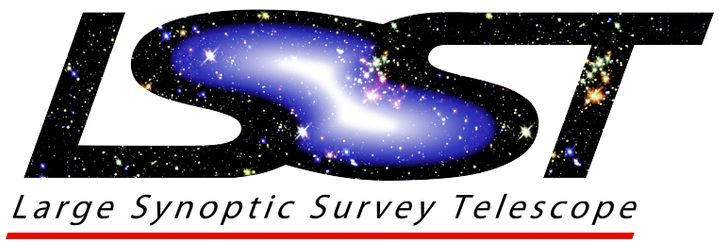
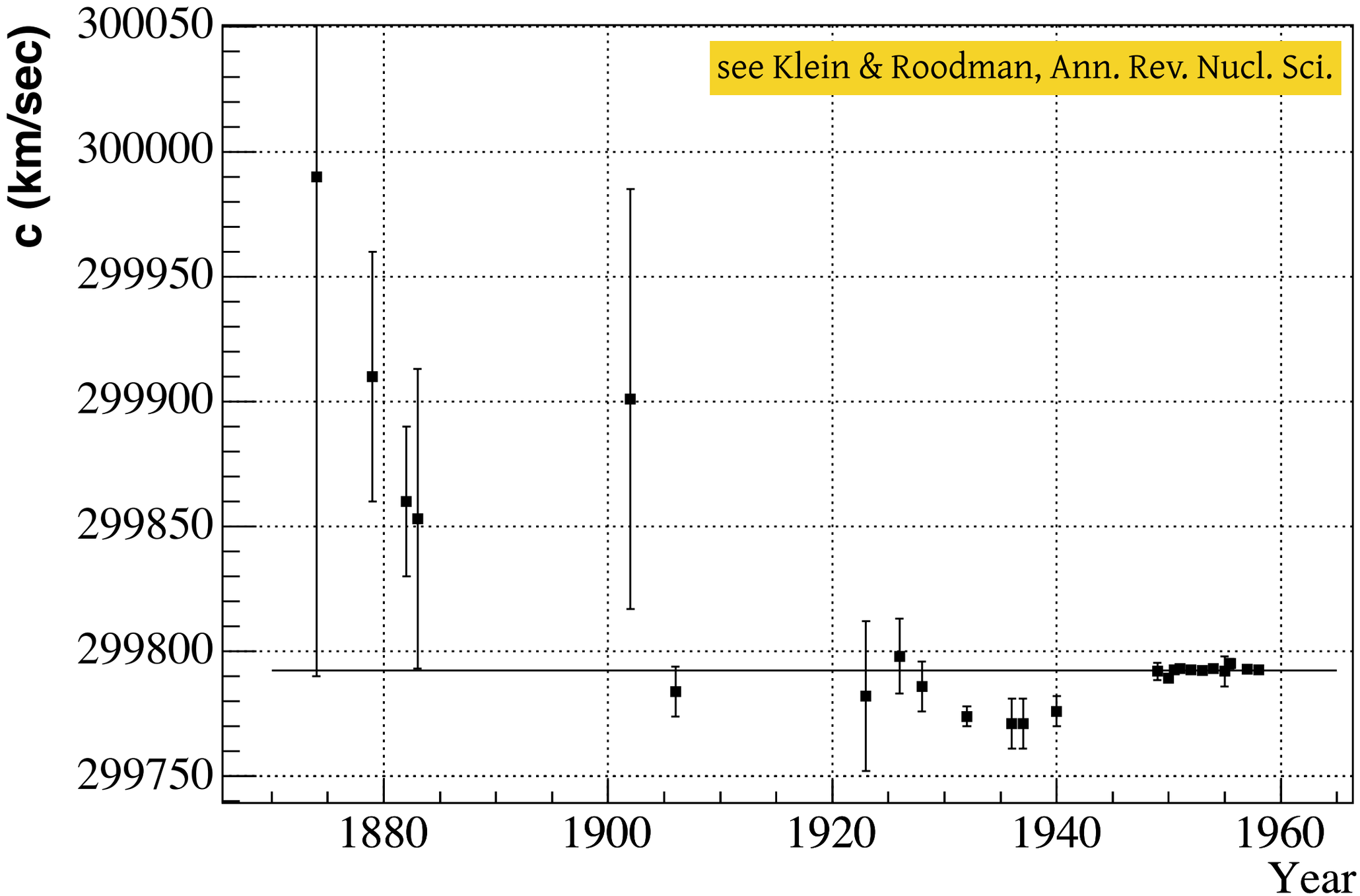
Introduction to Blind Analysis

Aaron Roodman

Blind Analysis in High-Stakes Survey Science: When, Why, and How?
March 13, 2017



History of the Speed of Light



Q1: What do you infer from this time history?

Experimenter's Bias

- *Bias:* A systematic distortion of an expected statistical result due to a factor not allowed for in its derivation. (OED)
- *Experimenter's Bias:* A Bias caused by those performing the Measurement.

Q2: How might Experimenter's Bias occur?

Experimenter's Bias

- *Bias:* A systematic distortion of an expected statistical result due to a factor not allowed for in its derivation. (OED)
- *Experimenter's Bias:* A Bias caused by those performing the Measurement.

Q2: How might Experimenter's Bias occur?

- Agreement with prior results, or theoretical predictions, guide:

Experimenter's Bias

- *Bias:* A systematic distortion of an expected statistical result due to a factor not allowed for in its derivation. (OED)
- *Experimenter's Bias:* A Bias caused by those performing the Measurement.

Q2: How might Experimenter's Bias occur?

- Agreement with prior results, or theoretical predictions, guide:
  - which data to use

Experimenter's Bias

- *Bias:* A systematic distortion of an expected statistical result due to a factor not allowed for in its derivation. (OED)
- *Experimenter's Bias:* A Bias caused by those performing the Measurement.

Q2: How might Experimenter's Bias occur?

- Agreement with prior results, or theoretical predictions, guide:
  - which data to use
  - what choices to make in the analysis

Experimenter's Bias

- *Bias:* A systematic distortion of an expected statistical result due to a factor not allowed for in its derivation. (OED)
- *Experimenter's Bias:* A Bias caused by those performing the Measurement.

Q2: How might Experimenter's Bias occur?

- Agreement with prior results, or theoretical predictions, guide:
  - which data to use
  - what choices to make in the analysis
  - when to stop & publish

Is Your Result Correct?

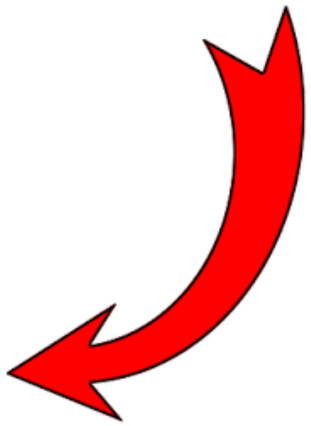
Think about a paper you are working on:

Q3: How do you know your results are correct?

Is Your Result Correct?

Think about a paper you are working on:

Q3: How do you know your results are correct?



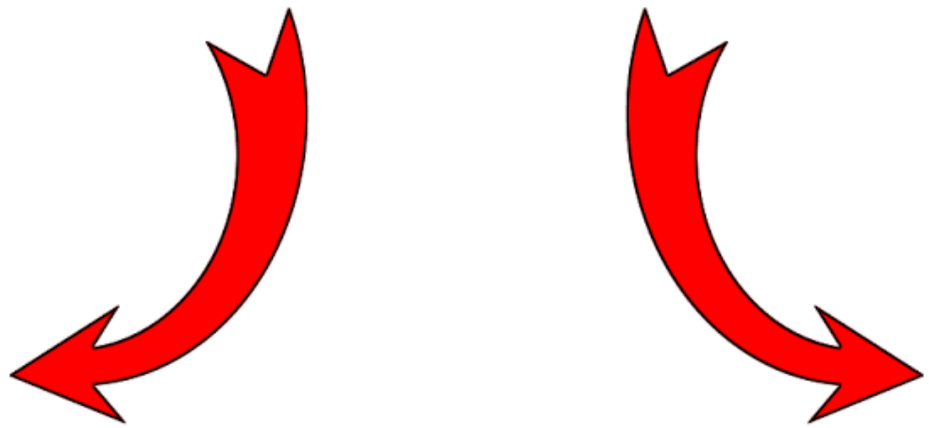
- Internal Consistency
- Pass Null Tests
- Physically sensible

Good

Is Your Result Correct?

Think about a paper you are working on:

Q3: How do you know your results are correct?



- Internal Consistency
- Pass Null Tests
- Physically sensible

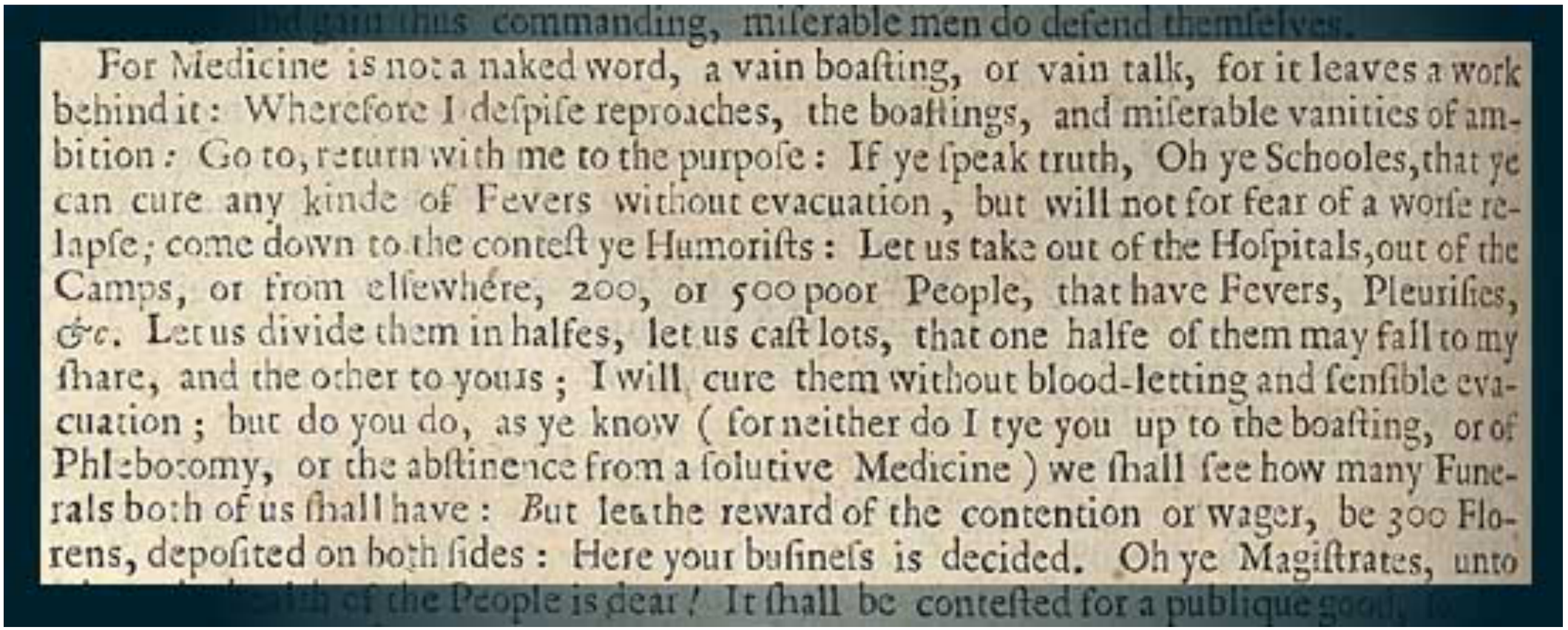
Good

- Agree w/ Previous Results
- Agree w/ Theory

Biased

Blind Analysis

- *Blind Analysis:* A method that hides some aspect of the data or result to prevent experimenter's bias.



commanding, miſerable men do defend themſelves.

For Medicine is not a naked word, a vain boaſting, or vain talk, for it leaves a work behind it: Wherefore I deſpiſe reproaches, the boaſtings, and miſerable vanities of ambition: Go to, return with me to the purpoſe: If ye ſpeak truth, Oh ye Schooles, that ye can cure any kinde of Fevers without evacuation, but will not for fear of a worſe relapſe; come down to the conteſt ye Humoriſts: Let us take out of the Hoſpitals, out of the Camps, or from elſewhere, 200, or 500 poor People, that have Fevers, Pleuriſies, &c. Let us divide them in halfes, let us caſt lots, that one halfe of them may fall to my ſhare, and the other to yours; I will cure them without blood-letting and ſenſible evacuation; but do you do, as ye know (for neither do I tye you up to the boaſting, or of Phlebotomy, or the abſtinence from a ſolutive Medicine) we ſhall ſee how many Funerals both of us ſhall have: But let the reward of the contention or wager, be 300 Florens, depoſited on both ſides: Here your buſineſs is decided. Oh ye Magiſtrates, unto

of the People is dear! It ſhall be conteſted for a publique good, to

van Helmont 1662 from R. Doll, British Medical Journal, 1998.

Blind Analysis

- *Blind Analysis:* A method that hides some aspect of the data or result to prevent experimenter's bias.

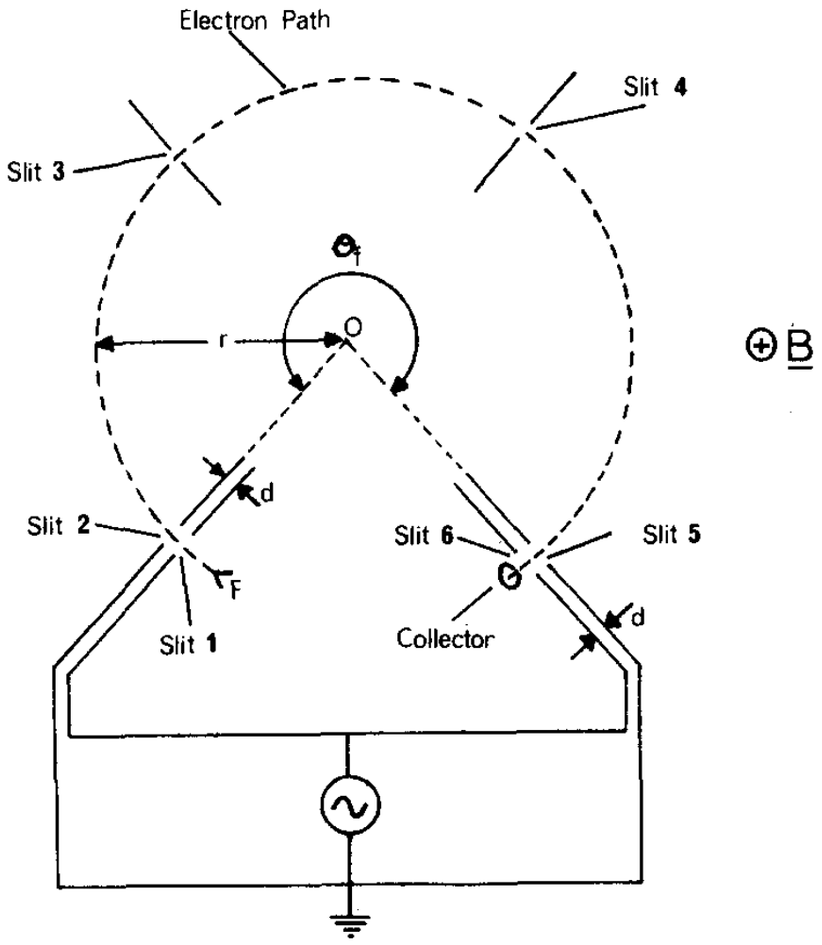


Fig. 1. Schematic diagram of apparatus.

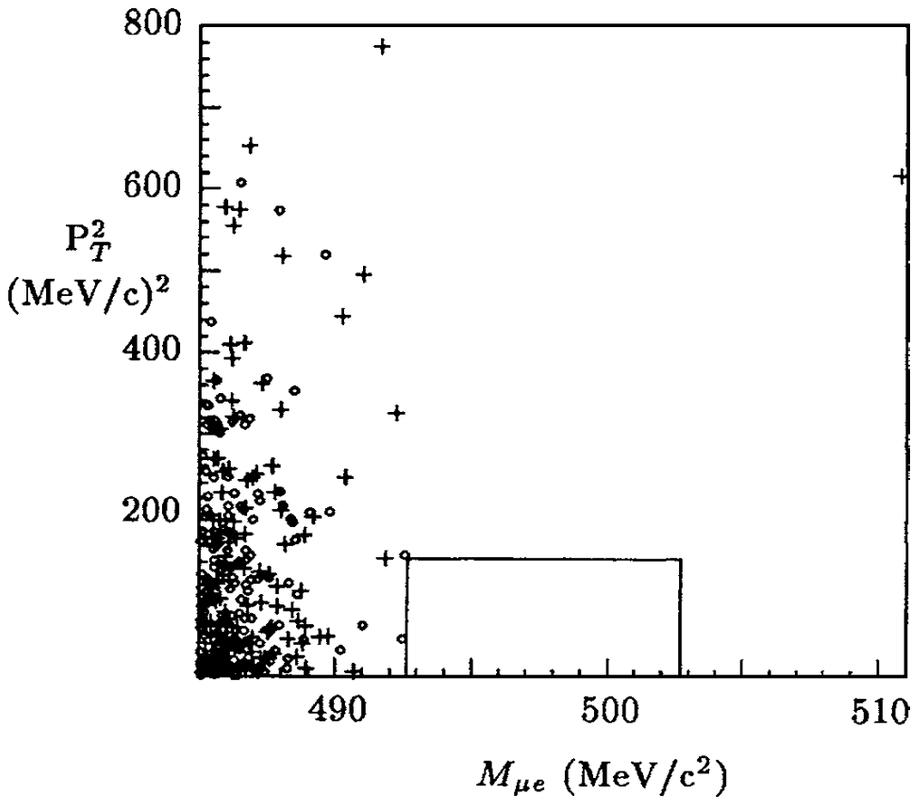
Dunnington, Electron e/m, 1933

$$e/m = \theta v/B_0.$$

exact value of θ chosen by Dunnington's machinist and kept hidden until the data analysis was complete

Blind Analysis Methods

- *Hidden Signal Box:*



- *Hidden Answer:*

$$\epsilon'/\epsilon(\text{Hidden}) = \left\{ \begin{array}{c} 1 \\ -1 \end{array} \right\} \times \epsilon'/\epsilon + C$$

sign and C: chosen by computer and hidden

- *Unblind sub-sample*

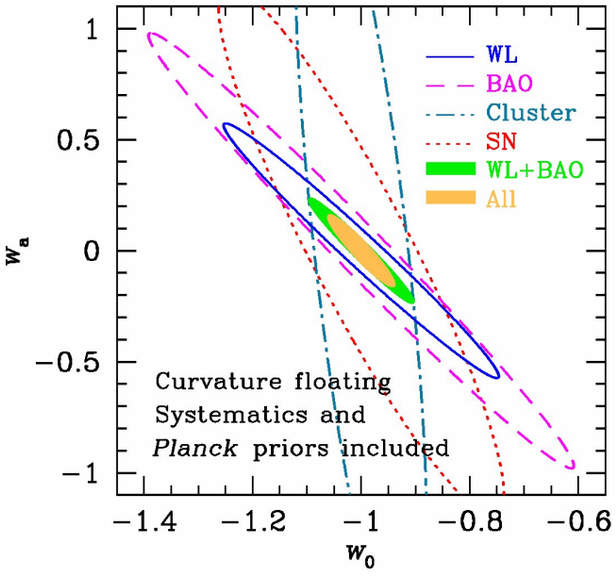
fraction of data studied unblind

- *Fake Signal*

fake signal added to data

High Stakes Survey Science

- Era of Precision Cosmology
  - Hubble Parameter
  - Baryon, Dark Matter & Dark Energy Density
  - Dark Energy Equation of State
  - Neutrino Masses
  - etc...



- Strong Experimental and Theoretical priors
- Ultimate results from just a few data sets
- Future replication may be difficult

Large Synoptic Survey Telescope

euclid

⇒ High Stakes Survey Science

Blind Analysis: Caveats & Cautions

Blind Analysis: Caveats & Cautions

- Blind Analysis may not always be possible:
  - well suited to measurements of precision parameters & searches with well understood background

Blind Analysis: Caveats & Cautions

- Blind Analysis may not always be possible:
  - well suited to measurements of precision parameters & searches with well understood background
- Blind Analysis should be as simple as possible:
  - blind only what is absolutely necessary
  - data checks & null tests must be possible

Blind Analysis: Caveats & Cautions

- Blind Analysis may not always be possible:
  - well suited to measurements of precision parameters & searches with well understood background
- Blind Analysis should be as simple as possible:
  - blind only what is absolutely necessary
  - data checks & null tests must be possible
- Many different Blind Analysis techniques have been invented:
  - custom designed for measurement

Blind Analysis: Caveats & Cautions

- Blind Analysis may not always be possible:
  - well suited to measurements of precision parameters & searches with well understood background
- Blind Analysis should be as simple as possible:
  - blind only what is absolutely necessary
  - data checks & null tests must be possible
- Many different Blind Analysis techniques have been invented:
  - custom designed for measurement
- Blind Analysis does take more time:
  - more time will be spent checking the analysis without the crutch of checking the final value

Goals for this Workshop

Q4: What do you want to take away from this workshop?

Goals for this Workshop

Q4: What do you want to take away from this workshop?

What Blind Analysis Method Will Be Effective For Your Measurement?