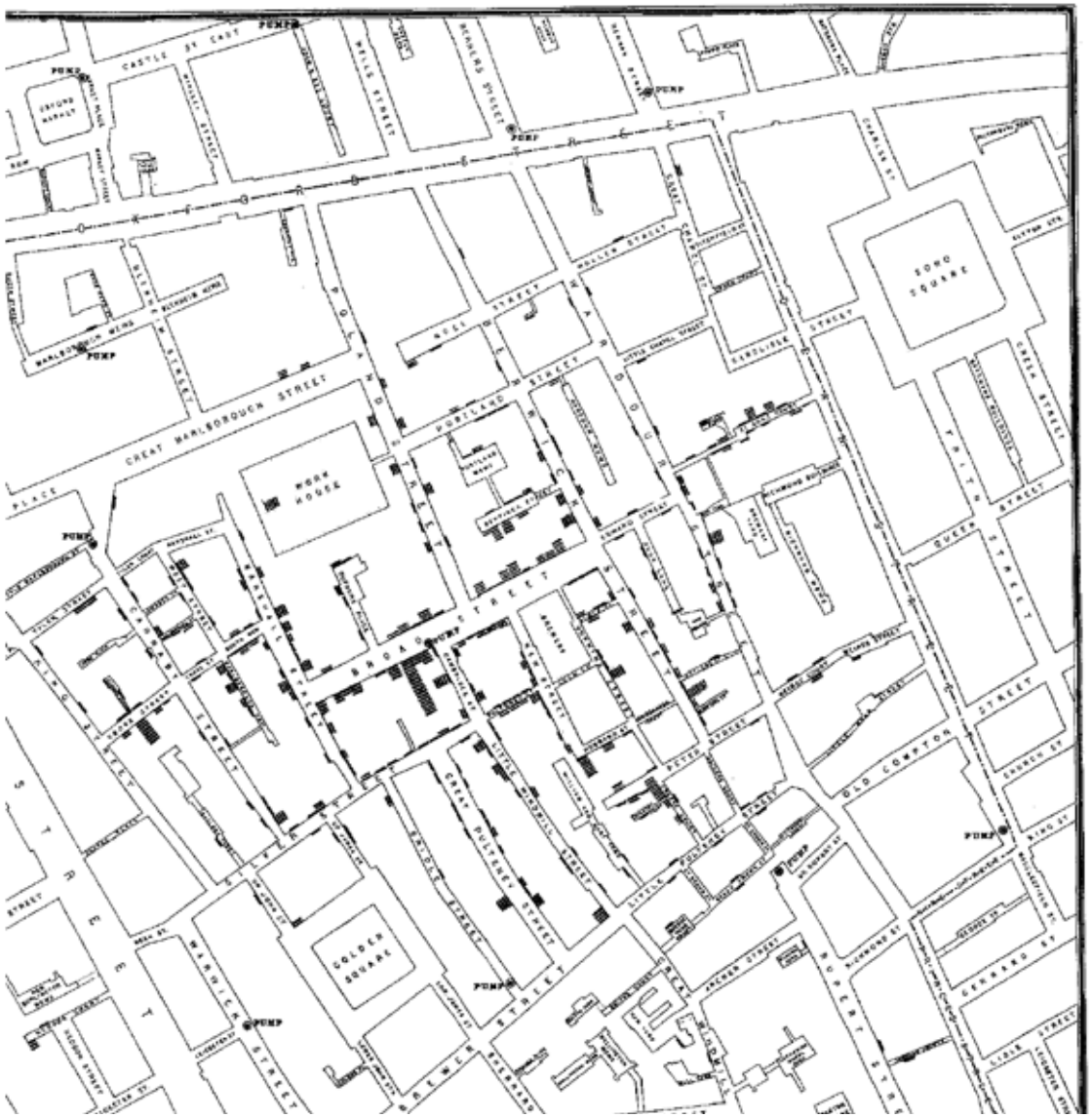
Snow's Designs and Methods:
He searches for correlations between water and cholera.

(1) Look for impurities in water → **No visible impurities** → **Dead End**

(2) Connect deaths with water sources → **Obtain a list of deaths from cholera from General Register Office** → **Convert original list of data (text) into a map**

John Snow's Cholera Visualization

Tufte, 2007

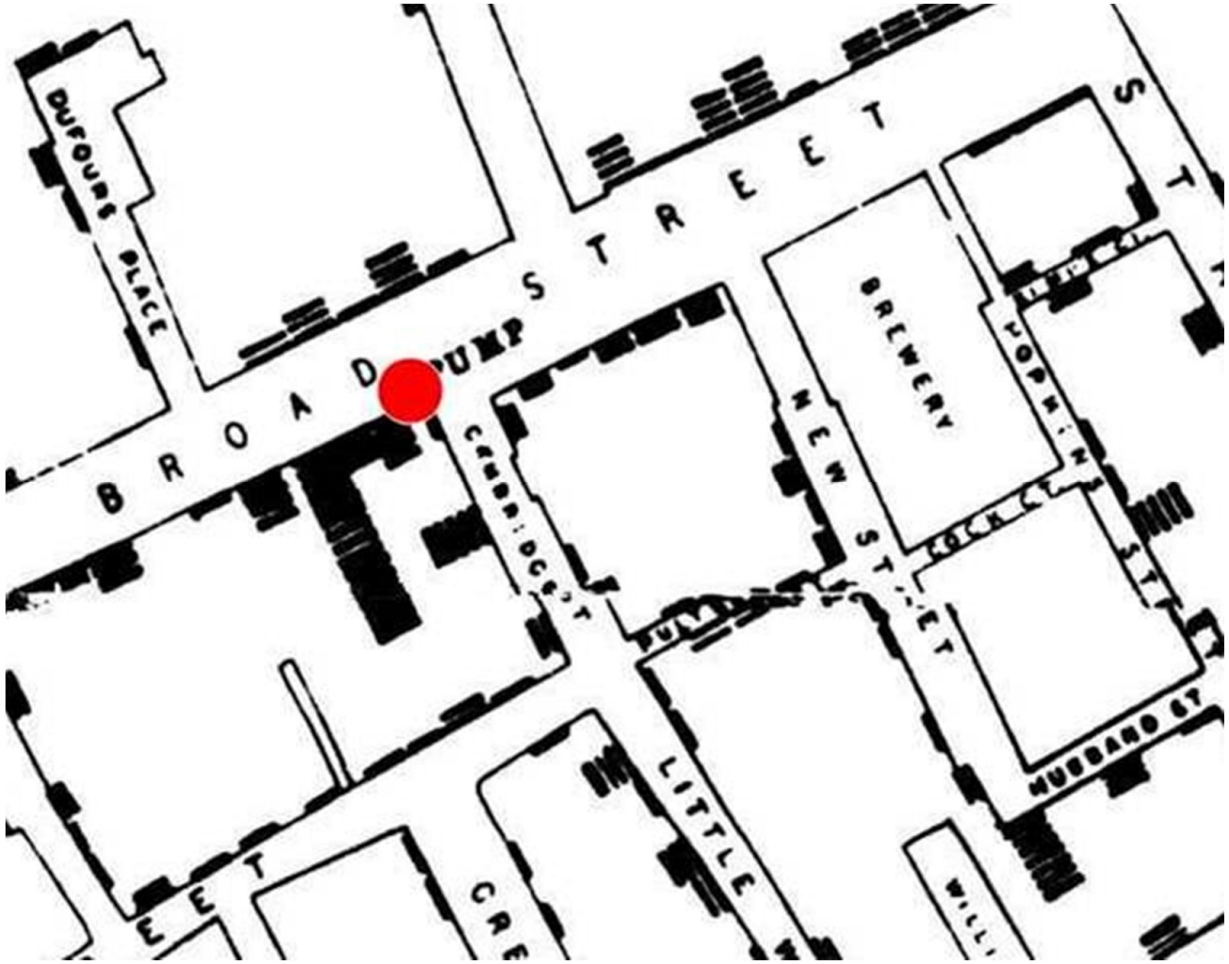


The graphical display was aimed at conveying information about a possible cause-effect relationship.

Snow marked

- deaths from cholera (lllllll)
- locations of 11 community water pumps.

Snow correlates deaths from cholera with locations of the water pumps



Water pump

Residence of cholera victim

Strong correlation of cholera victims near the Broad St water pump!

John Snow's Cholera Visualization

Tufte, 2007

The spatially arranged display allows inspection of alternative explanations and contrary evidence.

Dr. Fraser also first called my attention to the following circumstances, which are perhaps the most conclusive of all in proving the connexion between the Broad Street pump and the outbreak of cholera. In the 'Weekly Return of Births and Deaths' of September 9th, the following death is recorded: 'At West End, on 2nd September, the widow of a percussion-cap maker, aged 59 years, diarrhea two hours, *cholera epidemica* sixteen hours.' I was informed by this lady's son that she had not been in the neighbourhood of Broad Street for many months. A cart went from Broad Street to West End every day, and it was the custom to take out a large bottle of the water from the pump in Broad Street, as she preferred it. The water was taken on Thursday, 31st August, and she drank of it in the evening, and also on Friday. She was seized with cholera on the evening of the latter day, and died on Saturday. . . . A niece, who was on a visit to this lady, also drank of the water; she returned to her residence, in a high and healthy part of Islington, was attacked with cholera, and died also. There was no cholera at the time, either at West End or in the neighbourhood where the niece died.[10]

John Snow's Cholera Visualization

Tufte, 2007

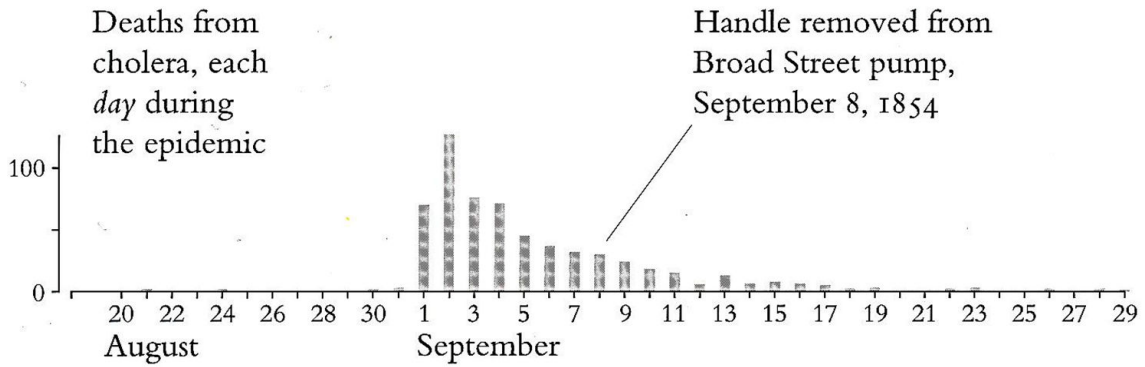
Snow's visualization enables quantitative comparisons to be made.

There is a brewery in Broad Street, near to the pump, and on perceiving that no brewer's men were registered as having died of cholera, I called on Mr. Huggins, the proprietor. He informed me that there were above seventy workmen employed in the brewery, and that none of them had suffered from cholera—at least in severe form—only two having been indisposed, and that not seriously, at the time the disease prevailed. The men are allowed a certain quantity of malt liquor, and Mr. Huggins believes they do not drink water at all; and he is quite certain that the workmen never obtained water from the pump in the street. There is a deep well in the brewery, in addition to the New River water. (p. 42)

"Saved by the Beer!"

Results and Conclusions:
Snow reports to the authorities

- **Snow described his findings to the authorities one week after epidemic.**
  - handle on the Broad Street water pump was removed on Sept 8
  - epidemic soon ended
- **But did Snow's intervention really cause the end of the epidemic?**



  - most people in central London had fled or died
- **Removing the pump handle probably prevented a recurrence.**
- **Snow's analysis and map provided strong evidence that cholera is transmitted by drinking contaminated water.**

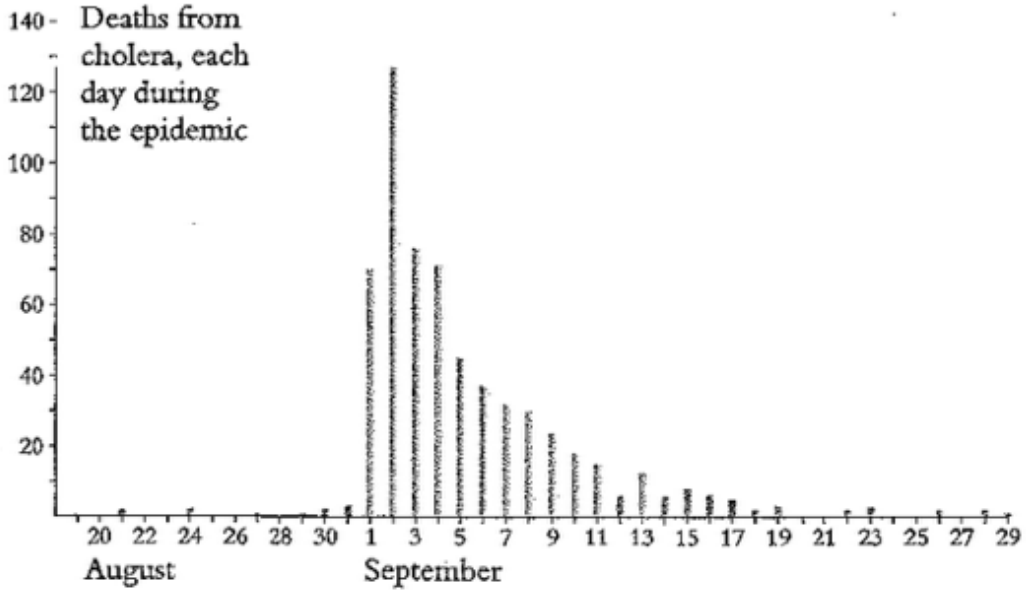
The Flip Side of Snow's Display

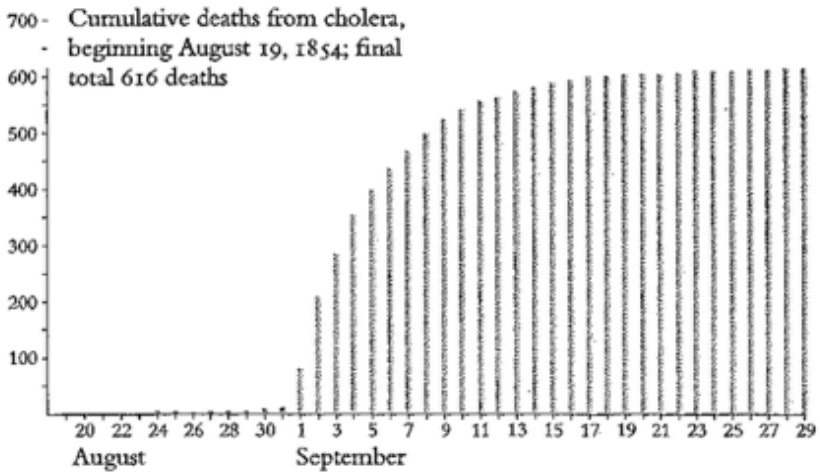
Tufte, 2007

The dot map

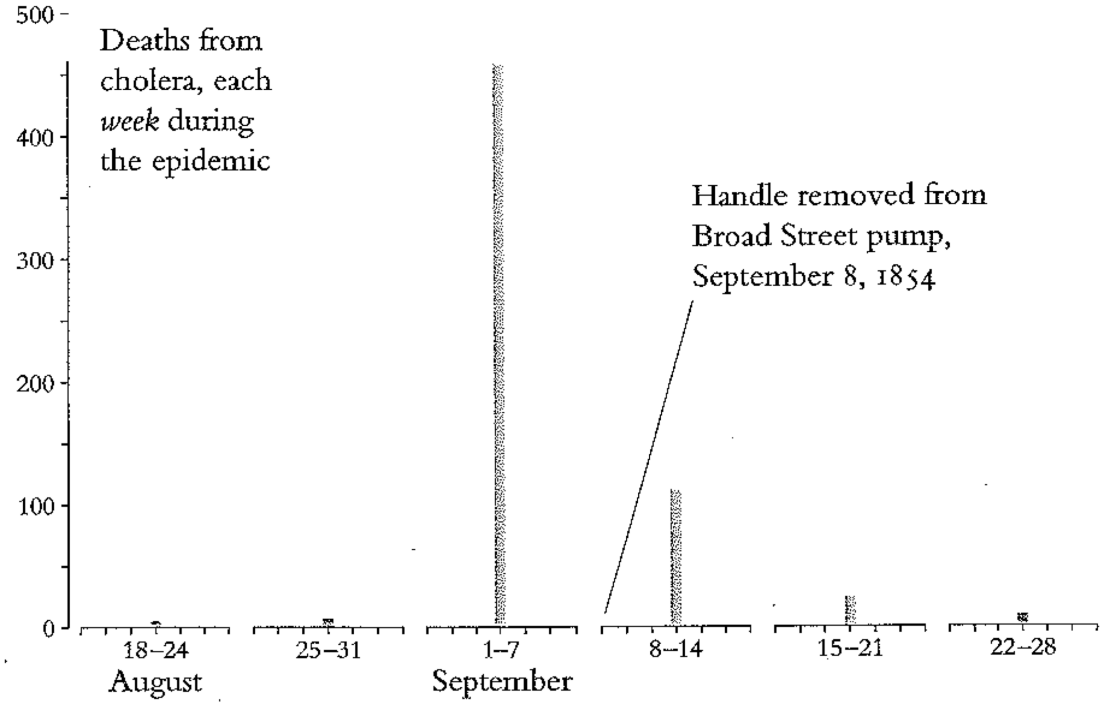
- does not take into account the number of people living in an area (e.g., an area may be free of cases because it is not populated"
- does not show death rates (e.g., maybe more people lived near Broad Street pump?)

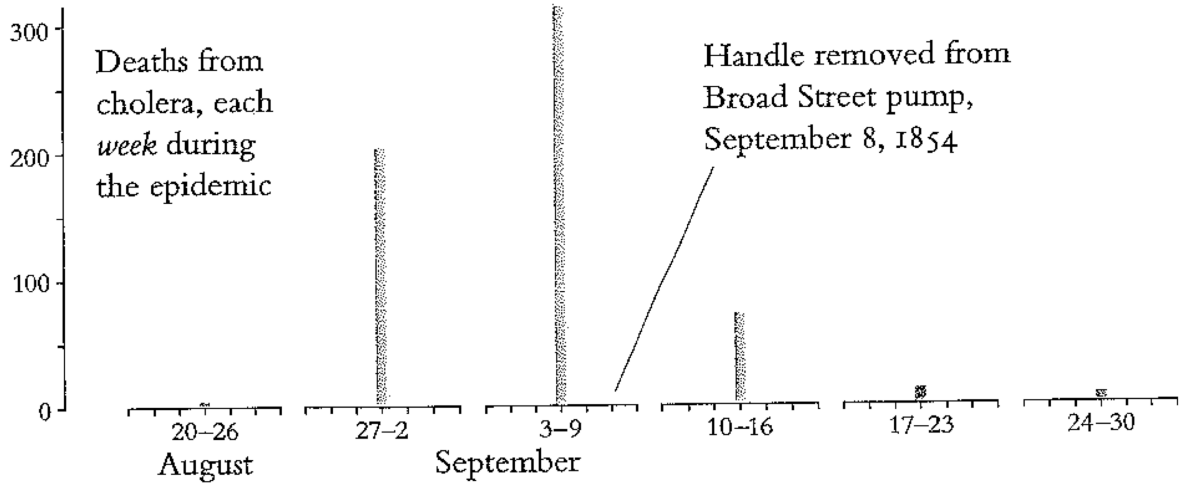
Different displays can lead to different conclusions, that is, the link between cause and effect





Different displays can lead to different conclusions, that is, the link between cause and effect

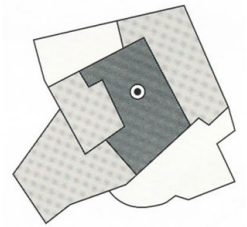


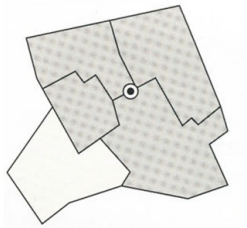


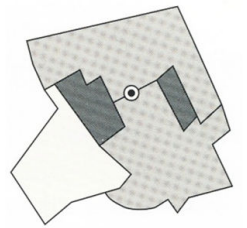
Lesson: How NOT to manipulate data

Mark Monmonier's *How to Lie with Maps*

aggregates of Snow's map:

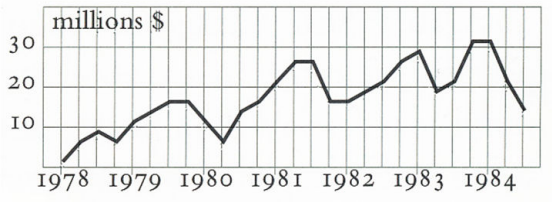




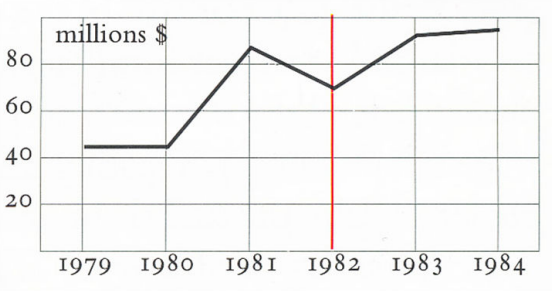


Gregory Joseph's *Modern Visual Evidence*

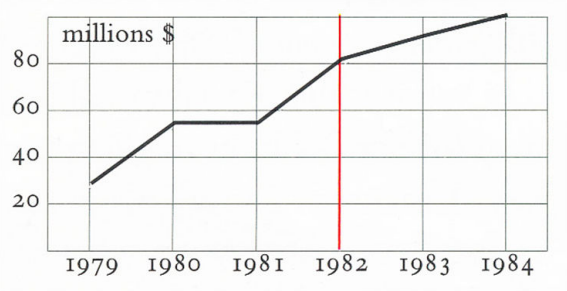
quarterly data



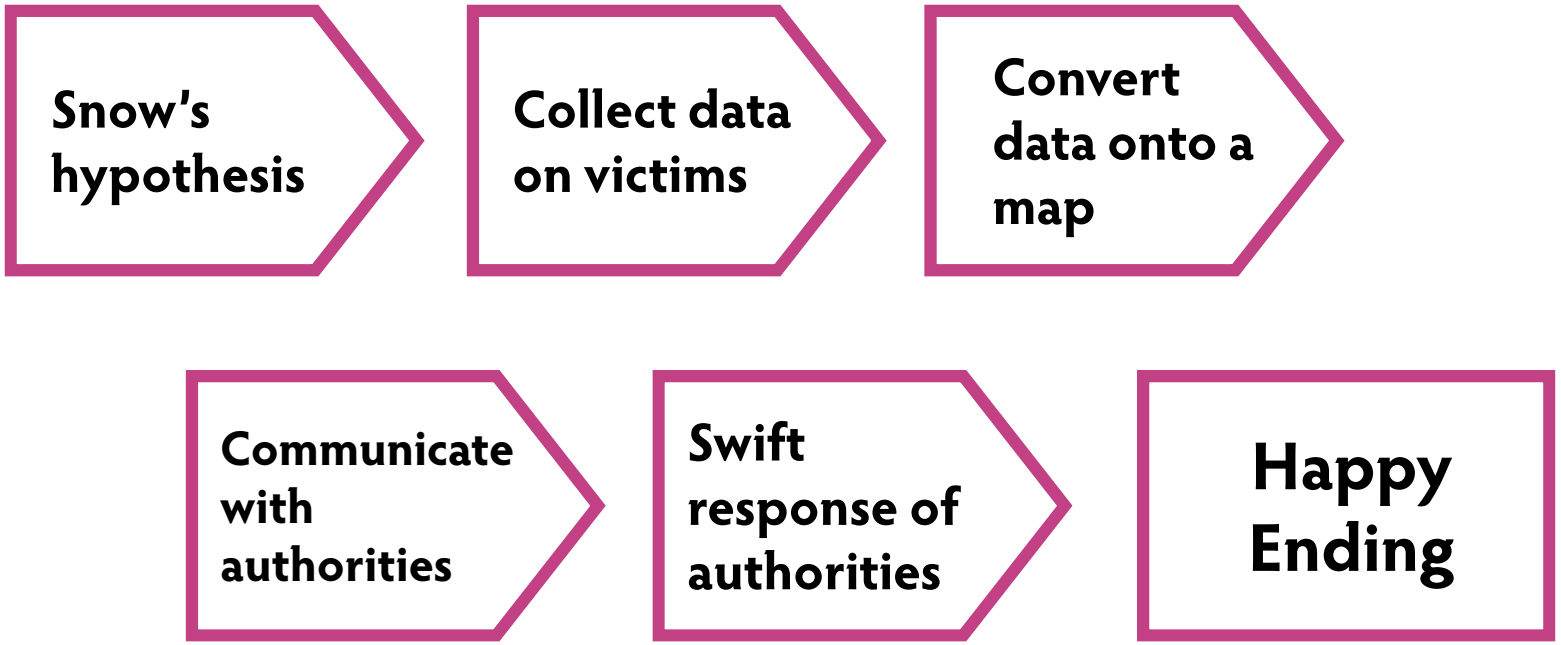
fiscal years



calendar years



Case 1: Hollywood Happy Ending



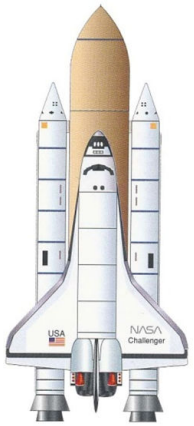
"For close upon 100 years we have been free in this country from epidemic cholera, and it is a freedom which, basically, we owe to the logical thinking, acute observations and simple sums of Dr. John Snow"

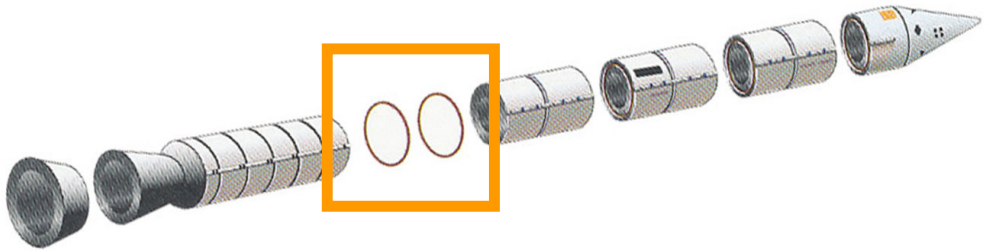
Bradford Hill
Proceedings of the Royal Society of Medicine, 1955

Case 2: Decision to Launch the Space Shuttle Challenger in January 1986

C

In the space shuttle, segments of the booster rockets are sealed with O-rings.
Previous launches showed damage to the O-rings.





C

All previous launches had occurred at temperatures of $\geq$53 °F.
Forecasted temperature of the launch was 26-29 °F.

Q

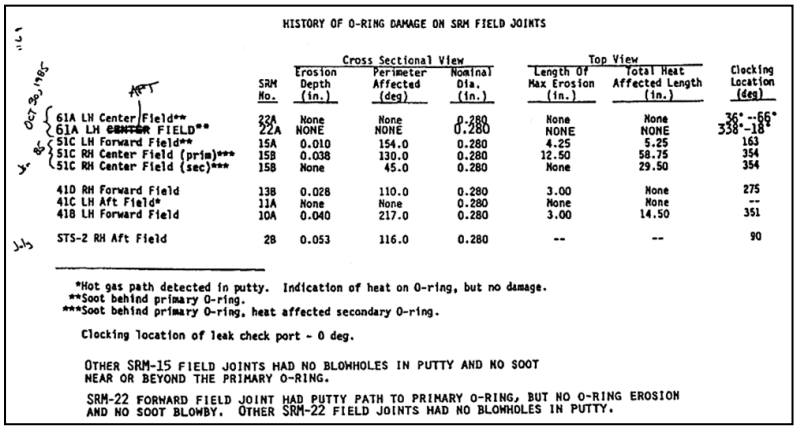
Will the O-rings maintain their seal at 26-29 °F?
Should the launch proceed?

H

Engineers at Morton Thiokol Inc (MTI): No, and then Yes
NASA officials: Yes

How did the engineers at Morton Thiokol Inc initially argue for their first decision?

- 13 slides were faxed from MTI to NASA



HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

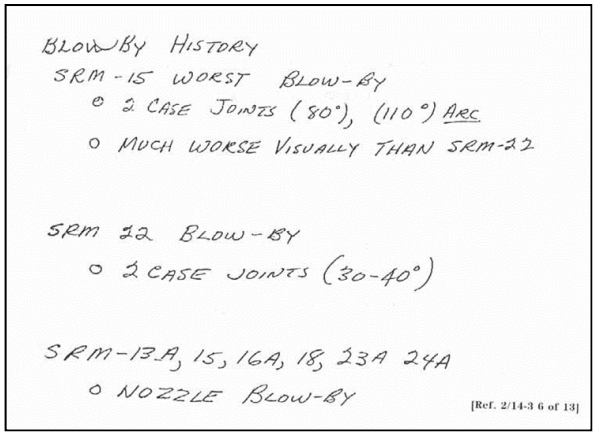
SRM No.	Cross Sectional View: Erosion Depth (in.)	Cross Sectional View: Perimeter Affected (deg)	Cross Sectional View: Nominal Dia. (in.)	Top View: Length Of Max Erosion (in.)	Top View: Total Heat Affected Length (in.)	Clocking Location (deg)
61A LH Center Field**	22A	None	None	0.280	None	None	36°-66°
61A LH ~~Center~~ FIELD**	22A	NONE	NONE	0.280	NONE	NONE	338°-18
51C LH Forward Field**	15A	0.010	154.0	0.280	4.25	5.25	163
51C RH Center Field (prim)***	15B	0.038	130.0	0.280	12.50	58.75	354
51C RH Center Field (sec)***	15B	None	45.0	0.280	None	29.50	354
41D RH Forward Field	13B	0.028	110.0	0.280	3.00	None	275
41C LH Aft Field*	11A	None	None	0.280	None	None	--
41B LH Forward Field	10A	0.040	217.0	0.280	3.00	14.50	351
STS-2 RH Aft Field	2B	0.053	116.0	0.280	--	--	90

*Hot gas path detected in putty. Indication of heat on O-ring, but no damage.
**Soot behind primary O-ring.
***Soot behind primary O-ring, heat affected secondary O-ring.

Clocking location of leak check port - 0 deg.

OTHER SRM-15 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY AND NO SOOT NEAR OR BEYOND THE PRIMARY O-RING.

SRM-22 FORWARD FIELD JOINT HAD PUTTY PATH TO PRIMARY O-RING, BUT NO O-RING EROSION AND NO SOOT BLOWBY. OTHER SRM-22 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY.



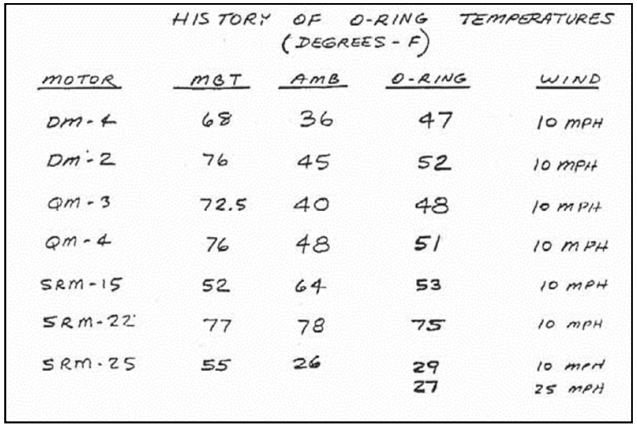
BLOW-BY HISTORY

SRM-15 WORST BLOW-BY
- 2 CASE JOINTS (80°), (110°) ARC
- MUCH WORSE VISUALLY THAN SRM-22

SRM 22 BLOW-BY
- 2 CASE JOINTS (30-40°)

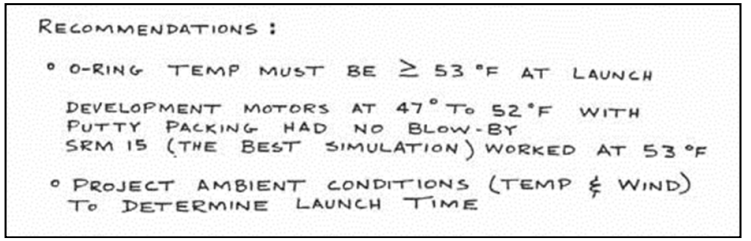
SRM-13A, 15, 16A, 18, 23A 24A
- NOZZLE BLOW-BY

[Ref. 2/14-3 6 of 13]



HISTORY OF O-RING TEMPERATURES (DEGREES-F)

MOTOR	MBT	AMB	O-RING	WIND
DM-1	68	36	47	10 MPH
DM-2	76	45	52	10 MPH
QM-3	72.5	40	48	10 MPH
QM-4	76	48	51	10 MPH
SRM-15	52	64	53	10 MPH
SRM-22	77	78	75	10 MPH
SRM-25	55	26	29	10 MPH
			27	25 MPH



RECOMMENDATIONS:

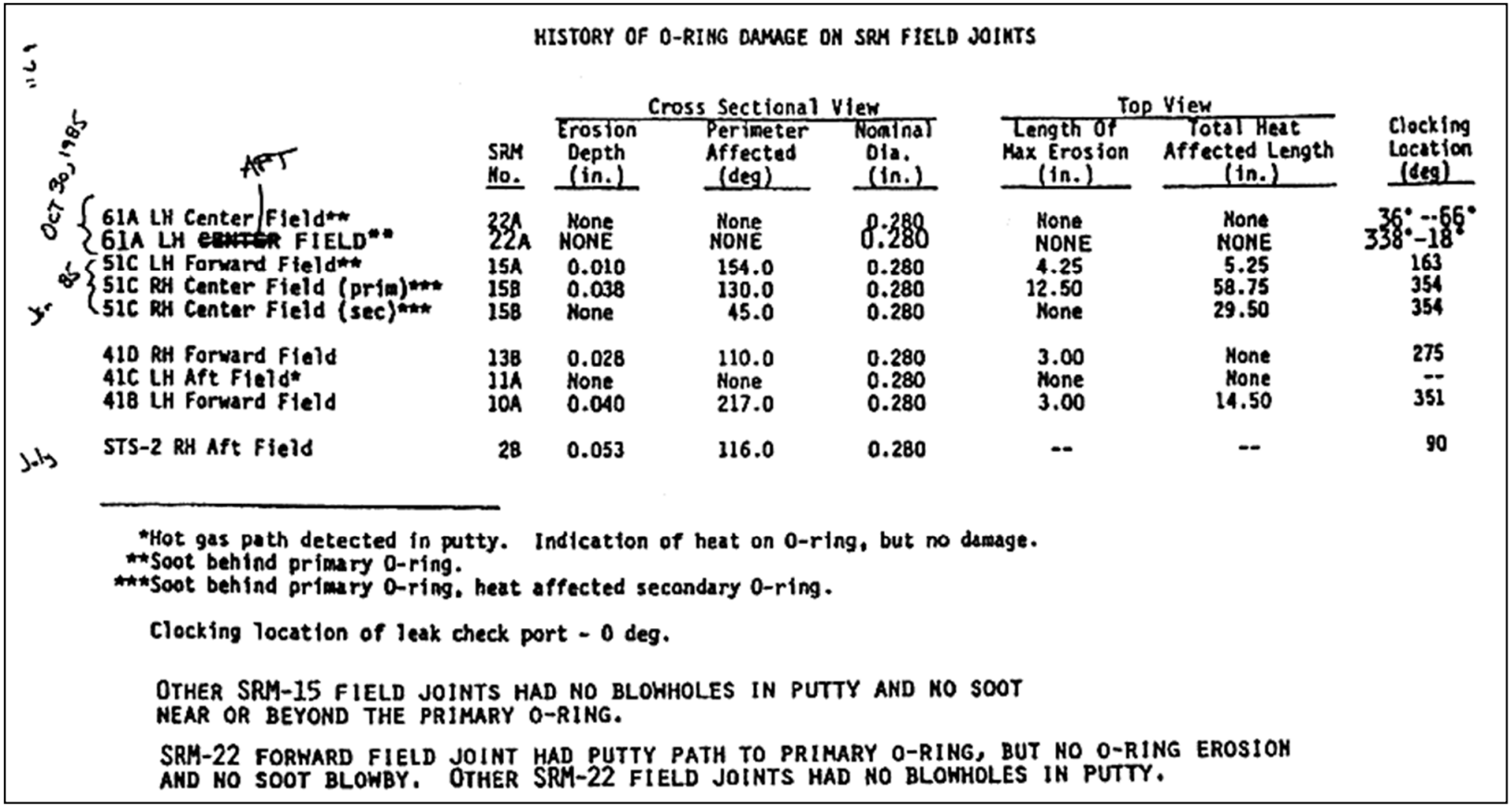
- O-RING TEMP MUST BE ≥ 53 °F AT LAUNCH

  DEVELOPMENT MOTORS AT 47° TO 52°F WITH PUTTY PACKING HAD NO BLOW-BY
  SRM 15 (THE BEST SIMULATION) WORKED AT 53 °F

- PROJECT AMBIENT CONDITIONS (TEMP & WIND) TO DETERMINE LAUNCH TIME

How did the engineers at Morton Thiokol Inc initially argue for their first decision?

- 13 slides were faxed from MTI to NASA



HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

	SRM No.	Cross Sectional View: Erosion Depth (in.)	Cross Sectional View: Perimeter Affected (deg)	Cross Sectional View: Nominal Dia. (in.)	Top View: Length Of Max Erosion (in.)	Top View: Total Heat Affected Length (in.)	Clocking Location (deg)
61A LH Center Field**	22A	None	None	0.280	None	None	36° --66°
61A LH CENTER FIELD**	22A	NONE	NONE	0.280	NONE	NONE	338°-18°
51C LH Forward Field**	15A	0.010	154.0	0.280	4.25	5.25	163
51C RH Center Field (prim)***	15B	0.038	130.0	0.280	12.50	58.75	354
51C RH Center Field (sec)***	15B	None	45.0	0.280	None	29.50	354
41D RH Forward Field	13B	0.028	110.0	0.280	3.00	None	275
41C LH Aft Field*	11A	None	None	0.280	None	None	--
41B LH Forward Field	10A	0.040	217.0	0.280	3.00	14.50	351
STS-2 RH Aft Field	2B	0.053	116.0	0.280	--	--	90

*Hot gas path detected in putty. Indication of heat on O-ring, but no damage.
**Soot behind primary O-ring.
***Soot behind primary O-ring, heat affected secondary O-ring.

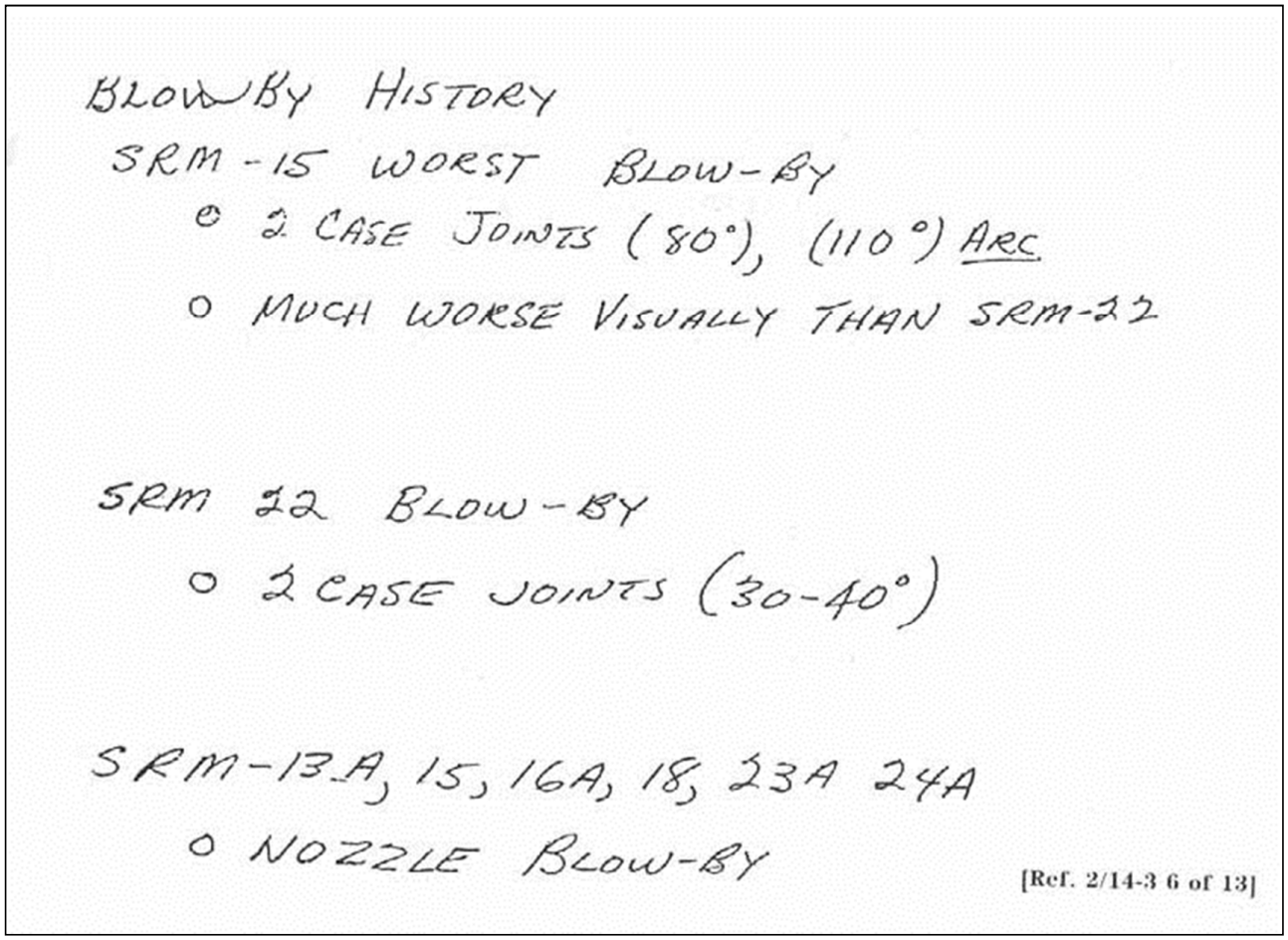
Clocking location of leak check port - 0 deg.

OTHER SRM-15 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY AND NO SOOT NEAR OR BEYOND THE PRIMARY O-RING.

SRM-22 FORWARD FIELD JOINT HAD PUTTY PATH TO PRIMARY O-RING, BUT NO O-RING EROSION AND NO SOOT BLOWBY. OTHER SRM-22 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY.

How did the engineers at Morton Thiokol Inc initially argue for their first decision?

- 13 slides were faxed from MTI to NASA



BLOW BY HISTORY

SRM-15 WORST BLOW-BY

- 2 CASE JOINTS (80°), (110°) ARC
- MUCH WORSE VISUALLY THAN SRM-22

SRM 22 BLOW-BY

- 2 CASE JOINTS (30-40°)

SRM-13A, 15, 16A, 18, 23A 24A

- NOZZLE BLOW-BY

[Ref. 2/14-3 6 of 13]

How did the engineers at Morton Thiokol Inc initially argue for their first decision?

- 13 slides were faxed from MTI to NASA

HISTORY OF O-RING TEMPERATURES
(DEGREES - F)

MOTOR	MBT	AMB	O-RING	WIND
DM-1	68	36	47	10 MPH
DM-2	76	45	52	10 MPH
QM-3	72.5	40	48	10 MPH
QM-4	76	48	51	10 MPH
SRM-15	52	64	53	10 MPH
SRM-22	77	78	75	10 MPH
SRM-25	55	26	29 27	10 MPH 25 MPH

NASA officials ask MTI to reconsider, and MTI reverses their original decision

MTI ASSESSMENT OF TEMPERATURE CONCERN ON SRM-25 (51L) LAUNCH

- CALCULATIONS SHOW THAT SRM-25 O-RINGS WILL BE 20° COLDER THAN SRM-15 O-RINGS
- TEMPERATURE DATA NOT CONCLUSIVE ON PREDICTING PRIMARY O-RING BLOW-BY
- ENGINEERING ASSESSMENT IS THAT:
  - COLDER O-RINGS WILL HAVE INCREASED EFFECTIVE DUROMETER ("HARDER")
  - "HARDER" O-RINGS WILL TAKE LONGER TO "SEAT"
    - MORE GAS MAY PASS PRIMARY O-RING BEFORE THE PRIMARY SEAL SEATS (RELATIVE TO SRM-15)
      - DEMONSTRATED SEALING THRESHOLD IS 3 TIMES GREATER THAN 0.038" EROSION EXPERIENCED ON SRM-15
  - IF THE PRIMARY SEAL DOES NOT SEAT, THE SECONDARY SEAL WILL SEAT
    - PRESSURE WILL GET TO SECONDARY SEAL BEFORE THE METAL PARTS ROTATE
      - O-RING PRESSURE LEAK CHECK PLACES SECONDARY SEAL IN OUTBOARD POSITION WHICH MINIMIZES SEALING TIME
- MTI RECOMMENDS STS-51L LAUNCH PROCEED ON 28 JANUARY 1986
  - SRM-25 WILL NOT BE SIGNIFICANTLY DIFFERENT FROM SRM-15

Joe C. Kilminster

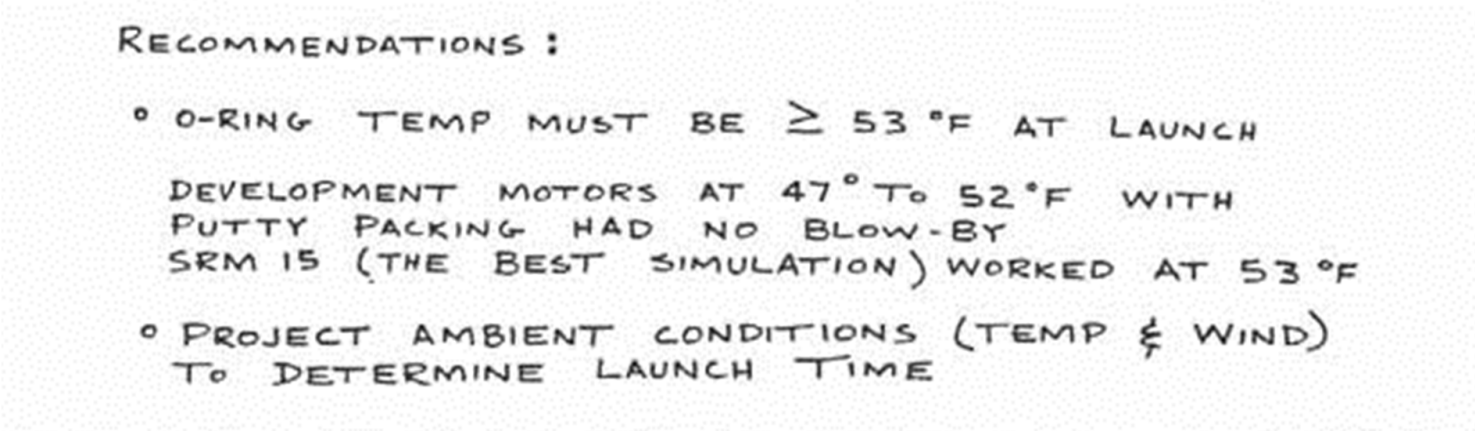
JOE C. KILMINSTER, VICE PRESIDENT
SPACE BOOSTER PROGRAMS

MORTON THIOKOL INC.
Wasatch Division

INFORMATION ON THIS PAGE WAS PREPARED TO SUPPORT AN ORAL PRESENTATION AND CANNOT BE CONSIDERED COMPLETE WITHOUT THE ORAL DISCUSSION

How did the engineers at Morton Thiokol Inc initially argue for their first decision?

- 13 slides were faxed from MTI to NASA



- How would you respond to this argument? Was this an effective argument?
- This was MTI's only no-launch recommendation in 12 years.
- A NASA official responded that he was "appalled" by MTI's recommendation not to launch.

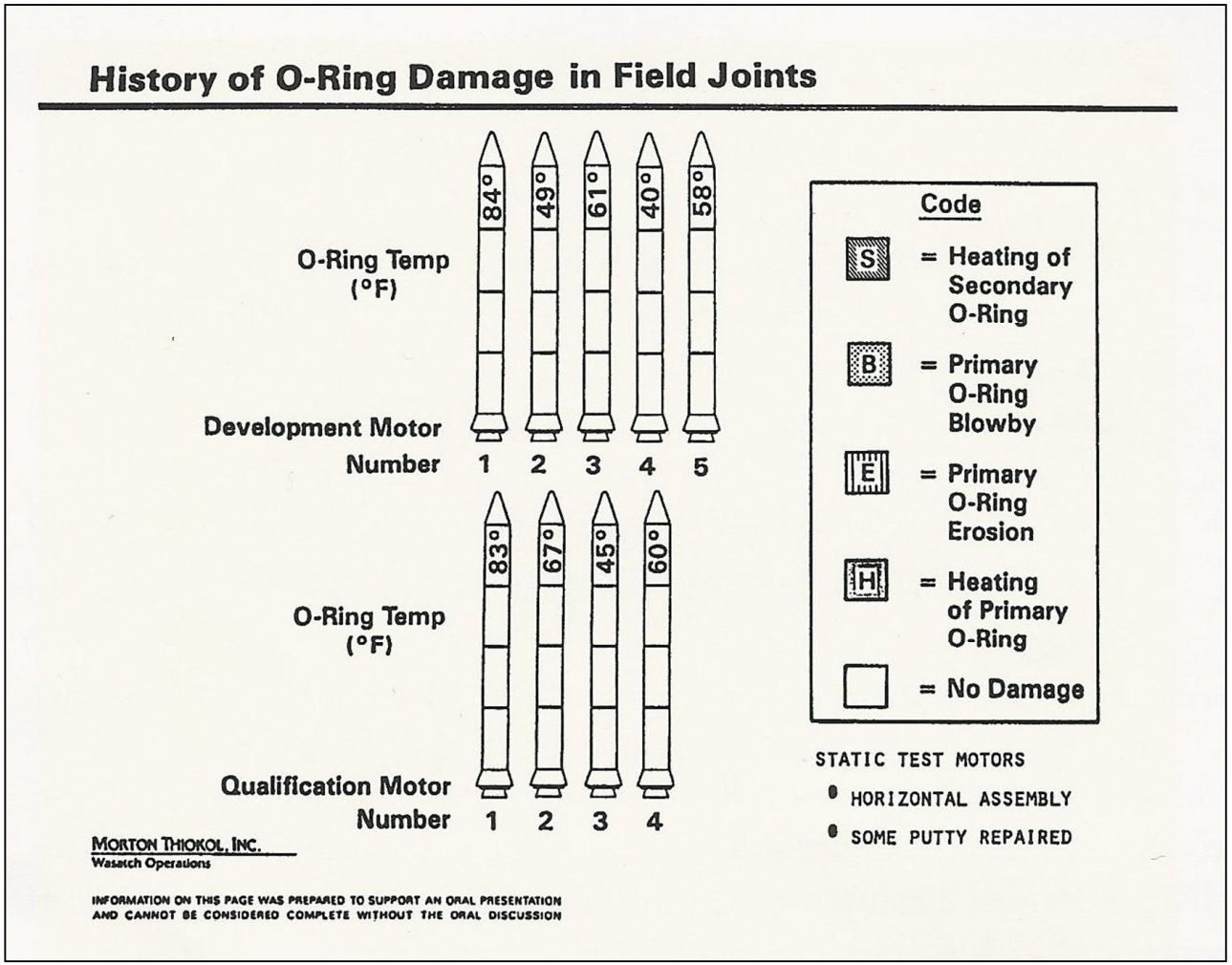
Post-Analysis

- MTI's engineers had originally reached the right conclusion, although with an ineffective argument.
- Commission investigating the accident:

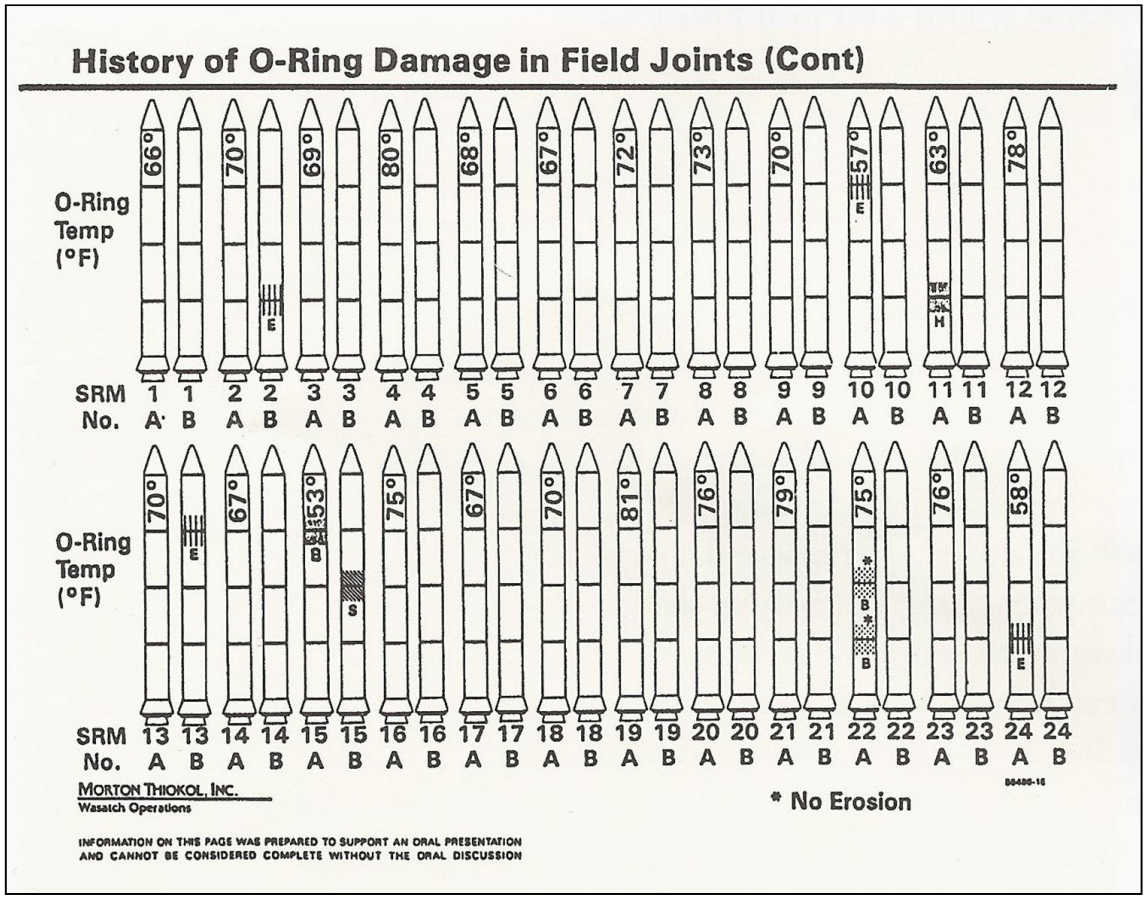
  "A careful analysis of the flight history of O-ring performance would have revealed the correlation of O-ring damage and low temperature. Neither NASA nor Thiokol carried out such an analysis; consequently, they were unprepared to properly evaluate the risks of launching the 51-L [Challenger] mission in conditions more extreme than they had encountered before."

- How might the data have been better analyzed, presented and communicated?

Attempt #1 shows a full analysis correlating temperature with damage to the O-rings



Attempt #1 shows a full analysis correlating temperature with damage to the O-rings



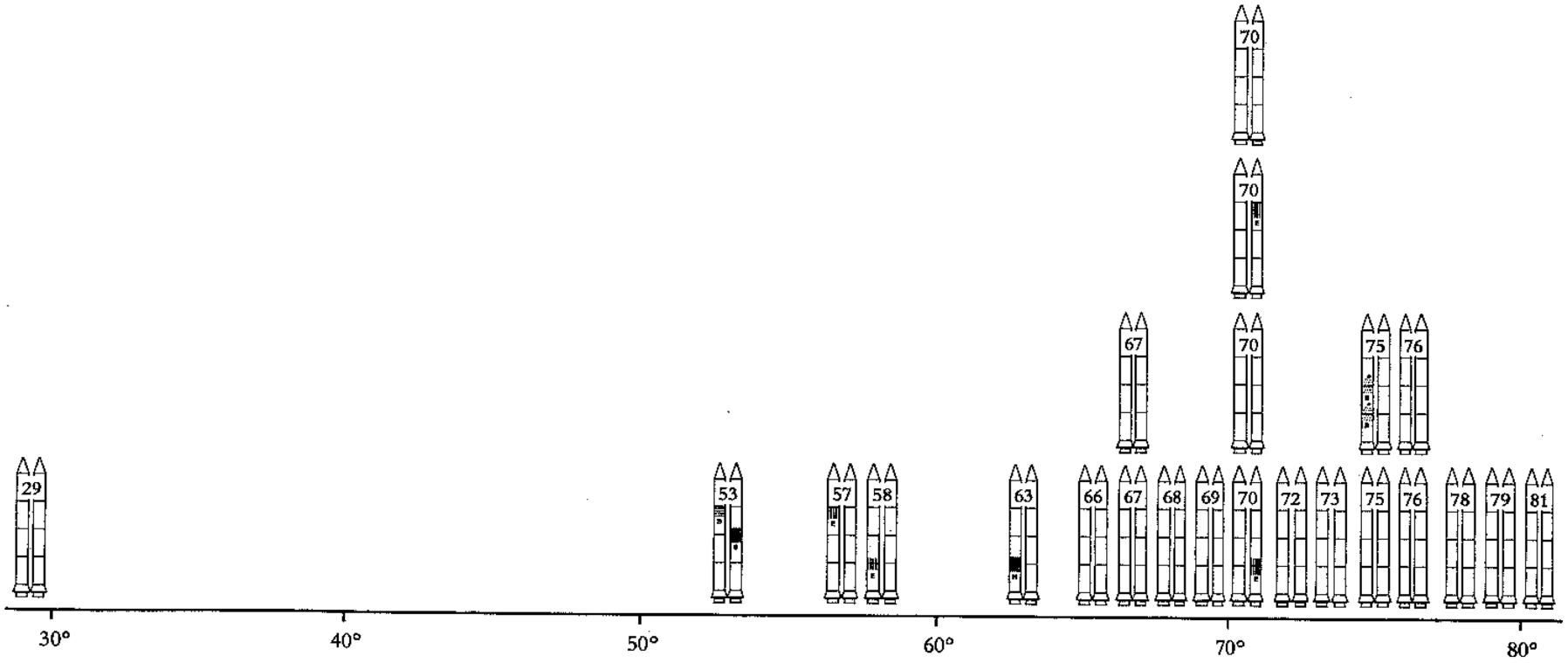
- What are the pro's and con's of this data display?
- Can it be improved?

Attempt #2: Tufte summarizes all data into a table with a "Damage Index"

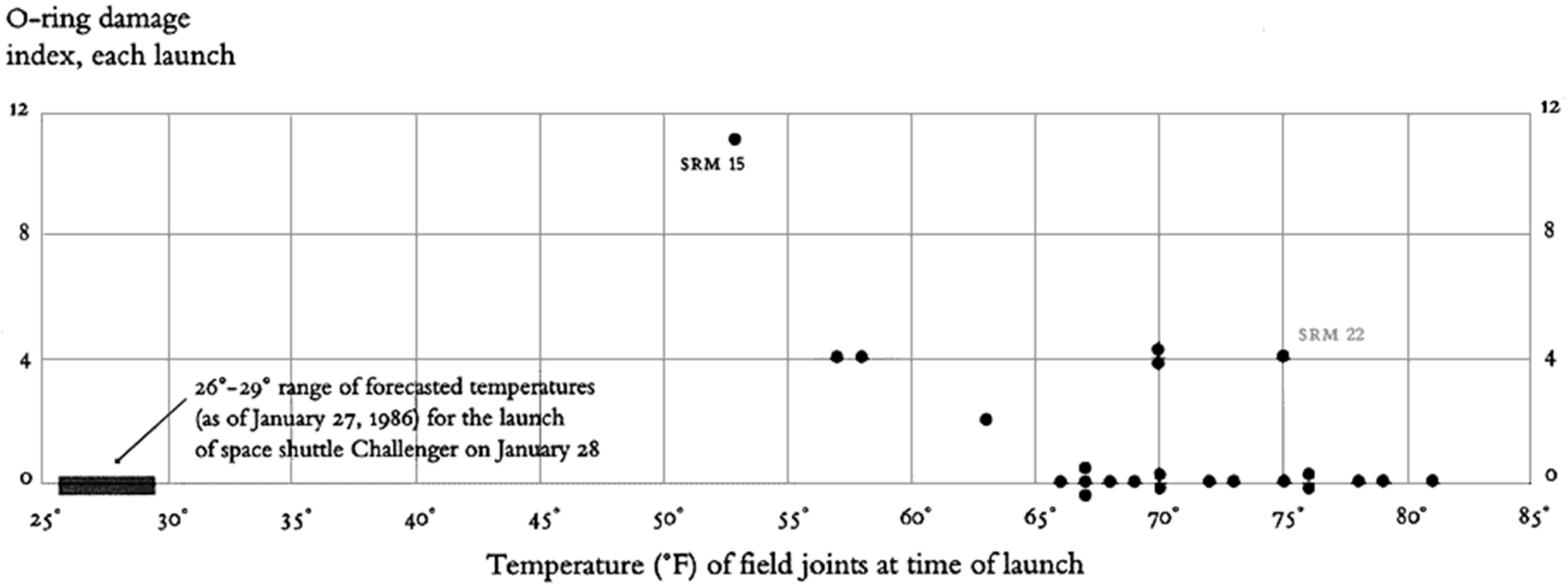
Flight	Date	Temperature °F	Erosion Incidents	Blow-by incidents	Damage Index	Comments
51-C	01.24.85	51°	3	2	11	Most erosion any flight; blow by; secondary rings heated
41-B	02.03.84	57°	1		4	Deep, extensive erosion
61-C	01.12.86	58°	1		4	O-ring erosion on launch two weeks before Challenger
41-C	04.06.84	63°	1		2	O-ring showed signs of heating, but no damage
1	04.12.81	66°			0	Coolest launch without O-ring problems
6	04.04.83	67°			0	
51-A	11.08.84	67°			0	
51-D	04.12.85	67°			0	
5	11.11.82	68°			0	
3	02.22.82	69°			0	
2	11.12.81	70°	1		4	Extent of erosion not fully known
9	11.28.83	70°			0	
41-D	08.30.84	70°			0	
51-G	06.17.85	70°	1		4	
7	06.18.83	72°			0	
8	08.30.83	73°			0	
51-B	04.29.85	75°		2	0	No erosion. Soot found behind two primary O-Rings
61-A	10.30.85	76°			0	
51-I	08.27.85	76°			0	
61-B	11.26.85	76°			0	
41-G	10.05.84	78°			0	
51-J	10.03.95	79°			0	
4	06.27.82	80°			?	O-ring condition unknown; rocket casing lost at sea
51-F	07.29.85	81°			0	

- What are the pro's and con's of this data display?
- Can it be improved?

Bad visual displays

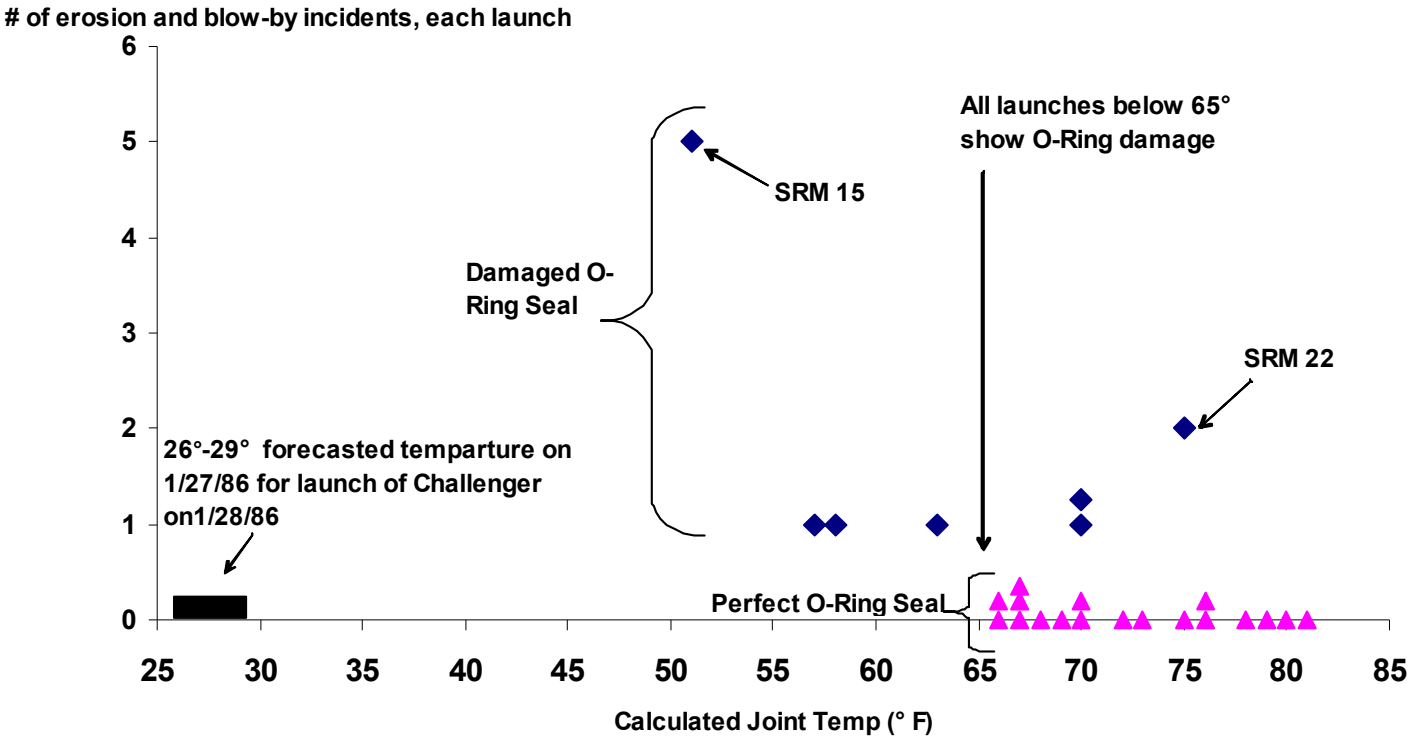


Attempt #2: Tufte summarizes all data into a graph with a "Damage Index"



- What are the pro's and con's of this data display?
- Can it be improved?

Attempt #3: Keller summarizes all data into a color graph



- What are the pro's and con's of this data display?
- Can it be improved?

Applying the 4 key tasks to the Challenger launch

1. Defining message	2. Choosing form	3. Creating design	4. Using software
What's the point of this display? What am I trying to communicate? What is my message? How do I make my message clear?	Should I use table, text, or graph, or a visual?	What design principles lead to quick cognitive processing and effective communication of the message?	How do I implement my ideas using software so that I control the software, and the software does not control the outcome?

Apply to Challenger Problem

Need to persuade mgmt. that low temperatures can cause O-ring damage	Table or graph to show relationship	Organize with complete dataset of events, ordered by temperature, ideally on one page	Excel scatter plot, with appropriate scale and highlights