How to do Research for Fun and Profit

Divesh Srivastava
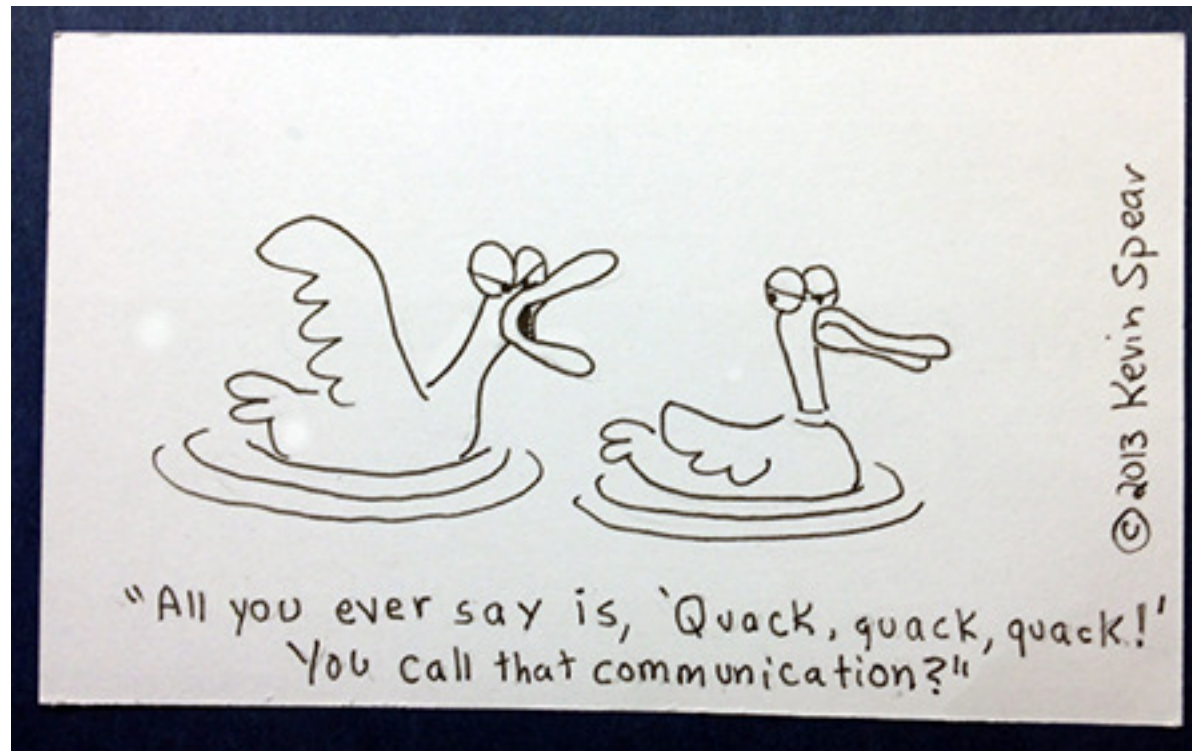
AT&T Labs-Research

Research for Fun and Profit

♦ Research is about posing new questions …

Research for Fun and Profit

◆ Research is about posing new questions, a creative inquiry to find answers …

Research for Fun and Profit

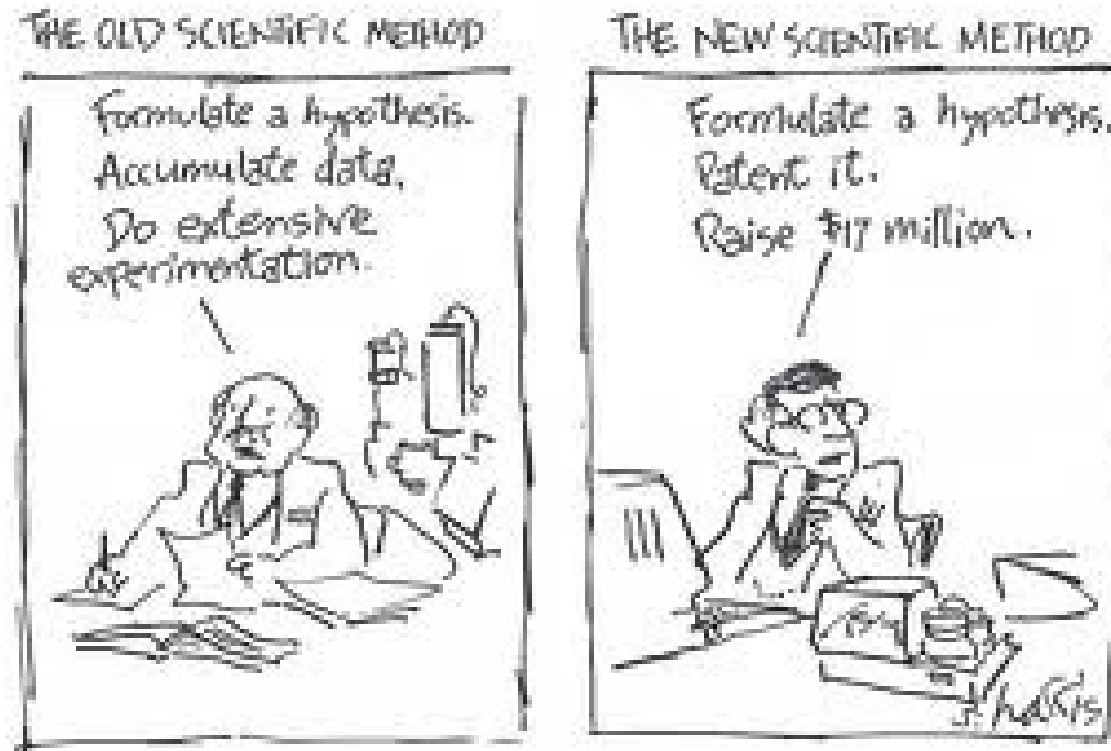"All you ever say is, 'Quack, quack, quack!' You call that communication?"

♦ Research is about posing new questions, a creative inquiry to find answers, communicating this knowledge to the community …

Research for Fun and Profit

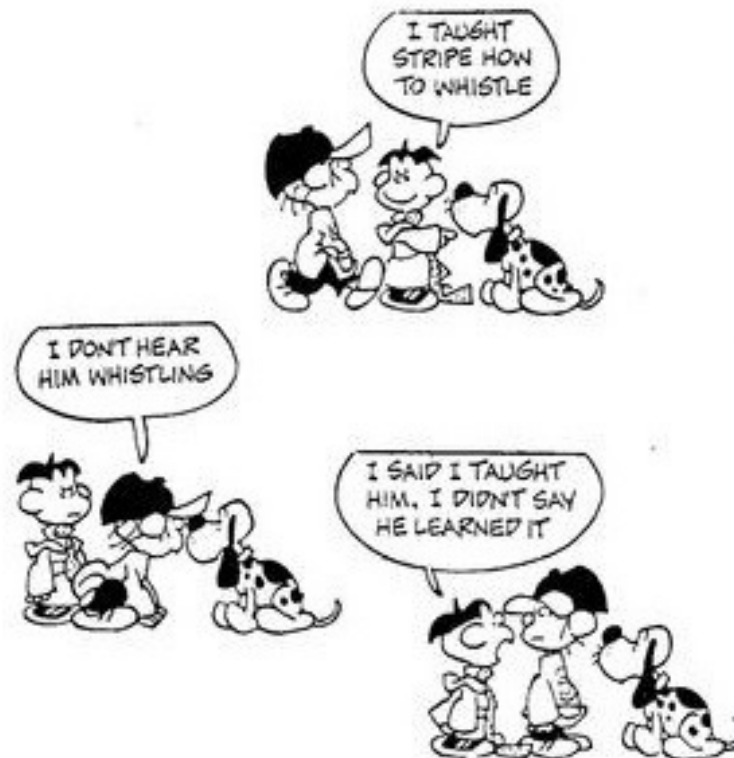♦ Research should also be fun, and …

Research for Fun and Profit

THE OLD SCIENTIFIC METHOD

Formulate a hypothesis.
Accumulate data.
Do extensive experimentation.

THE NEW SCIENTIFIC METHOD

Formulate a hypothesis.
Patent it.
Raise $17 million.
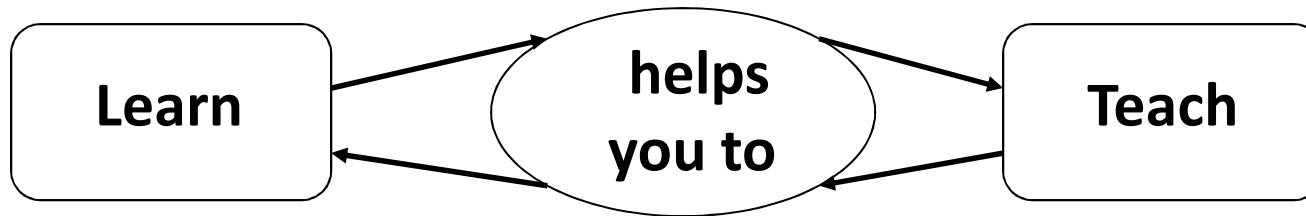
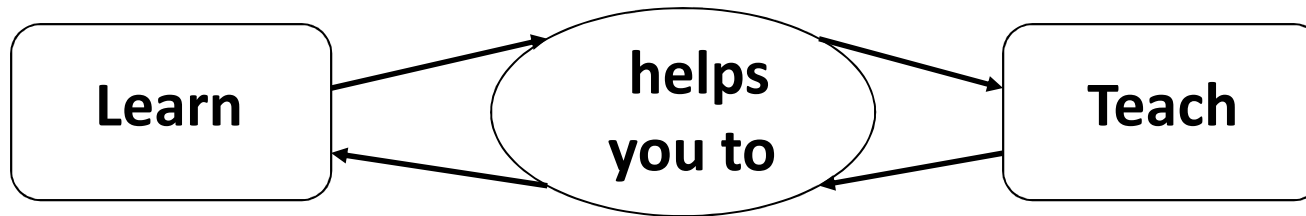◆ Research should also be profitable*

   – * Your results may vary

Outline of Talk

♦ Preparing for research

♦ Choosing research problems

♦ Doing the actual research

Preparing for a Research Career

Learn → *helps you to* → **Teach**

Preparing for a Research Career

```
┌──────────┐        ╭─────────╮        ┌──────────┐
│          │ ──────▶│  helps  │──────▶ │          │
│  Learn   │        │ you to  │        │  Teach   │
│          │ ◀──────│         │◀────── │          │
└──────────┘        ╰─────────╯        └──────────┘
```

◆ **Read** voraciously

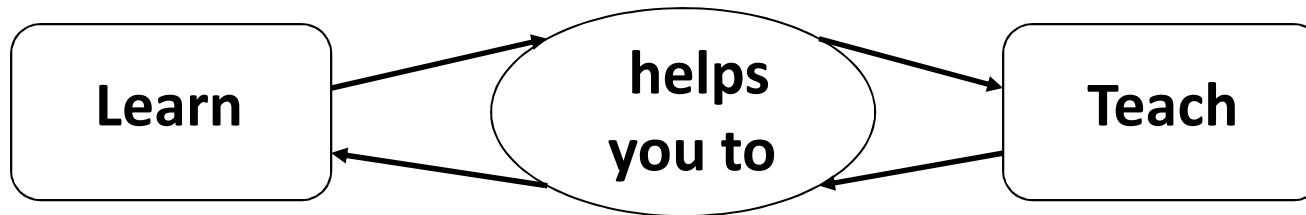◆ **Listen** attentively

◆ **Write** extensively

◆ **Present** effectively

On Reading

"It's called 'reading'. It's how people install new software into their brains."

Preparing for a Research Career

◆ **Read** voraciously

- – Of course, everything your advisor asks you to read, but also …
- – Well-written, highly-cited papers
- – The Elements of Style, by E. B. White and William Strunk, Jr, 1920
- – Draft papers by student colleagues

On Writing

Preparing for a Research Career

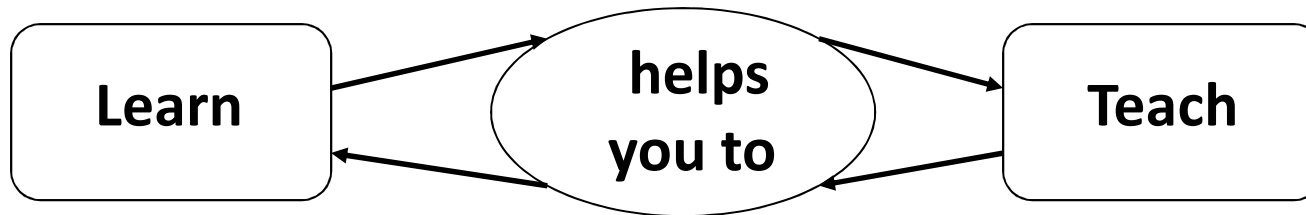Learn → **helps you to** → Teach

- ◆ **Write** extensively
  - – Even when your advisor does not ask you to write, such as …
  - – Summaries, reviews of papers that you have read
  - – Feedback on draft papers by your student colleagues
  - – Write clearly, eschew obfuscation ☺

On Listening

PEANUTS CLASSICS By Charles M. Schulz

Preparing for a Research Career

Learn ⟷ helps you to ⟷ Teach

♦ **Listen** attentively

 – Of course to your advisor, but also …

 – To Turing award lectures **http://amturing.acm.org/lectures.cfm**

 – To TED talks **https://www.ted.com/talks/browse**

 – To presentations by your student colleagues

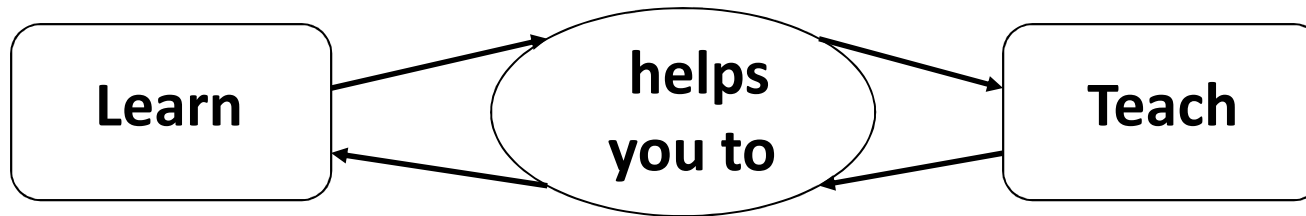# On Presenting

"OK, I'm now going to read out loud every single slide to you, word for word, until you all wish you'd just die."

Preparing for a Research Career
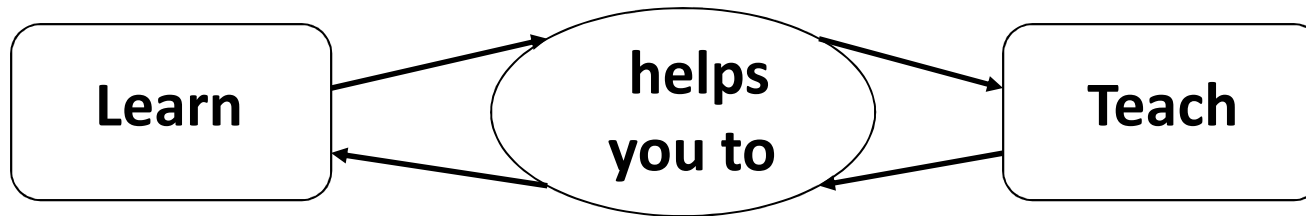


```
┌─────────┐          ╱‾‾‾‾‾‾‾╲          ┌─────────┐
│  Learn  │  ⇄      (  helps  )      ⇄  │  Teach  │
└─────────┘          ( you to )          └─────────┘
                      ╲_____╱
```

◆ **Present** effectively

    – Even when your advisor does not ask you to present, such as …

    – Papers you have read, using your own slides, to student colleagues

Preparing for a Research Career: Recap

Learn → helps you to → Teach

- ◆ **Read** voraciously

- ◆ **Listen** attentively

- ◆ **Write** extensively

- ◆ **Present** effectively

Outline of Talk

♦ Preparing for research

♦ **Choosing research problems**

♦ Doing the actual research

Goal: High Quality, High Impact Research

♦ Prediction is very difficult, especially if it's about the future

  – Niels Bohr, sometimes attributed to Yogi Berra

Choosing Problems: Goldilocks Principle

♦ Goldilocks and the three bears
  – Robert Southey (1837)

Choosing Problems: Goldilocks Principle

♦ Porridge: not too hot, not too cold, just right!

Choosing Problems: Goldilocks Principle

♦ Porridge: metaphor for **taste** in research problems

♦ Research problem should not be **too hot**

  – Porridge too hot → you can burn your mouth

  – Problem "too hot" → ???

♦ Advantage: many people seem to be excited about the problem

♦ Disadvantage: may just be the fad of the moment, no longevity

  – Difficult to do **high quality** research

Choosing Problems: Goldilocks Principle

♦ Porridge: metaphor for **taste** in research problems

♦ Research problem should not be **too cold**

  – Porridge too cold → it is not tasty

  – Problem "too cold" → ???

♦ Advantage: can spend the time to do technically deep work

♦ Disadvantage: very few people are excited about the problem

  – Difficult to do **high impact** research

Choosing Problems: Goldilocks Principle

◆ Porridge: metaphor for **taste** in research problems

◆ Research problem

– Should not be **too hot**, should not be **too cold**

– Should be "just right": potential for **high quality** and **high impact**

◆ Example: Data stream research in AT&T [2001—now]

– GS Tool system, algorithms and systems research, used in AT&T

– Open-sourced as Tigon SQL **http://tigon.io/**

Choosing Problems: Goldilocks Principle

♦ Bed: not too hard, not too soft, just right!

Choosing Problems: Goldilocks Principle

◆ Bed: metaphor for **difficulty** of research problems

◆ Research problem should not be **too hard**

 – Bed too hard → you toss and turn, cannot sleep

 – Problem "too hard" → ???

◆ Advantage: if you succeed, you get a lot of recognition

◆ Disadvantage: unlikely to succeed in a reasonable amount of time

 – Difficult to do **high impact** research

Choosing Problems: Goldilocks Principle

♦ Bed: metaphor for **difficulty** of research problems

♦ Research problem should not be **too soft**

   – Bed too soft → you sleep, but wake up feeling tired

   – Problem "too soft" → ???

♦ Advantage: easy to completely solve the problem

♦ Disadvantage: difficult to get recognition and acceptance

   – Difficult to do **high quality** research

Choosing Problems: Goldilocks Principle

♦ Bed: metaphor for **difficulty** of research problems

♦ Research problem

    – Should not be **too hard**, should not be **too soft**

    – Should be "just right": potential for **high quality** and **high impact**

♦ Example: Data quality research in AT&T [2003—now]

   – Many tools, algorithms and statistics research, used in AT&T

Choosing Problems: Recap

♦ Goldilocks principle

♦ **Taste** in research problems
  – Should not be **too hot**, should not be **too cold**
  – Should be "just right": potential for high quality and high impact

♦ **Difficulty** of research problems
  – Should not be **too hard**, should not be **too soft**
  – Should be "just right": potential for high quality and high impact

Outline of Talk

♦ Preparing for research

♦ Choosing research problems

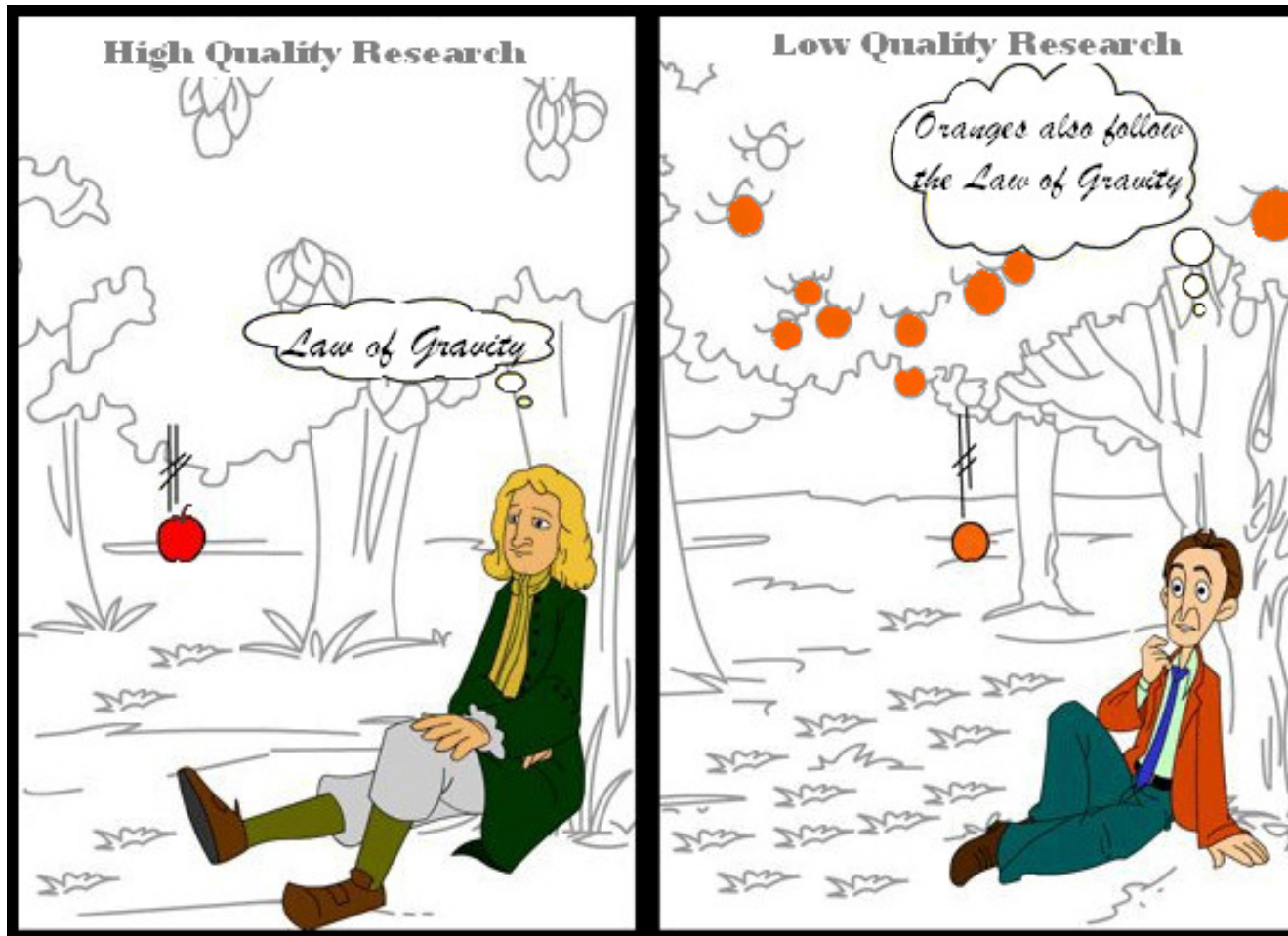♦ **Doing the actual research**

Doing the Actual Research
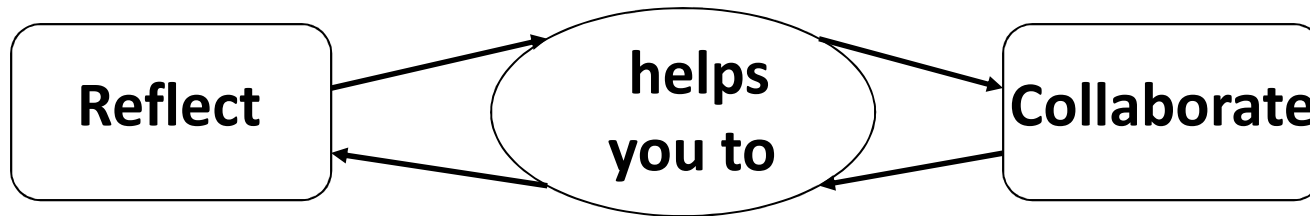
◆ Goal 1: Doing **high quality** research

◆ Goal 2: Communicating research results to have **high impact**

    – Papers, presentations, systems, …

Research: High Quality

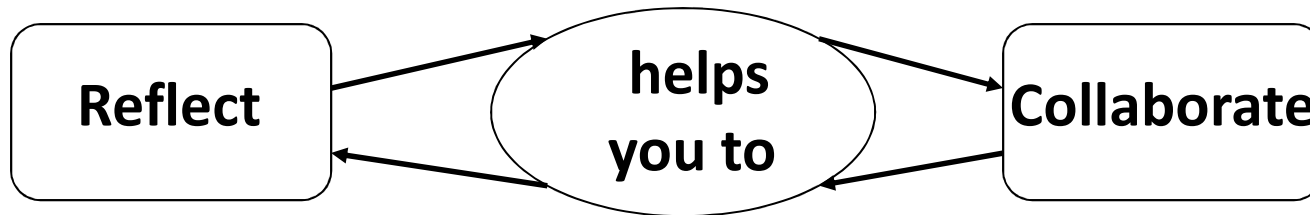# Research: High Quality vs Low Quality

Need for Communicating Research Results

- ◆ If a tree falls in a forest and no one is around to hear it, does it make a sound?   -- George Berkeley (1710)
  - – Communication makes **high quality** research have **high impact**
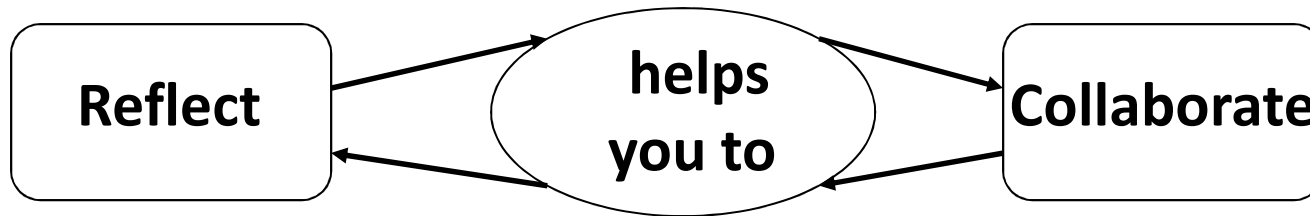
Doing the Actual Research

```
┌──────────┐        ╭─────────╮        ┌──────────────┐
│          │ ─────► │  helps  │ ─────► │              │
│ Reflect  │        │ you to  │        │ Collaborate  │
│          │ ◄───── │         │ ◄───── │              │
└──────────┘        ╰─────────╯        └──────────────┘
```

♦ **Analyze** systematically

♦ **Read** voraciously

♦ **Listen** attentively

♦ **Criticize** constructively

♦ **Write** extensively

♦ **Present** effectively

On Analyzing

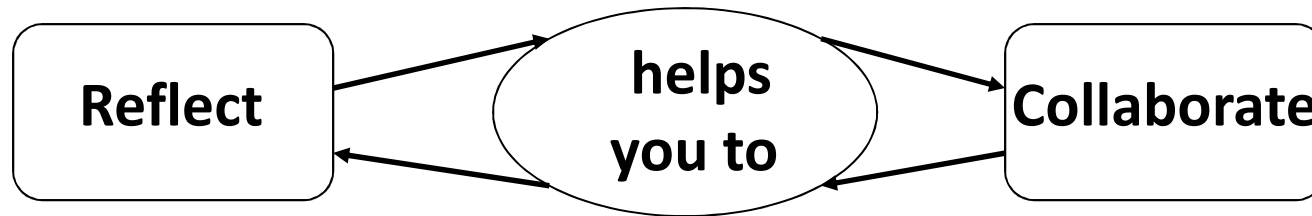# Doing the Actual Research
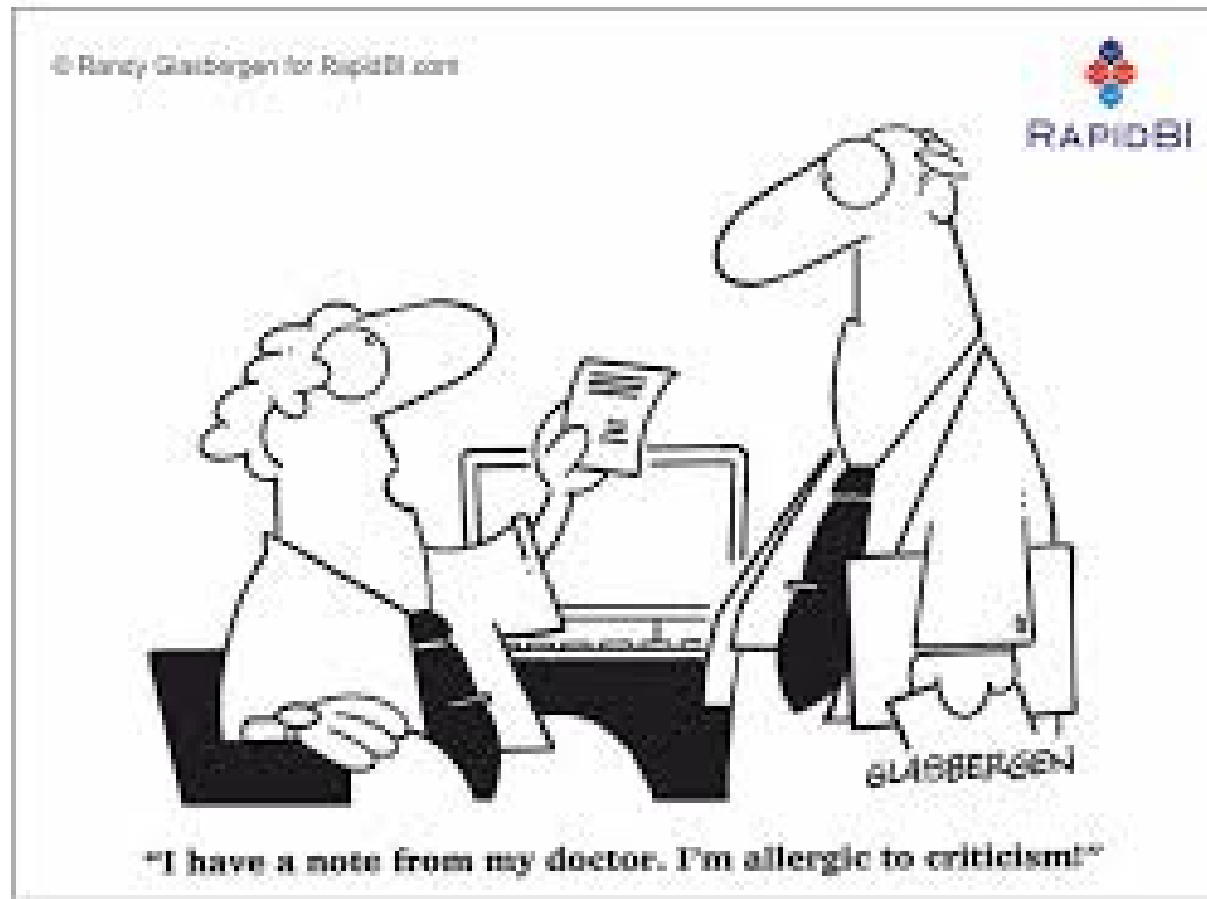

Reflect → helps you to → Collaborate

♦ **Analyze** systematically

   – Of course, correctness, complexity of algorithms, but also …

   – Think twice about how to solve a problem

   – Start simple, layer-in complexity

Doing the Actual Research

```
┌──────────┐         ╭─────────╮         ┌──────────────┐
│          │ ──────→ │  helps  │ ──────→ │              │
│ Reflect  │         │ you to  │         │ Collaborate  │
│          │ ←────── │         │ ←────── │              │
└──────────┘         ╰─────────╯         └──────────────┘
```
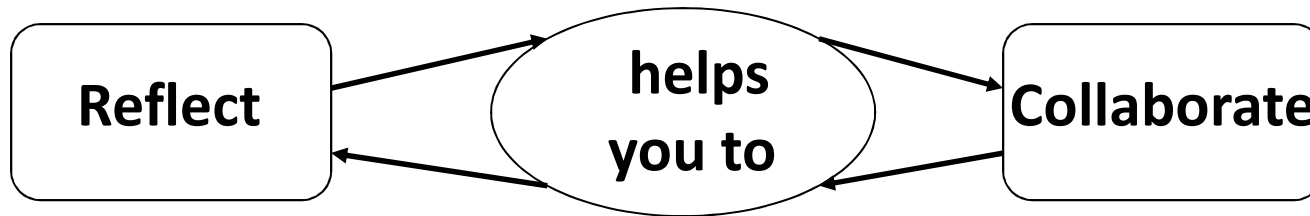
◆ **Read** voraciously

   – Everything in your area of research

◆ **Listen** attentively

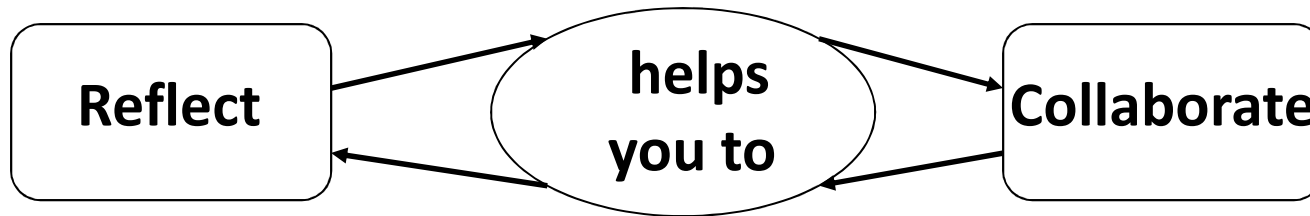   – To criticisms from collaborators, reviewers

On Criticizing

"I have a note from my doctor. I'm allergic to criticism!"

Doing the Actual Research


```
  ┌──────────┐      ╭─────────╮      ┌─────────────┐
  │          │ ───▶ │  helps  │ ───▶ │             │
  │ Reflect  │      │ you to  │      │ Collaborate │
  │          │ ◀─── │         │ ◀─── │             │
  └──────────┘      ╰─────────╯      └─────────────┘
```
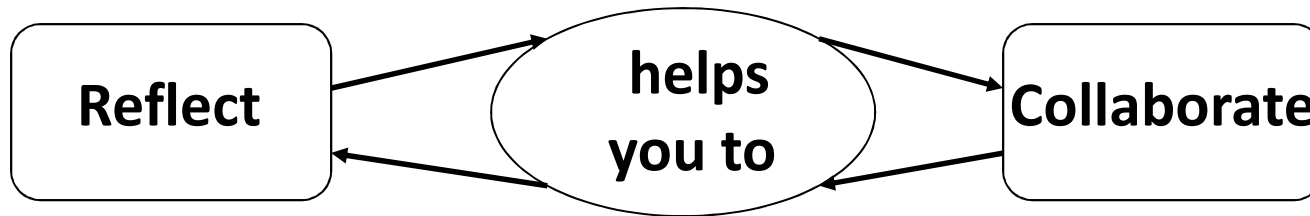
♦ **Criticize** constructively

 – When you collaborate, don't just agree with collaborators, but …

 – If result does not seem correct, come up with counter-examples

 – If algorithm looks too complex, design a simpler algorithm

 – Provide feedback: what works, what does not

Doing the Actual Research

```
┌──────────┐              ╭─────────╮              ┌─────────────┐
│          │  ──────────▶ │  helps  │ ──────────▶  │             │
│ Reflect  │              │ you to  │              │ Collaborate │
│          │  ◀────────── │         │ ◀──────────  │             │
└──────────┘              ╰─────────╯              └─────────────┘
```

♦ **Write** extensively

  – Survey papers, with your own perspective, seek feedback

♦ **Present** effectively

  – Tutorials on your topic, interact with the audience

Doing the Actual Research: Recap

Reflect → **helps you to** → **Collaborate**

♦ **Analyze** systematically

♦ **Read** voraciously

♦ **Listen** attentively

♦ **Criticize** constructively

♦ **Write** extensively

♦ **Present** effectively

Summary

◆ Preparing for research

- – Learn, teach

◆ Choosing research problems

- – Goldilocks principle

◆ Doing the actual research

- – Reflect, collaborate

Research: The Last Word