
REPRODUCIBLE RESEARCH @UF

for UF Research & Scholarship Council

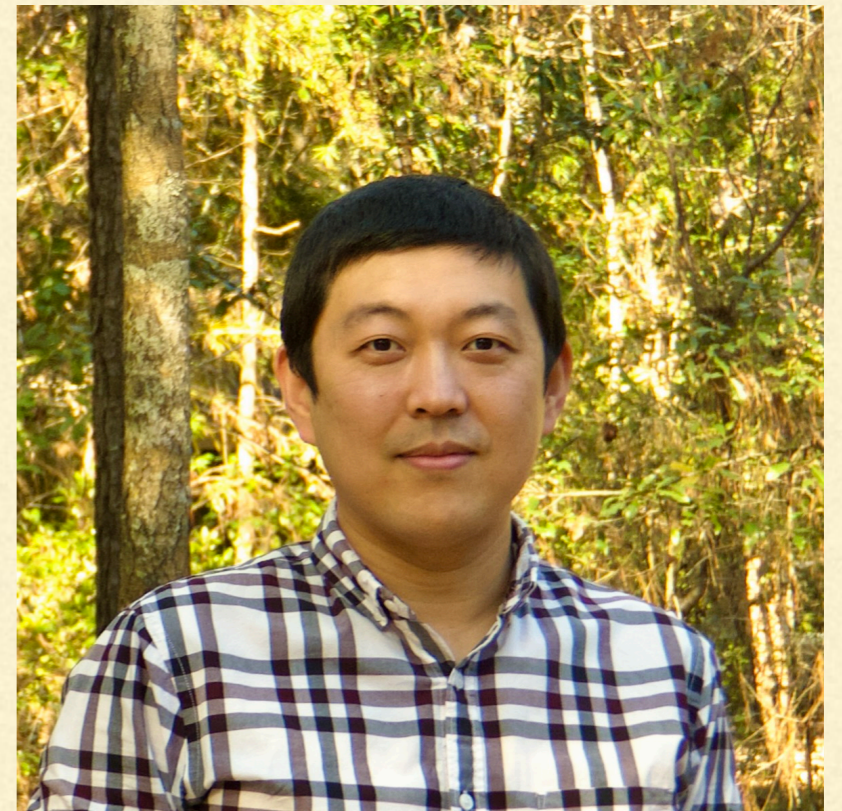
2021-04-02

Short Bio - Hao Ye

- ◆ 2015 PhD, Scripps Institution of Oceanography
- ◆ 2017 Postdoc & Moore Data Fellow, University of Florida
- ◆ 2020 Reproducibility Librarian

I believe open practices are a pathway to more rigorous, transparent, impactful, and equitable research and systems.

I promote research reproducibility and open science by educating and empowering researchers.



What does reproducibility mean anyway?

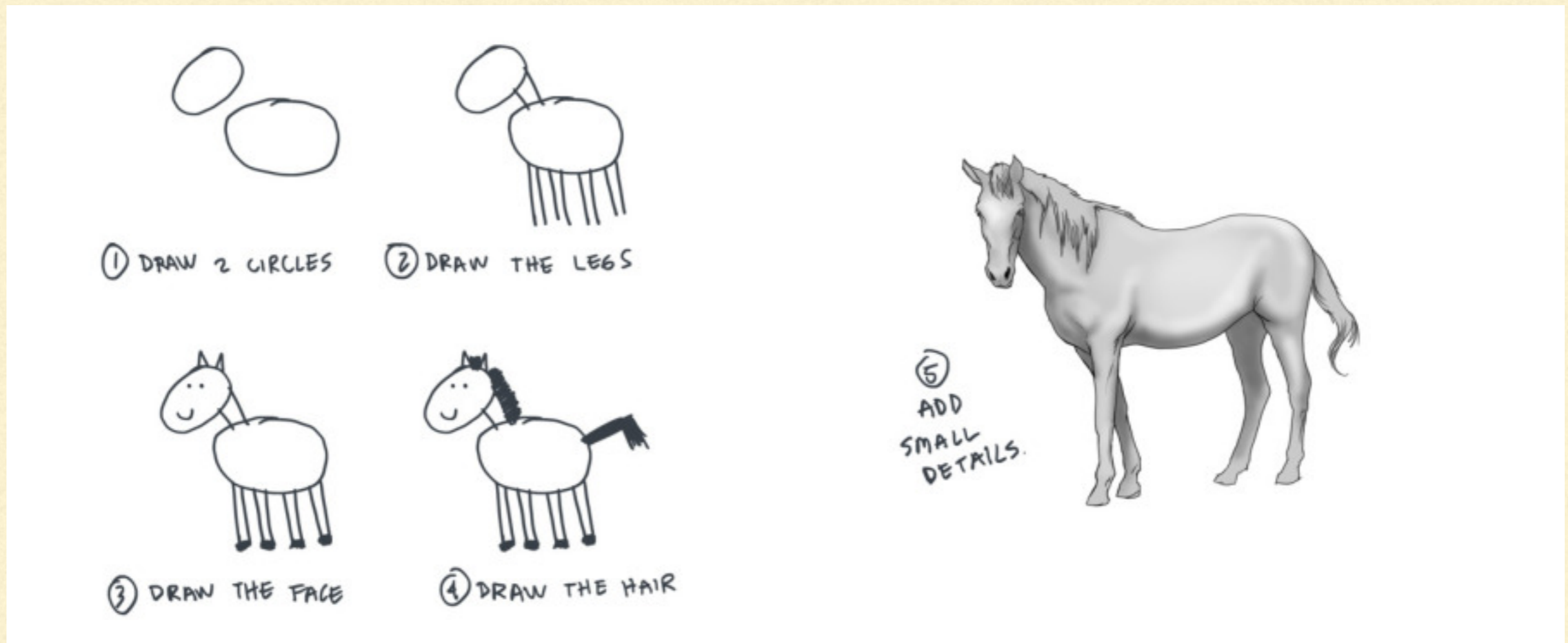
"green jelly beans cause acne ($p < 0.05$)"

- ◆ Generalizability—the effect is observed across different populations under different settings
- ◆ Replicability—the effect is observed when repeating the experiment in a different sample (but identical population)
- ◆ Reproducibility—the statistics and figures in the original paper can be reproduced using the same data, statistical software, analysis code, etc.

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

© The Turing Way, shared under CC-BY 4.0

Challenges in reproducing methods



Oktop, Van. "HOW TO: DRAW A HORSE" Van Oktop. 2012-01-05.
Web. 2020-01-27 <https://oktop.tumblr.com/post/15352780846>

The "Replication Crisis"

- ◆ many studies do not replicate and/or are not reproducible
 - distributions of p-values in publications show discontinuities at the 0.05 threshold
 - multi-lab replication efforts (~10–50% of results replicate)
 - "Many Labs" studies
 - efforts to reproduce from openly available code/data show wide-ranging success (~ 0%–80%)
 - multi-lab analysis efforts show variability in results from common data due to variation in statistical methodology
 - "Many Analysts" studies

Causes

◆ Systemic Factors

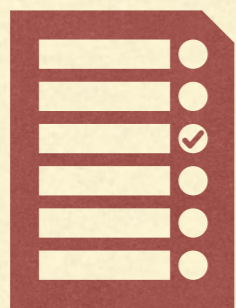
- publication bias (for novel, statistically significant results)
- funding, promotion, hiring incentives (for publications in prominent journals)

◆ Individualized Factors

- issues in experimental design and/or methodology
- incomplete or inaccurate reporting
- social and technological barriers for sharing data, code, etc.

Solutions

- ◆ mandates from funders and publishers
 - NIH guidelines for Rigor and Reproducibility
- ◆ transparency practices
 - pre-registration and registered reports (public declarations of data and analysis protocols)
 - open access, data, and code (+ publications and credit)
- ◆ tools and platforms
 - preprint servers, reproducible computing environments,



Selected Activities (to date)

- ◆ 15 virtual workshops (2020-05-01 to 2021-04-01)
 - data organization & cleaning, programming practices, version control, sharing & publishing work
<https://guides.uflib.ufl.edu/reproducibility/lessons>
- ◆ 3 guest lectures (2 upcoming in April)
 - rigor and reproducibility, data cleaning, version control
- ◆ 2020 Research Reproducibility conference
- ◆ (submitted) planning grant for open repository of grant proposals

UF

Rigor and Reproducibility Seminar Series

UF Interdisciplinary T32 in Movement Disorders and Neurorestoration

SCHEDULE

SPEAKERS

QUESTIONS OR CONCERNS?



<https://uf-repro.github.io/movementdisorders-seminar/>

Next Steps

- ◆ coordinate ReproducibiliTea Journal Club
- ◆ organize Reproducibility Hackathon
- ◆ foster local communities of practice
- ◆ map out UF digital and social resources in related and intertwined topics
 - open access, research integrity, data science, AI
- ◆ develop training and mentoring for lab/departmental culture re: open research practices
 - project management, publication and authorship guidelines, communications, skill sharing, etc.