"The fault lies not in our ✳✳✳ but in ourselves."

- Cassius to Brutus in Shakespeare's *Julius Caesar*

Danny Kaplan, Macalester College
USCOTS 2019

"[T]he tool has become the tyrant."

— Wasserstein, Schirm & Lazar (2019)
"Moving to a World Beyond 'p < 0.05'"
The American Statistician **73**

"The ASA Statement on P-Values and Statistical Significance stopped just short of recommending that declarations of 'statistical significance' be abandoned. We take that step here. We conclude, based on our review of the articles in this special issue and the broader literature, that **it is time to stop using the term 'statistically significant' entirely**. Nor should variants such as 'significantly different,' 'p < 0.05,' and 'nonsignificant' survive, whether expressed in words, by asterisks in a table, or in some other way.

"**In sum, "statistically significant" — don't say it and don't use it**."

Nature "Retire Statistical Significance"

March 2019

"How do statistics so often lead scientists to deny differences that those not educated in statistics can plainly see? For several generations, researchers have been warned that a statistically non-significant result does not 'prove' the null hypothesis (the hypothesis that there is no difference between groups or no effect of a treatment on some measured outcome). Nor do statistically significant results 'prove' some other hypothesis. Such misconceptions have famously warped the the literature with overstated claims and, less famously, led to claims of conflicts between studies where none exists."

Losing confidence in confidence

- Big n means small p and narrow CI

$$F \equiv \left(\frac{n - p}{p - 1} \right) \frac{R^2}{1 - R^2}$$

- Machine learning empowers pattern hacking.

- Many researchers, many studies, many ways through "The Garden of the Forked Paths"!

Ways to fix things?

- .p < 0.01 / no threshold / use log(p)

- ~~Significance~~

- Use effect size, "practical benefit"

- G values (Ziliak's "Guinnessometrics")

- *** ⟶ ★★★★★

-

Proposal:
Tie the scale to a responsible research process and teach that process

★★★★★ CI above SEWM*

★★★★★ Covariates approp. incorporated, DAG

★★★★★ Effect size has "practical benefit"

★★★★★ Experiment or Confounding Interval

★★★★★ Finding independently confirmed

* Smallest Effect Worth Mentioning

"**So let's do it.**"

"Statistics education will require major changes at all levels to move to a post 'p < 0.05' world. We are excited that, with support from the ASA, the **US Conference on Teaching Statistics (USCOTS)** will focus its 2019 meeting on teaching inference."

"The change that needs to happen demands change to editorial practice, to the teaching of statistics at every level where inference is taught, and to much more. **It Is Going to Take Work, and It Is Going to Take Time**.

— Wasserstein, Schirm & Lazar (2019)
"Moving to a World Beyond 'p < 0.05'"
The American Statistician **73**

2019
USCOTS

EVALUATING
EVIDENCE