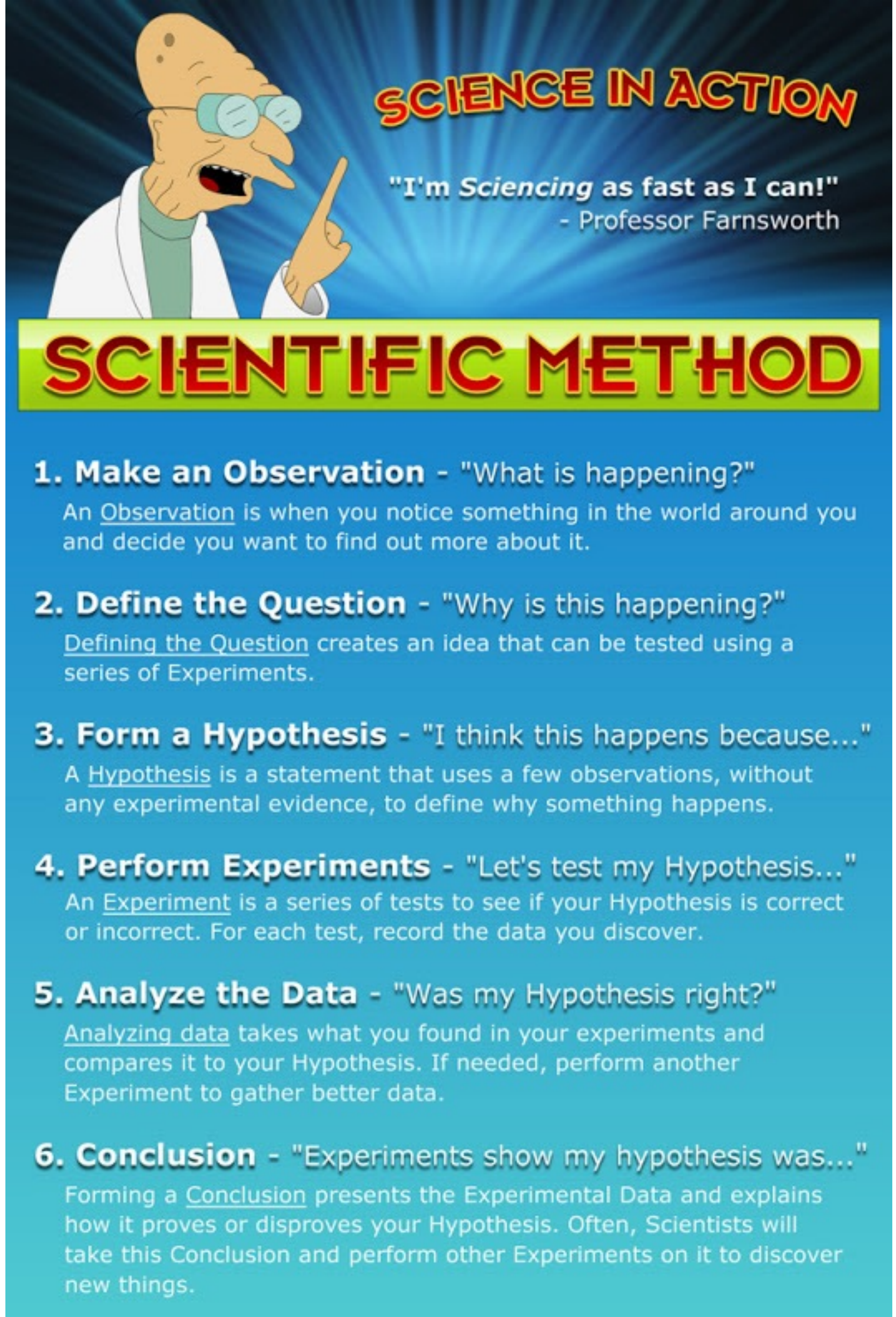
PH125 lab 1

uncertainty analysis & statistics

Sciencing

- So you have an idea.
- This idea must be *testable* ... or it is not science
- So we test it.

- How good is our test? How well did it work?
  - a measure of the result & accuracy
  - does it make any sense? predict something else ...



flickr user @neotsn https://www.flickr.com/photos/neotsn/4947989561/

Example

- Does a heavier or lighter object fall faster?
- can't do just one experiment
  - by "chance" one will land first
- do it many times
  - what is the variability?
  - how much is too much?
  - at what point are they basically the same?

Measurements

- we don’t do just one
- make a series of measurements, average them
- this should improve accuracy, right?
- how much? when to stop?

- need to quantify degree of uncertainty

Today's lab: counting cards

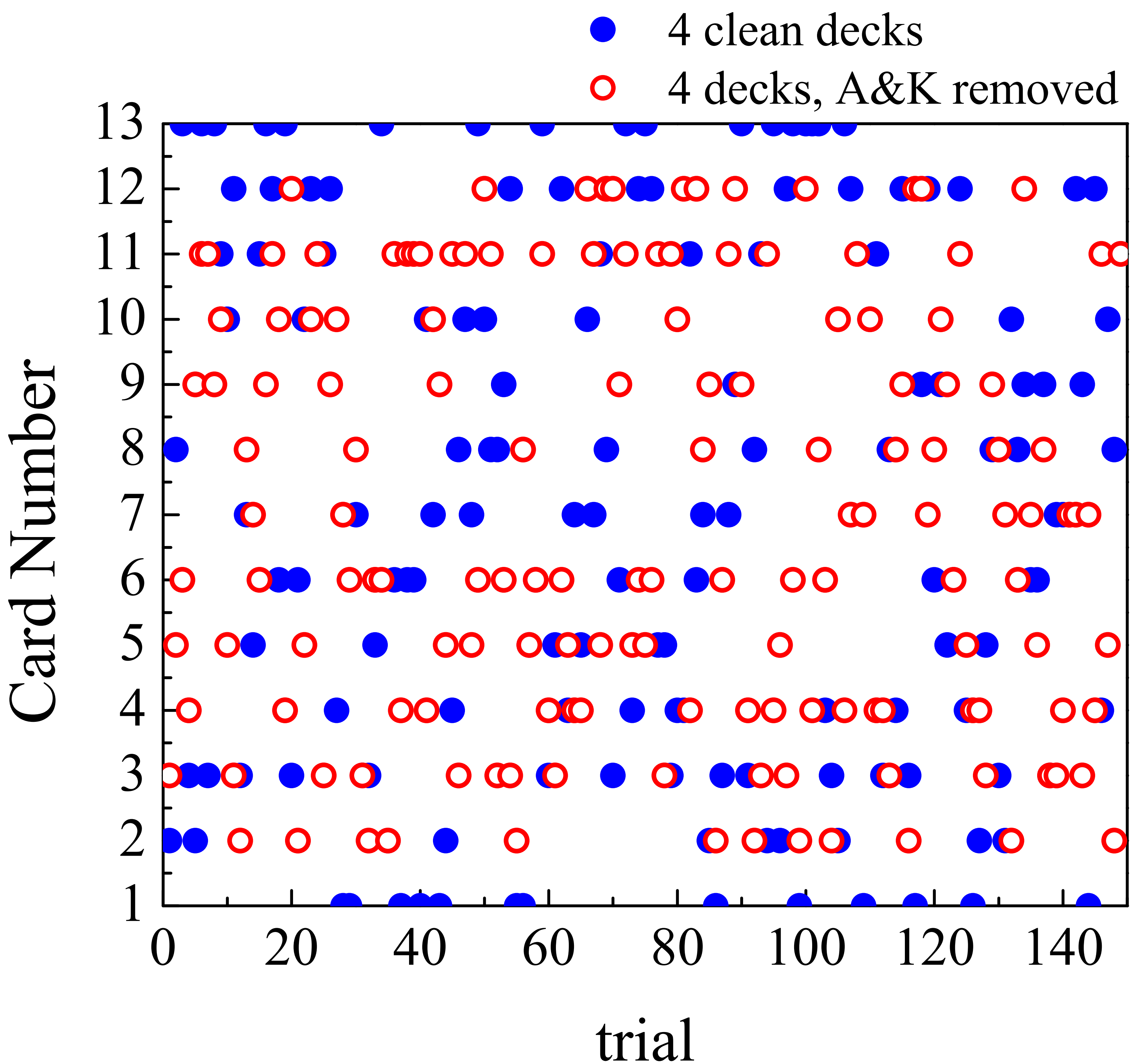
- one measurement vs. many
- how does accuracy improve?
- how to measure accuracy?
- statistical measures of uncertainty & dispersion

- if you don't see the whole deck at once, what can you still say?

The experiment: picking cards

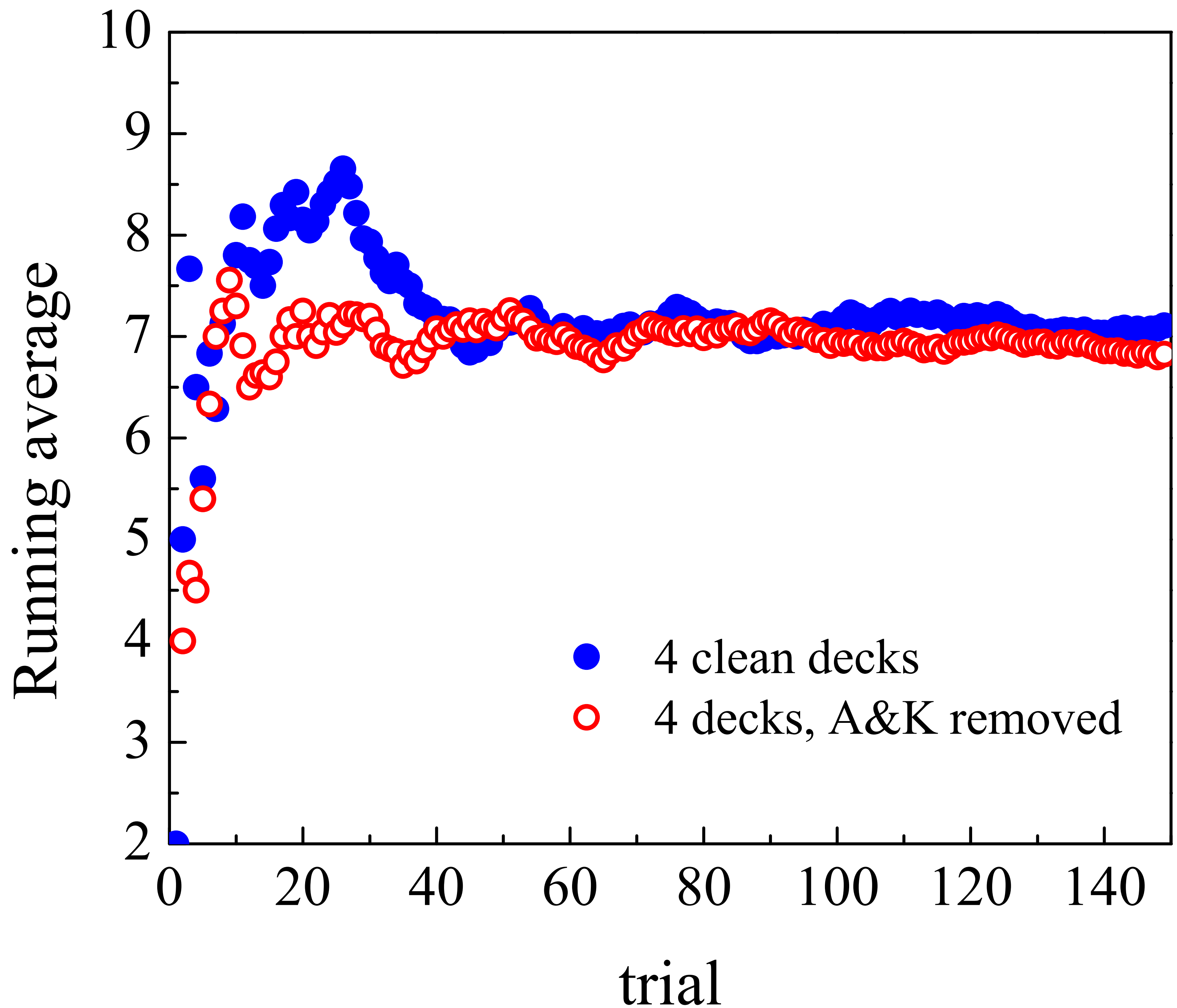
- give each one a number
- Ace = 1, 2 = 2 ... Jack = 11 ... King = 13
- what is the average card?
  - we expect it must be 7 ...
- what is the spread? how to define this?

150 trials ...



4 clean decks
4 decks, A&K removed
Card Number
trial
13
12
11
10
9
8
7
6
5
4
3
2
1
0
20
40
60
80
100
120
140

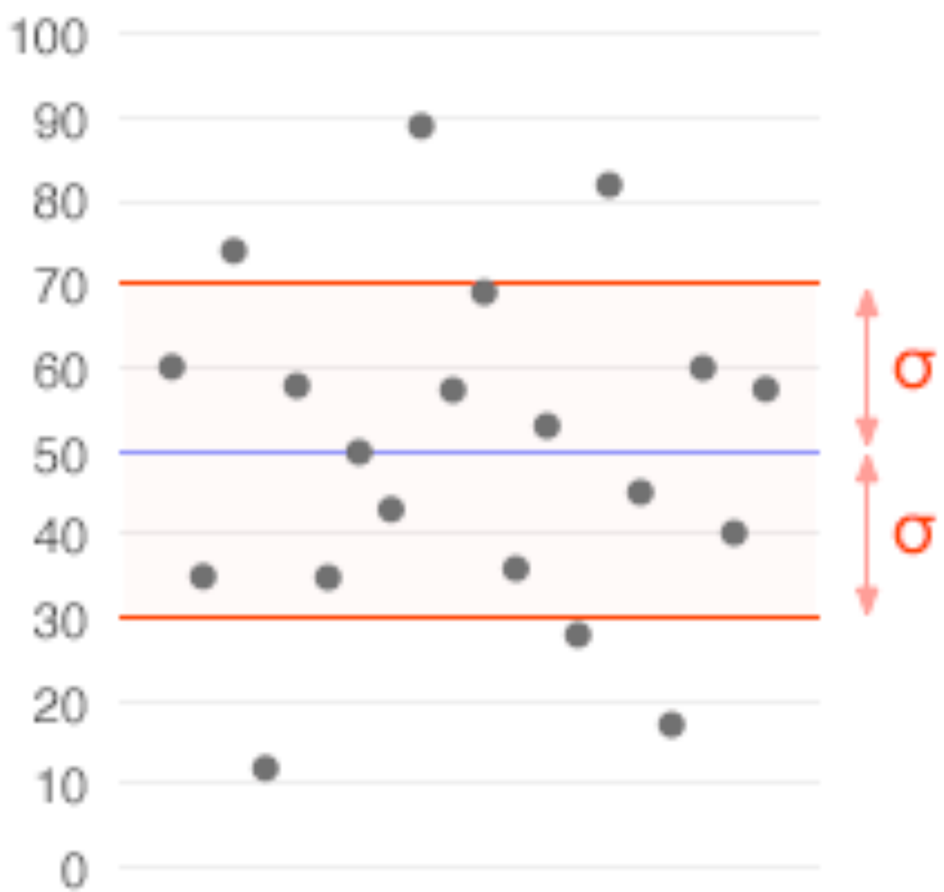
draw 1 card, record, shuffle, repeat
cards have values 1-13, equal number of each
average must be 7, if one chooses enough cards
takes ~50 before 'luck' is irrelevant!



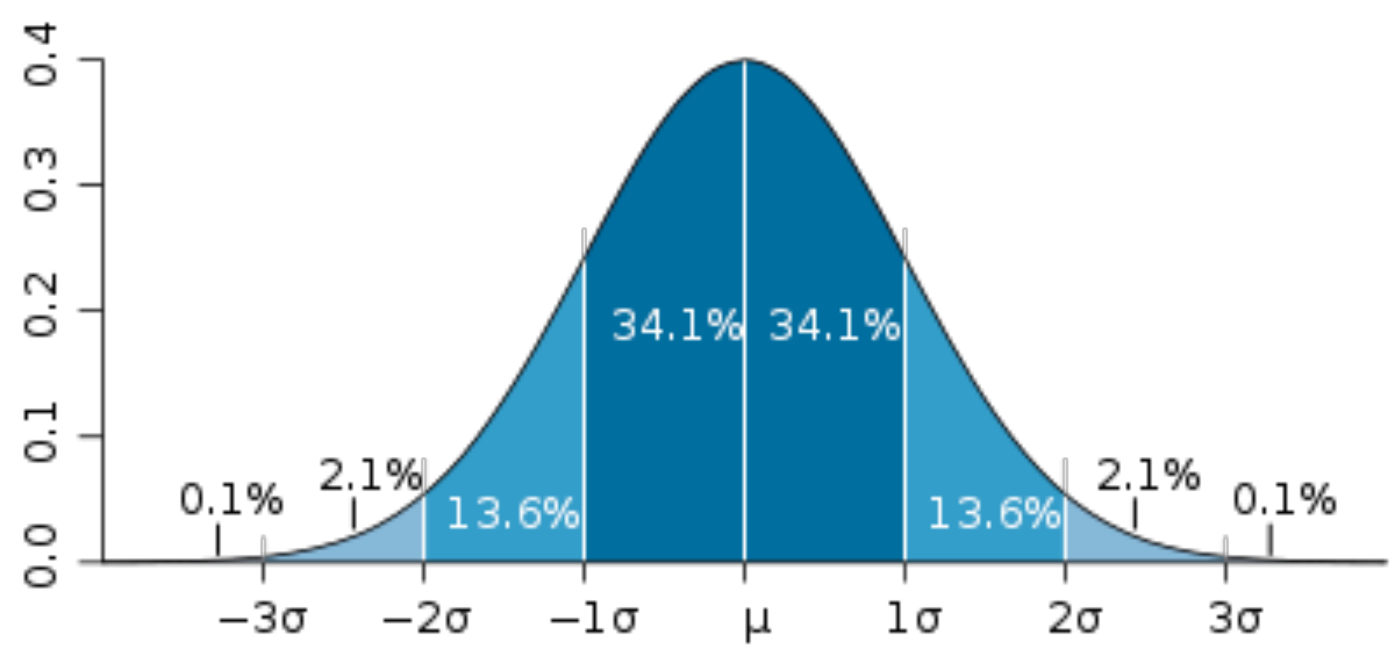
The mean isn’t enough.
how about the dispersion?

standard deviation is a measure of the variability dispersion in a population or data set

low standard deviation: data tends to lie close to the average (mean)
high standard deviation: data spread over a large range



data set: data clustered about average



many trials: follow a *distribution*

~68% within +/- 1 standard deviation
~95% within +/- 2 standard deviations
~99.7% within +/- 3 ...

so what?

- knowing the standard deviation tells you
  - if subsequent measurements are outliers
  - what to expect next, on average
  - *accuracy* of a set of data
  - variability in a large batch
- “six sigma” - quality control
  - means one out of 500 million!

so what?

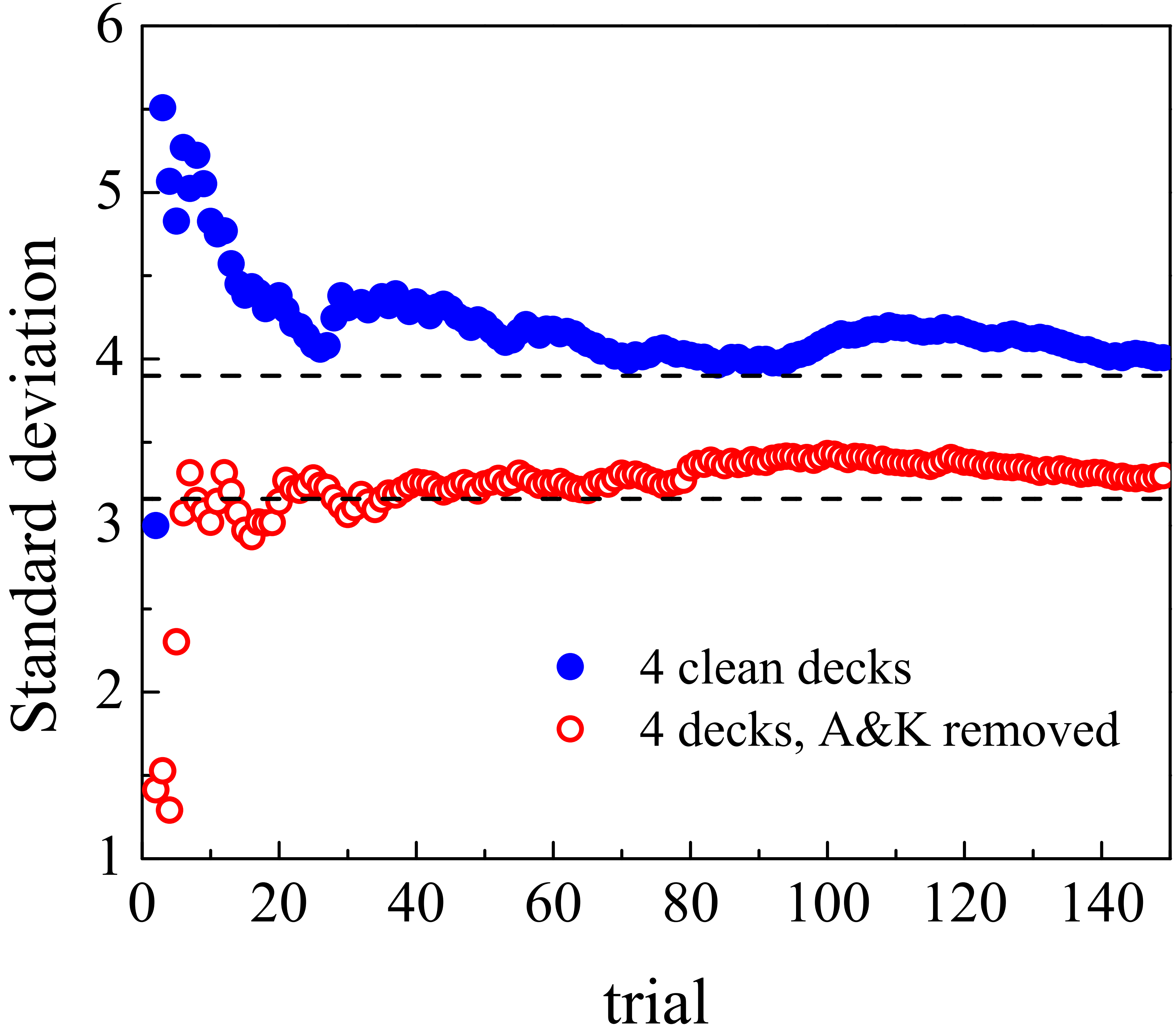
if the mean of the measurements is too far away from the prediction compared to the standard deviation, then the theory being tested probably needs to be revised!

particle physics: 5-sigma standard typical
1 out of 1.7M chance of false positive

larger signal than that ... probably a new effect!

for the cards?

take out highest and lowest cards,
data is more tightly distributed
lower standard deviation!



wait, there's more

expect 75% of cards within 2 standard deviations of average

or, 75% are within about 4 cards from the average after 100 trials

or, 75% of cards should be between 3 and Jack (inclusive)

It works!

flip side: we could estimate the distribution of cards without prior knowledge

- from a run of cards, if long enough, could say what's missing or in excess
- could take many samples of the deck though (e.g., removing all 7s)

what else?

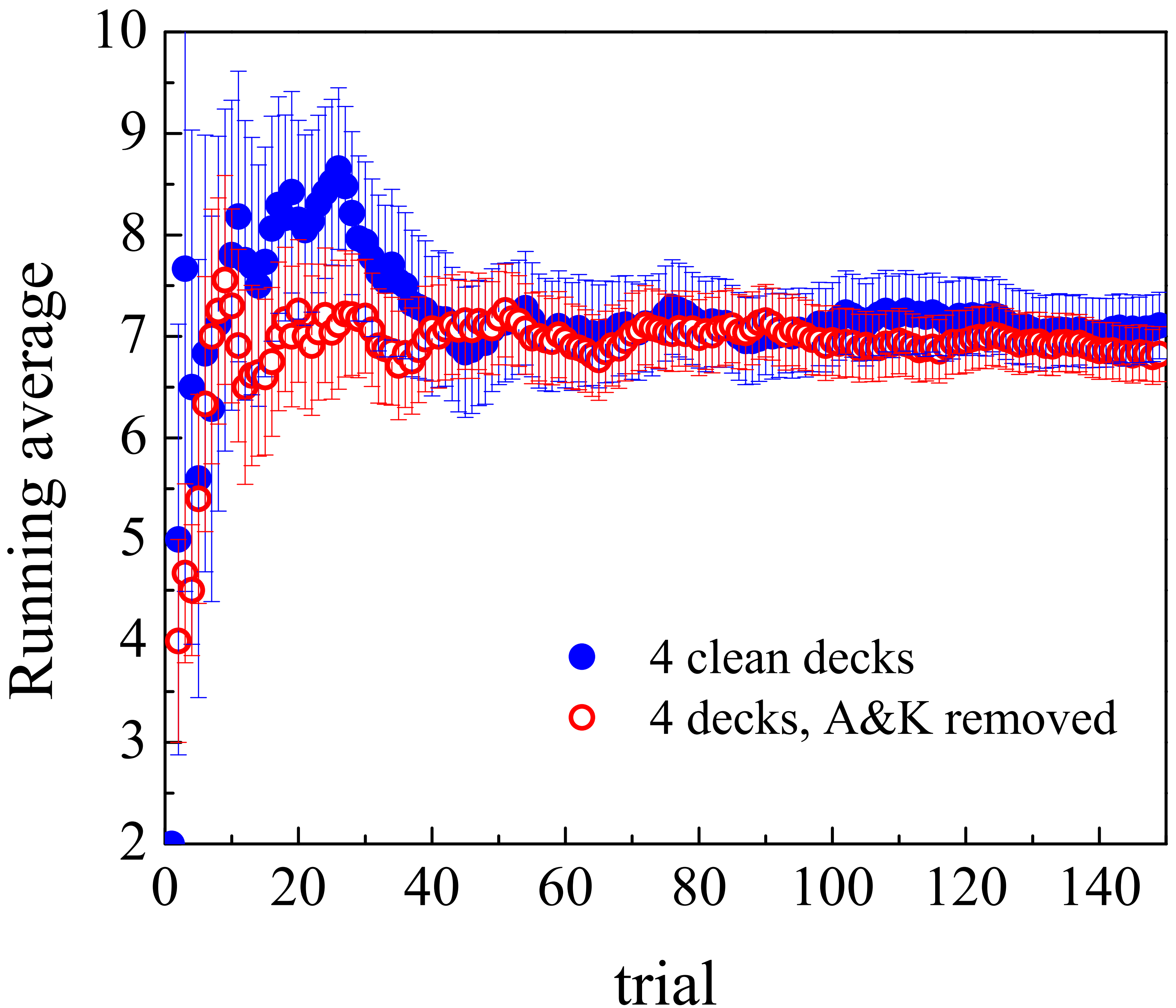
- standard deviation gives accuracy of averages
- if you do n measurements, average is more accurate for higher n. makes sense!
- uncertainty of the average is standard deviation over the root of the number of measurements

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

$$(\text{best value of } x) = \overline{x} \pm \sigma_{\overline{x}}$$

Now we can add error bars to our running averages

- error bars are large where N is small
- more data - more accurate average
- initial differences are not significant!



detailed explanation & examples in lab procedure

... so read it first. really.

excel example included too

draw	card	running average	running standard deviation	
1	2	2.00		
2	8	5.00	4.24	
3	13	7.67	5.51	
4	3	6.50	5.07	
5	2	5.60	4.83	

	A	B	C	D	E
1	draw	card	running average	running standard deviation	
2	1	2	2.00		
3	2	8	=AVERAGE($B$2:B3)	=STDEV($B$2:B3)	
4	3	13	^ drag down	^ drag down	
5	4	3	6.50	5.07	
6	5	2	5.60	4.83	
7					