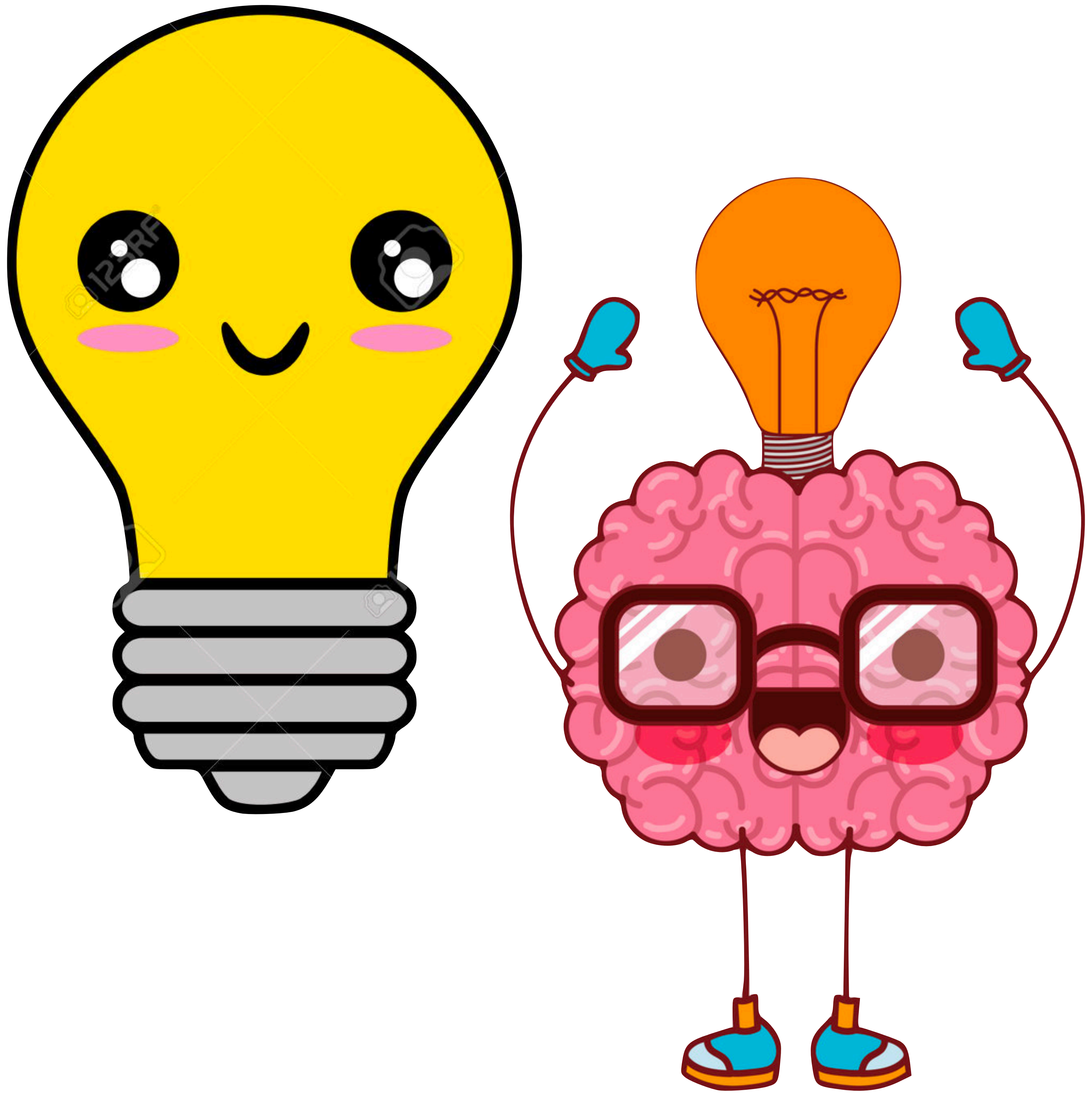


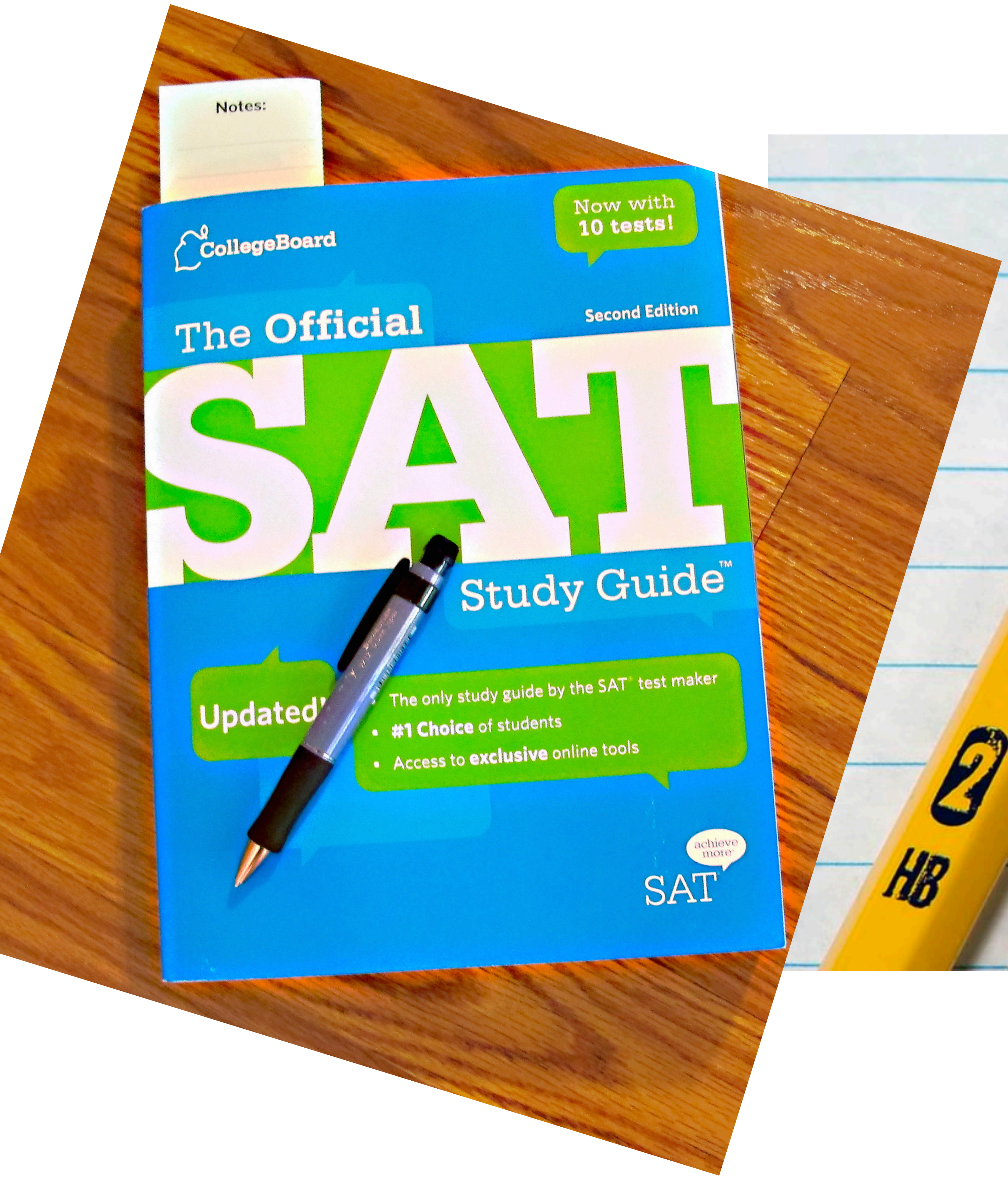
Introduction to Testing

March 15th, 2019

What is testing?

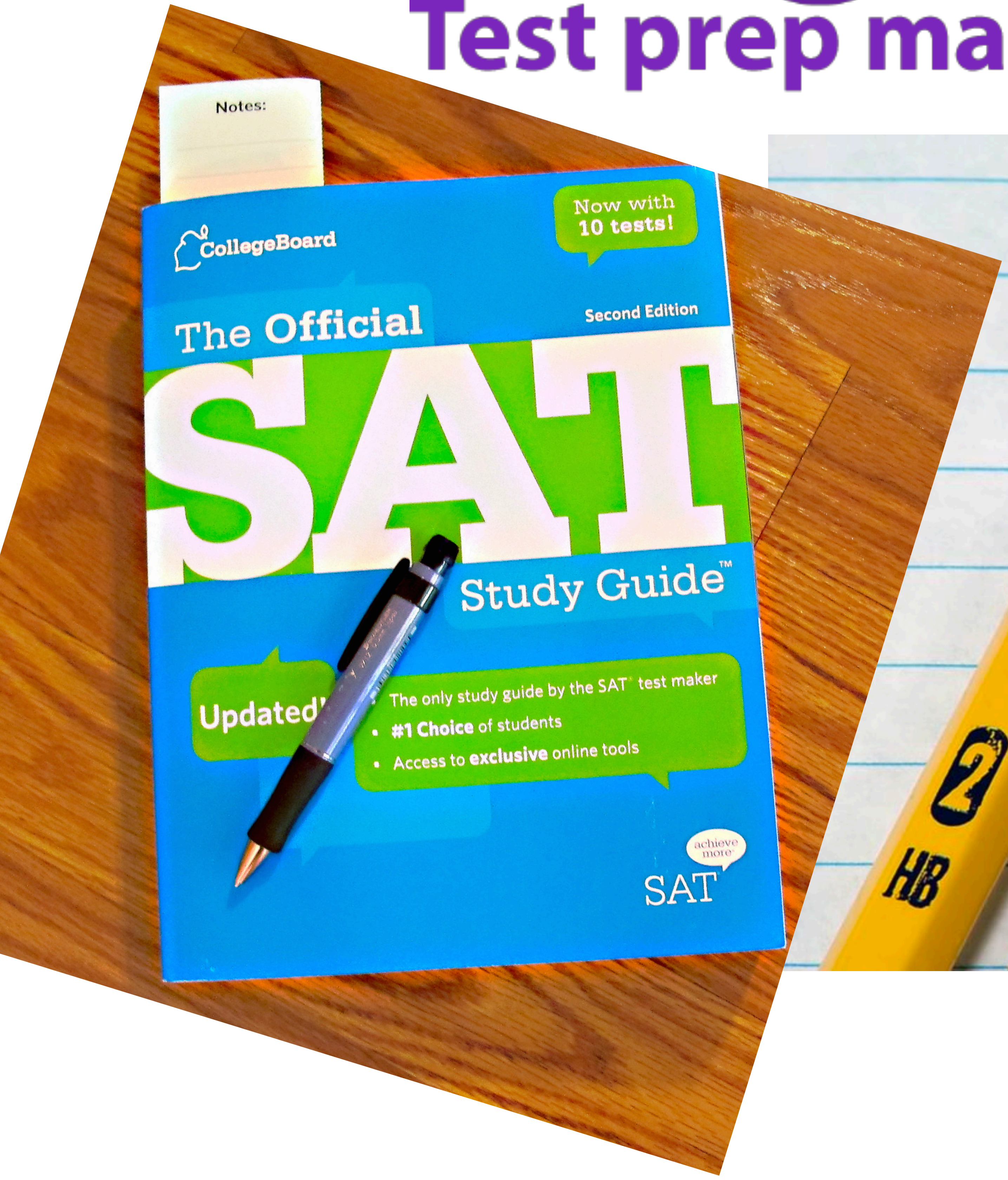
What is testing?

What is testing?



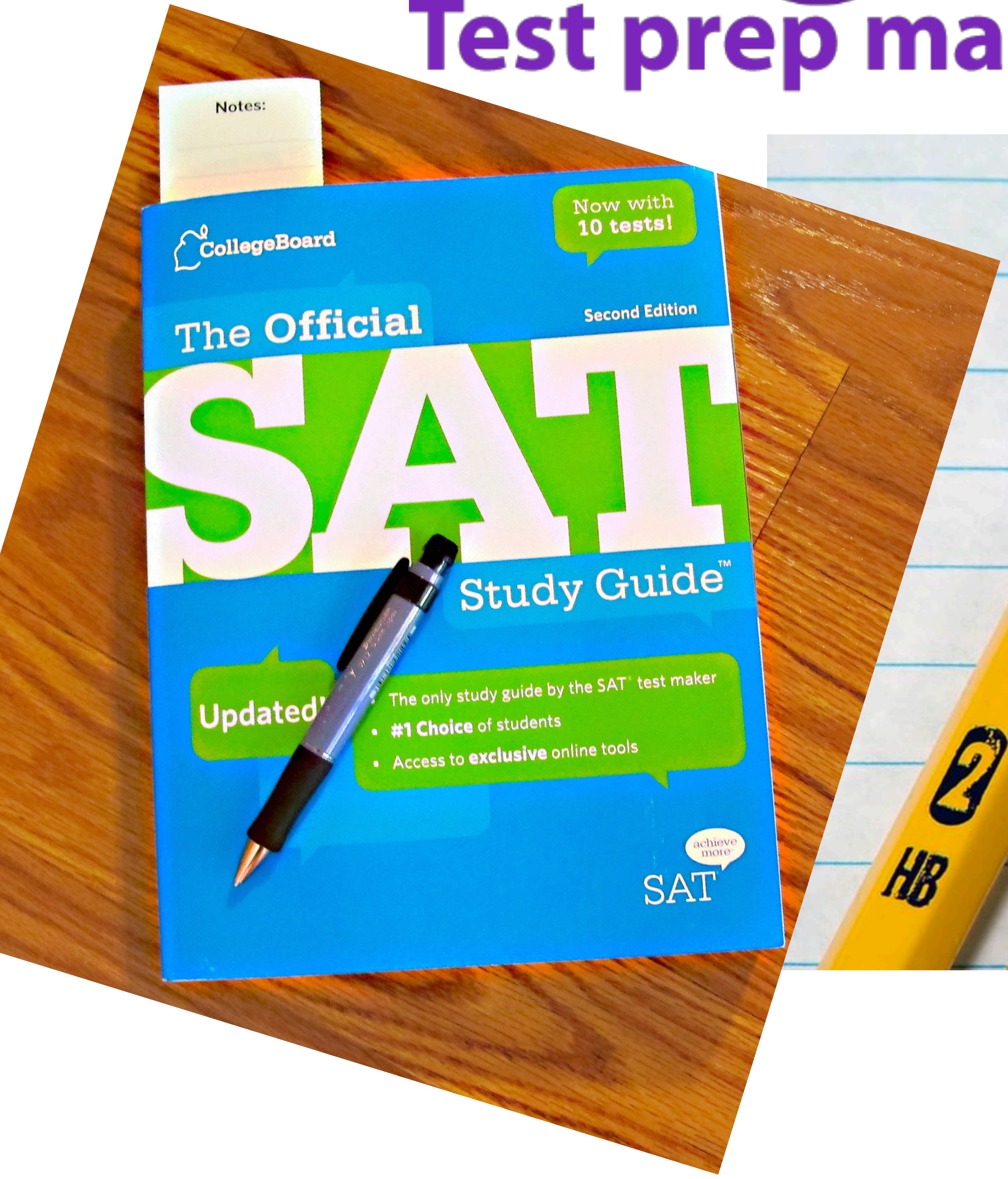
What is testing?

Magoosh
Test prep made simple



What is testing?

Magoosh
Test prep made simple



What is testing?

What is testing?

What is testing?

Hypotheses

Test statistic

or

Confidence Interval

What is testing?

Hypotheses

Test statistic

or

Confidence Interval

Interpretation

and

Conclusion

Hypotheses

- **Null hypothesis**
  - The status quo
- **Alternative hypothesis**
  - A competing opinion
- **Notation**
  - $H_0 : \mu = 10$

    $H_1 : \mu \neq 10$

Hypotheses

- **One sided**
  - $$H_0 : \mu \geq 9$$
    $$H_1 : \mu < 9$$
- **Two sided**
  - $$H_0 : \mu = 10$$
    $$H_1 : \mu \neq 10$$

Hypotheses

DISCLAIMER FOR THIS ENTIRE SECTION AND BEYOND

We are not proving anything in hypothesis tests. We are only saying whether or not we reject the null hypothesis based on our data.

Hypotheses

WE CAN ONLY SUGGEST THAT THE NULL HYPOTHESIS BE REJECTED

because our data can tell another story… but as all samples work, not all samples tell the truth. We just understand probabilities associated with finding significant results.

Confidence Interval

- **Three parts**
  - Sample estimate (either sample mean or proportion)
  - Critical value $z_{\alpha/2}$ **found using qnorm()**
  - Standard error (see associated sampling distribution)
  - Format

$$\bar{x} \pm \textbf{(Margin of Error)} = \bar{x} \pm 1.96\textbf{(SE)}$$

For 95% confidence intervals

Test statistic

- **z tests** The most basic of the tests we are covering in this class!
  - 1. Check conditions
    - You have a SRS
    - The underlying population distribution is normal
    - You know the true population standard deviation
  - 2. Make a test statistic
  - 3. Calculate the p-value
  - 4. Interpret p-value
  - 5. Conclude (Will you reject the null?)

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

p-value

- **p-value**
  - The probability of rejecting the null hypothesis given that the null hypothesis is true
  - The probability of observing our data or more extreme given that the null is true
  - We'll visualize this on the normal distribution for z-tests
  - In general, smaller p-values will imply that we have more evidence against the null hypothesis

Relationship

There is a very specific relationship between confidence intervals and z-tests. For the same data and the same hypotheses, the conclusions of the analyses will be the same.

A 95% confidence interval corresponds to a z test with $\alpha = 0.05$.

Interpretation

- **Confidence Intervals**
  - We are 95% confidence that our true parameter lies within the interval.
  - [Report interval.] This interval was made using a method that creates confidence intervals that contain the true parameter.
- **Test statistic**
  - Our p-value was [this value]. That is, there is a [this value * 100]% chance of observing the data we did or more extreme under the null hypothesis.

Conclusion

- **Confidence Intervals**
    - If our null hypothesized parameter is not within our confidence interval, then we reject the null.
- **Test statistic**
    - If p-value is less than significance level (or very small), then we reject the null hypothesis

Recap… your test will require these things.

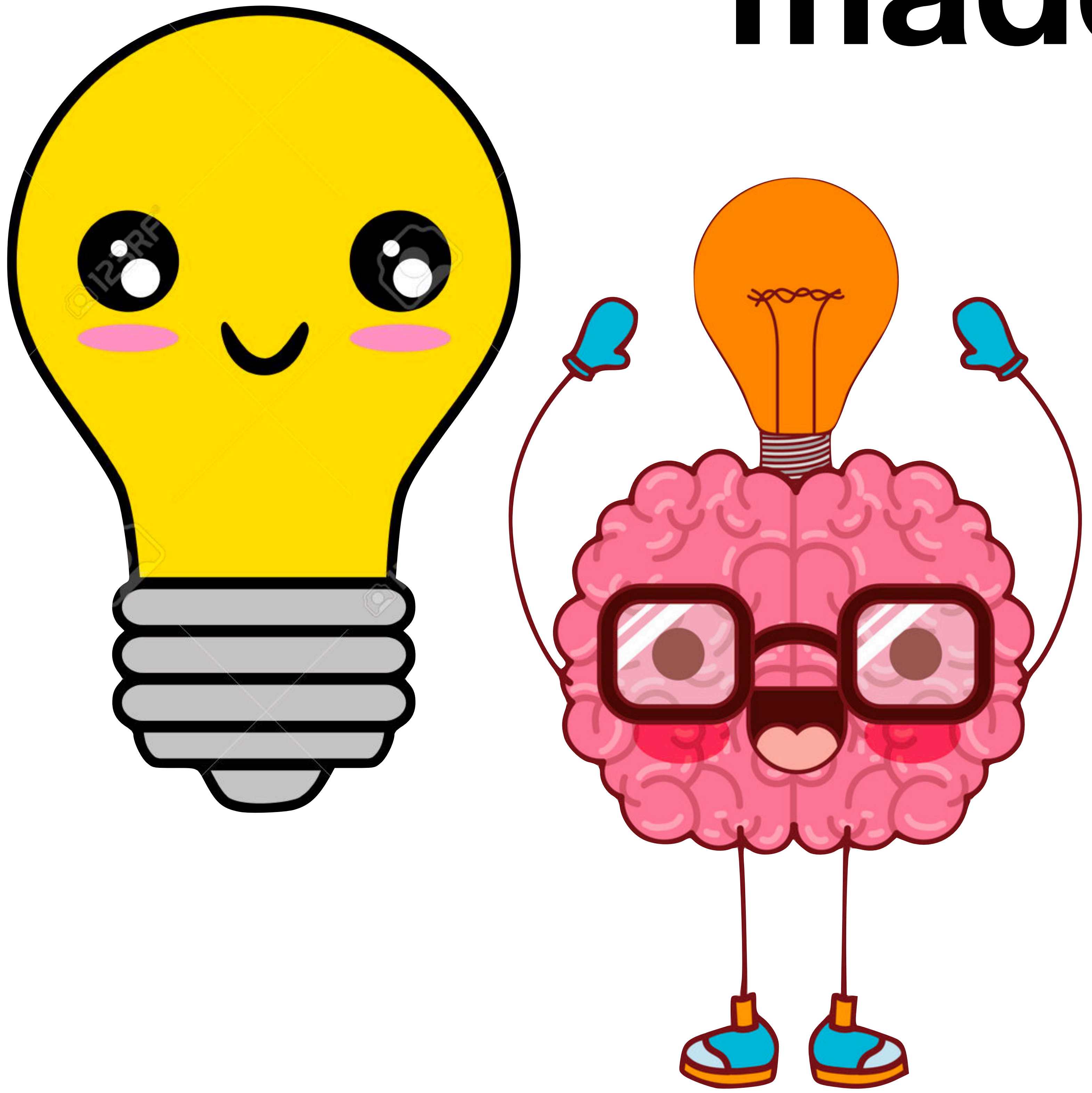
Hypotheses

Test statistic
or
Confidence Interval

Interpretation
and
Conclusion

Why is this true?

Your hypotheses must be made before seeing the data.



If you don’t, you’re treading into a bad place.

If you don’t, you’re treading into a bad place.

If you don’t, you’re treading into a bad place.

Meet this guy.

Meet this guy.
He’s a mad scientist.

Meet this guy.
He’s a mad scientist.
He will do anything to prove a point.

Meet this guy.
He's a mad scientist.
He will do anything to prove a point.

Joobika, Jumba
300 lbs
Current location: Hawaii

The golden rule in research is that if you have a p-value of less than 0.05, then you found a significant discovery.

By chance alone,
we may be able to get a small p-value based on your sample.

People who abuse the above are called “p-hackers”.

The golden rule in research is that if you have a p-value of less than 0.05, then you found a significant discovery.

By chance alone,
we may be able to get a small p-value based on your sample. If you just keep running your experiment a million times, then at least one of your tests can be significant.

And by dishonesty, you can fudge your data.

People who abuse the above are called "p-hackers".

Those who abuse the science of p-values are called "p-hackers".

But if you are a RESEARCHER with a DOCTORATE, why would you do this?!

BLIND PASSION

But if you are a RESEARCHER with a DOCTORATE, why would you do this?!

BLIND PASSION

But if you are a **RESEARCHER** with a **DOCTORATE**, why would you do this?!

IGNORANCE

BLIND PASSION

But if you are a RESEARCHER with a DOCTORATE, why would you do this?!

IGNORANCE

TENURE

BLIND PASSION

But if you are a RESEARCHER with a DOCTORATE, why would you do this?!

IGNORANCE

TENURE

FAME

BANK
$$$

BLIND
PASSION

But if you are a RESEARCHER with a DOCTORATE, why would you do this?!

IGNORANCE

TENURE

FAME

BANK
$$$

BLIND
PASSION

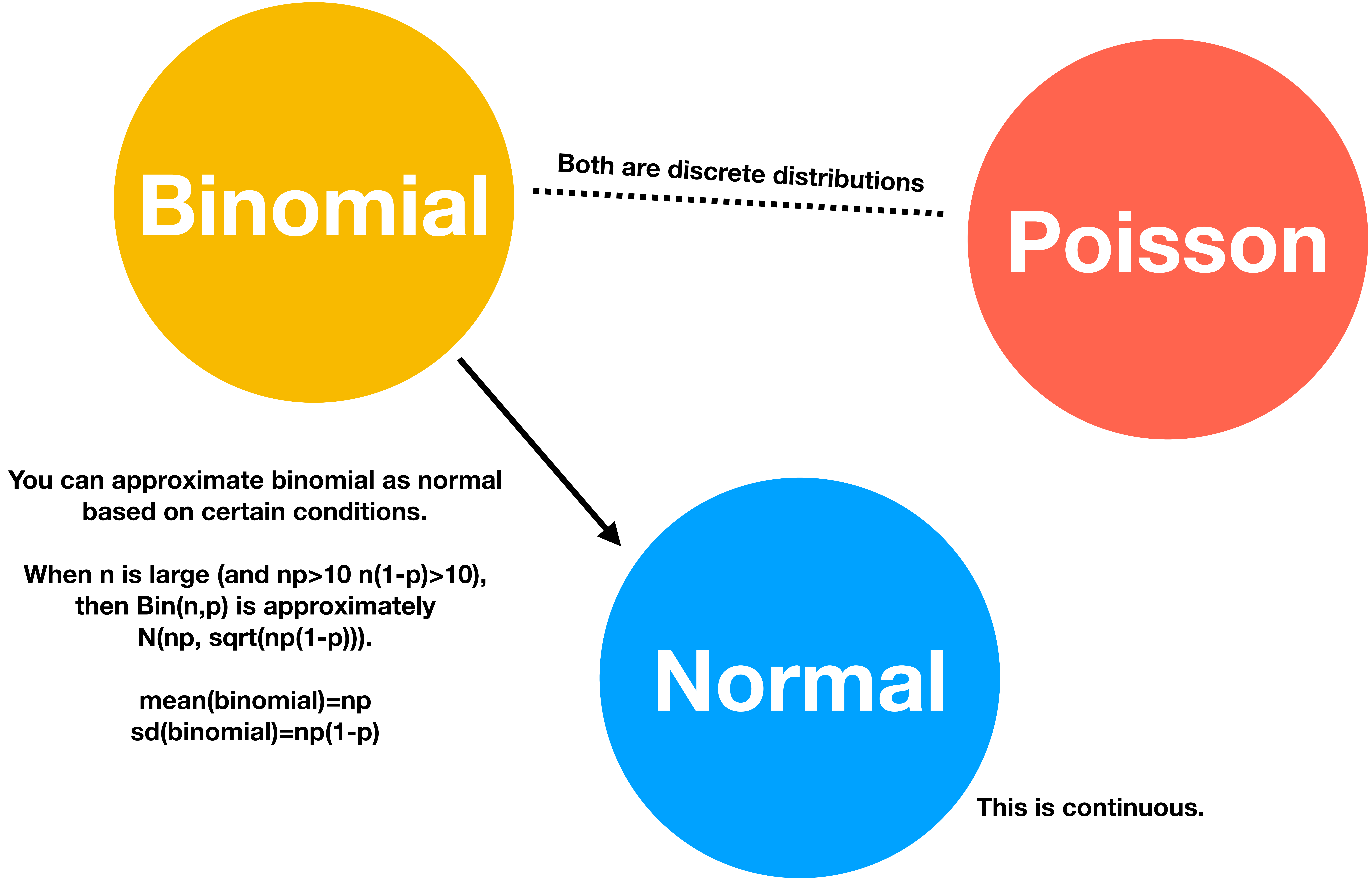
Top Cornell food researcher [Brian Wansink] did it.

CITATIONS

TENURE

FAME

Review


Binomial
Both are discrete distributions
Poisson
You can approximate binomial as normal based on certain conditions.
When n is large (and np>10 n(1-p)>10), then Bin(n,p) is approximately N(np, sqrt(np(1-p))).
mean(binomial)=np
sd(binomial)=np(1-p)
Normal
This is continuous.