An **awesome new test** has been invented for an early detection of cancer. The probability that it correctly identifies someone with cancer as positive is 95%, and the probability that it correctly identifies someone without cancer as negative is 99%. The incidence of this type of cancer in the general population is 10^{-4}. A random person in the population takes the test, and the result is positive.

What is the probability that he/she has cancer?

A. 99%
B. 95%
C. 30%
D. 1%

Get your i-clickers

participants

10^6 < 100 - cancer - 95 positive tests

$10^6 - 100 \approx 10^6$ no cancer

10^6 participants with no cancer → 10,000 positive tests

$$P(C|P) = \frac{95}{10{,}000 + 95} \approx 1\%$$

Events: C - cancer, C' - no cancer

Test events Y - positive, N - negative

We know:

$P(C) = 10^{-4}$, $P(Y|C) = 0.95$

$P(N|C') = 0.99$

We need

$P(C|Y)$

Bayes:

$$P(C|Y) = P(Y|C) \cdot \frac{P(C)}{P(Y)}$$

?

P(Y) – probability that a random person will test positive

$$P(Y) = P(Y \cap C) + P(Y \cap C') =$$

$$= P(Y|C)P(C) + P(Y|C')P(C') =$$

$$= 0.95 \times 10^{-4} + (1 - 0.99) \times (1 - 10^{-4}) \approx$$

$$\approx 10^{-4} + 10^{-2} \approx 10^{-2} = 1\%$$

$$P(C|Y) = P(Y|C) \cdot \frac{P(C)}{P(Y)} = 0.95 \times \frac{10^{-4}}{10^{-2}} \approx$$

$$\approx 1\%$$

An **awesome new test** has been invented for an early detection of cancer. The probability that it correctly identifies someone with cancer as positive is 95%, and the probability that it correctly identifies someone without cancer as negative is 99%. The incidence of this type of cancer in the general population is 10^{-4}. A suspected cancer patient with likelihood of cancer 50% takes the test, and the result is positive.

What is the probability that he/she has cancer?

A. 99%
B. 95%
C. 30%
D. 1%

Get your i-clickers

An **awesome new test** has been invented for an early detection of cancer. The probability that it correctly identifies someone with cancer as positive is 95%, and the probability that it correctly identifies someone without cancer as negative is 99%. The incidence of this type of cancer in the general population is 10^{-4}. A suspected cancer patient with likelihood of cancer 50% takes the test, and the result is positive.

What is the probability that he/she has cancer?

A. 99%
B. 95%
C. 30%
D. 1%

Get your i-clickers

What if a doctor is already 50% sure of cancer based on other tests?

That changes things!

Now $P(C) = P(C') = 0.5$

$$P(C|Y) = \frac{P(Y|C) \cdot P(C)}{P(Y|C) \cdot P(C) + P(Y|C')\, P(C')} =$$

$$= \frac{0.95 \times 0.5}{0.95 \times 0.5 + (1-0.99) \times 0.5} \approx 0.99$$

Sensitivity/specificity of the standard test for prostate cancer: PSA level > 4.0ng/mL

- Sensitivity of the test is 71.9%
  - fraction of cancer patients who will test positive
  - False negative rate is 28.1%
- Specificity of the test is 90%
  - fraction of healthy patients who will test negative
  - False positive rate is 10%

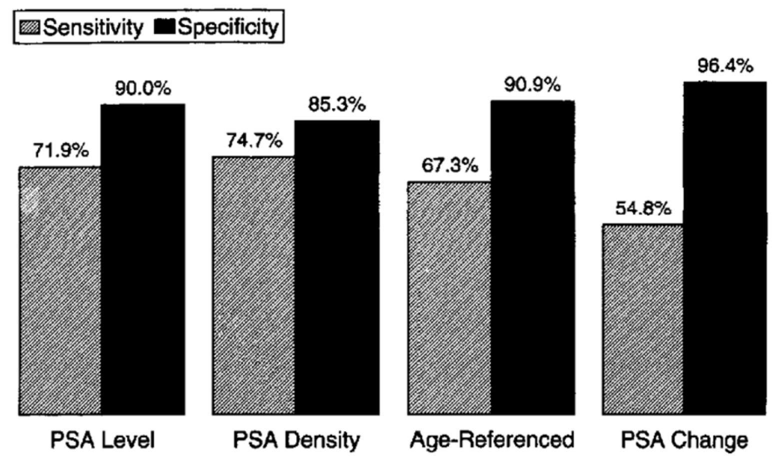


Figure 1. The relative sensitivity and specificity of different indexes of prostate specific antigen (PSA). Except for PSA change, sensitivity is the proportion of 171 known cancers cases for whom the index was positive and specificity is the proportion of 2011 men with normal transrectal ultrasound and digital rectal examinations not known to have prostate cancer who were negative on the index. The sensitivity and specificity of PSA change was evaluated in 84 men with prostate cancer and 1473 men without prostate cancer for on whom multiple PSA measures were available. A PSA level of 4.0 ng/ml or less was considered normal. A PSA density of 0.1 ng/ml per cubic centimeter of ultrasound-measured gland volume was considered normal. Age-referenced PSA was considered normal if it was 3.5 ng/ml or less in men aged 50–59, 4.5 ng/ml in men aged 60–69, and 6.5 ng/ml in men aged 70–79. PSA change was considered normal if the annual rate of PSA change was … or less per year.

Mettlin C, Littrup PJ, Kane RA, Murphy GP, Lee F, Chesley A, et al. Cancer. 1994;74: 1615–1620.

Prostate cancer is the most common type of cancer found in males. It is checked by PSA test that is notoriously unreliable. The probability that a noncancerous man will have an elevated PSA level >4.0 ng/mL is approximately 0.1, with this probability increasing to approximately 0.719 if the man does have prostate cancer. If, based on other factors, a physician is 50 percent certain that a male has prostate cancer, what is the conditional probability that he has the cancer given that the test indicates an elevated PSA level?

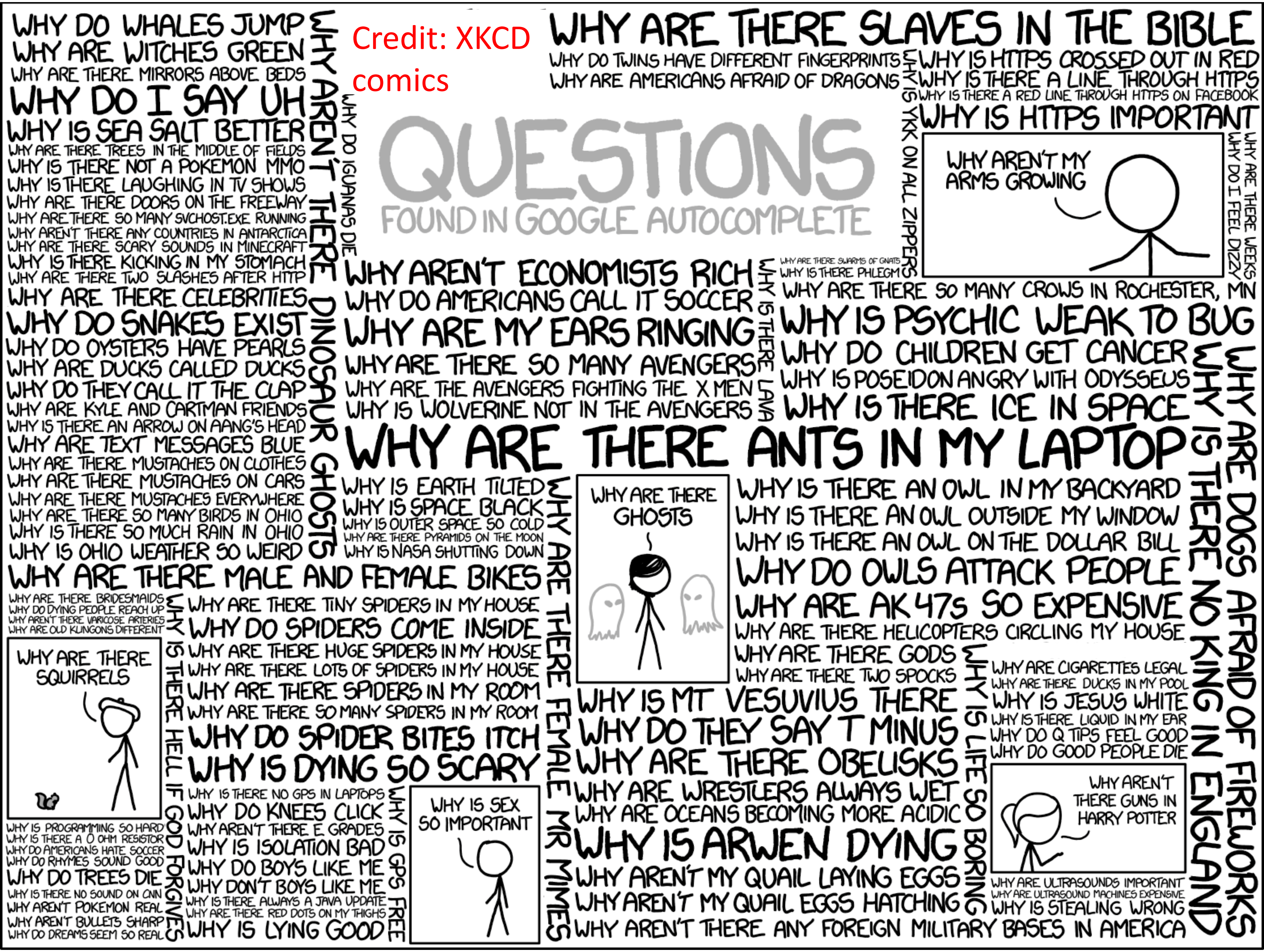
A. 99.9%
B. 95%
C. 88%
D. 55%

Get your i-clickers

All this trouble for a lousy 38% gain in confidence? I don't believe you!

- Let C – be the event that the patient has cancer; C' – patient is cancer free, E – events that the PSA test was elevated
- We know **doctor's prior belief**: P(C)=0.5
- We know test stats: P(E|C)=0.719, P(E|C')=0.1
- We need to find P(C|E)=P(E|C)*P(C)/P(E)
- P(E)=P(E|C)*P(C)+P(E|C')*P(C')= =0.719*0.5+0.1*0.5=0.41
- P(C|E)=0.5*0.719/0.41=0.88 or 88%

Credit: XKCD comics



Simpson’s paradox

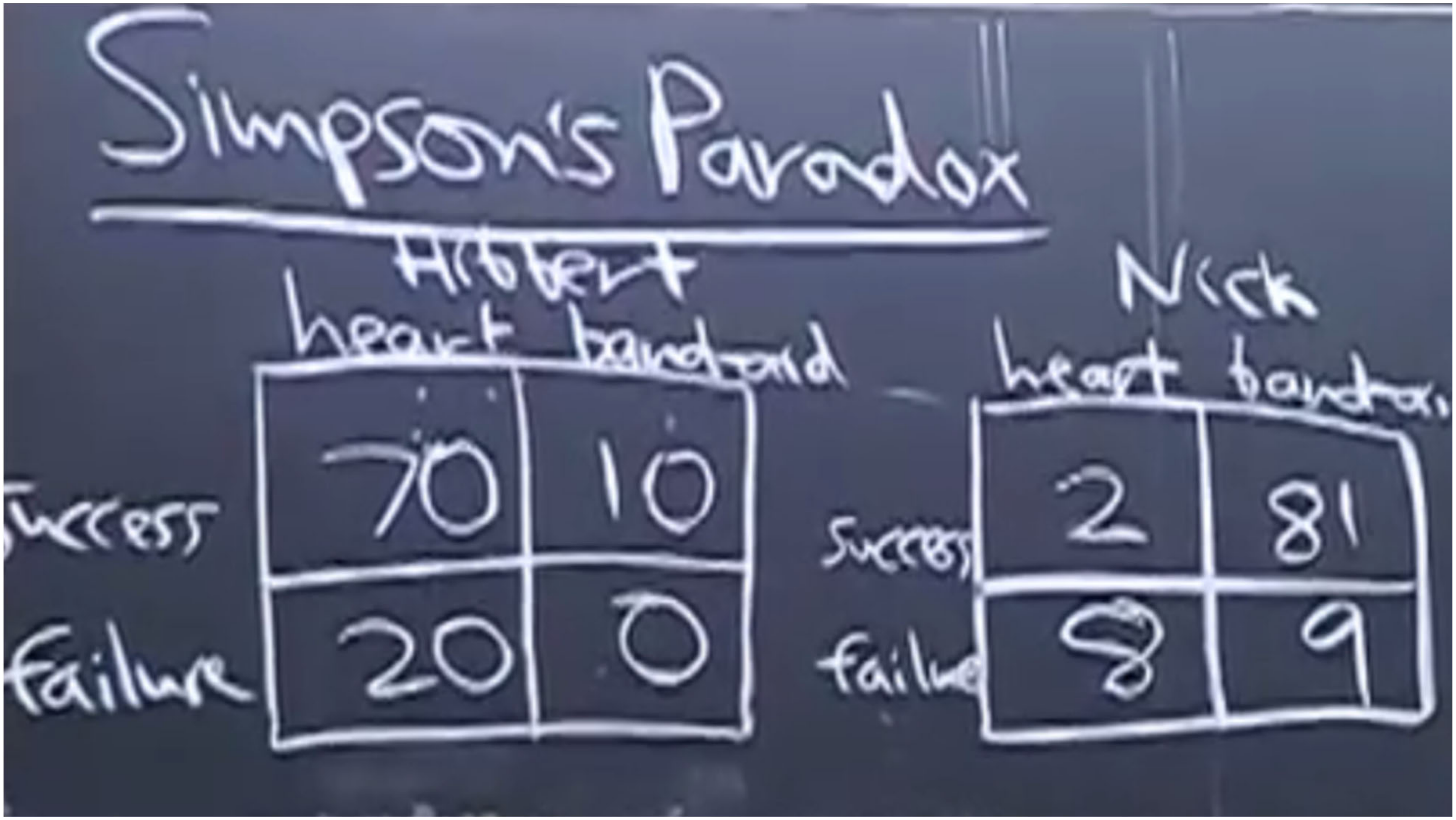
Edward Hugh Simpson
(10 December 1922 – 5 February 2019)
was a British codebreaker, statistician
and civil servant.

"The Interpretation of Interaction
in Contingency Tables“, Journal of the Royal Statistical Society, 1951

Is it possible for one doctor to have a **higher** success rate than another doctor in **every** type of treatment he performs but to have a **lower** **overall** success rate across all treatment types?

Dr. Hibbert

Dr. Nick

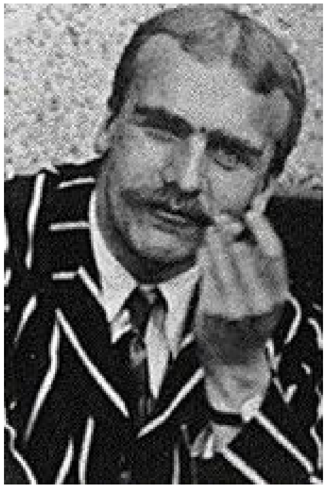


Dr. Hibbert: success rate =80%

Dr. Nick: success rate =83%

Simpson’s paradox might explain altruism

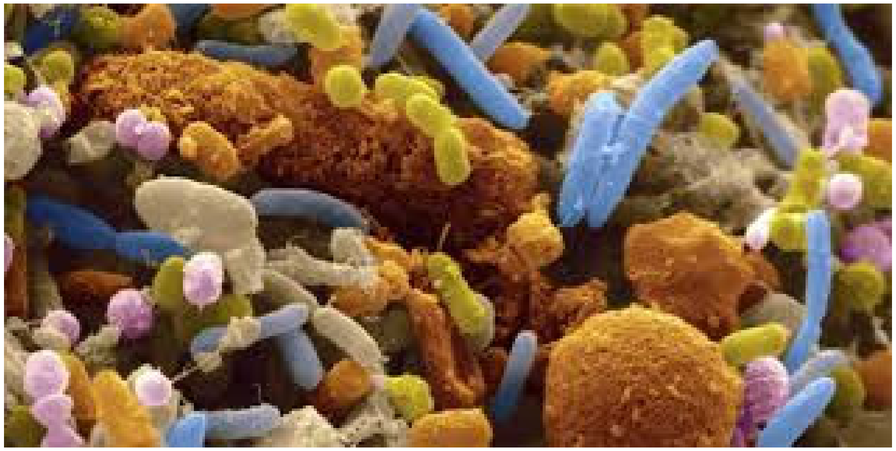
- Darwinian evolution has a problem with altruism
- “Selfish genes” do not care about others
- J. B. S. Haldane, (1892-1964) British geneticist, evolutionary biologist



- When asked if he would give his life to save a drowning brother answered: “No, but I would to save two brothers or eight cousins”
- Altruism in some insect colonies like ants is because they are all genetically similar.

Altruism in bacteria

- Bacteria live in communities in close proximity to each other
- Individual bugs spend significant resources to produce extracellular molecules, excrete them outside of the cell to share with others. **That slows their growth**
  - Examples: extracellular enzymes, biofilm components, antimicrobial and anti-immune agents
- Cheaters have faster growth rate
  - They can take over by not producing any shared molecules
- **Evolutionary paradox: how bacteria can be altruistic?**



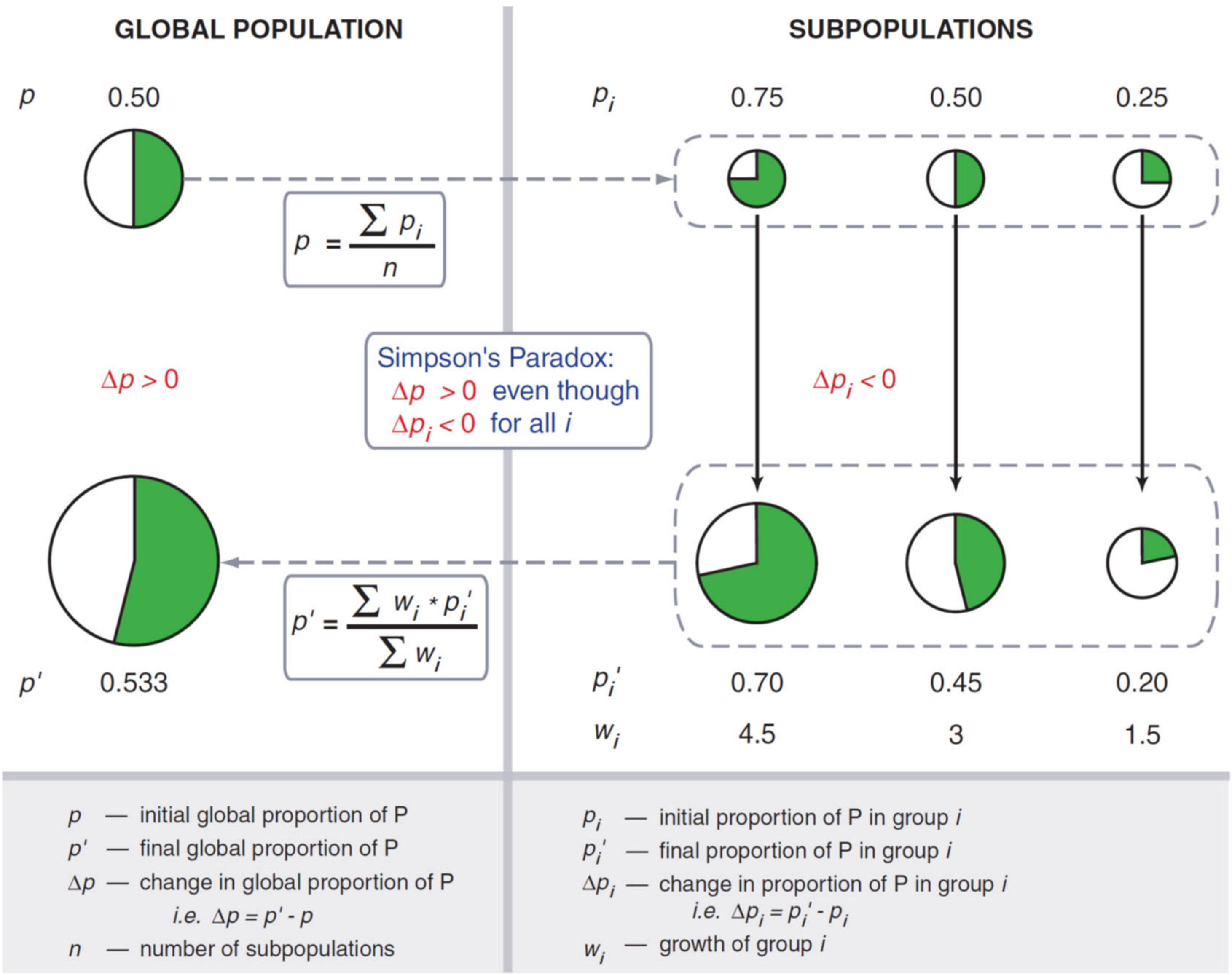
Chuang, Rivoire, and Leibler's answer

Simpson's Paradox in a Synthetic Microbial System

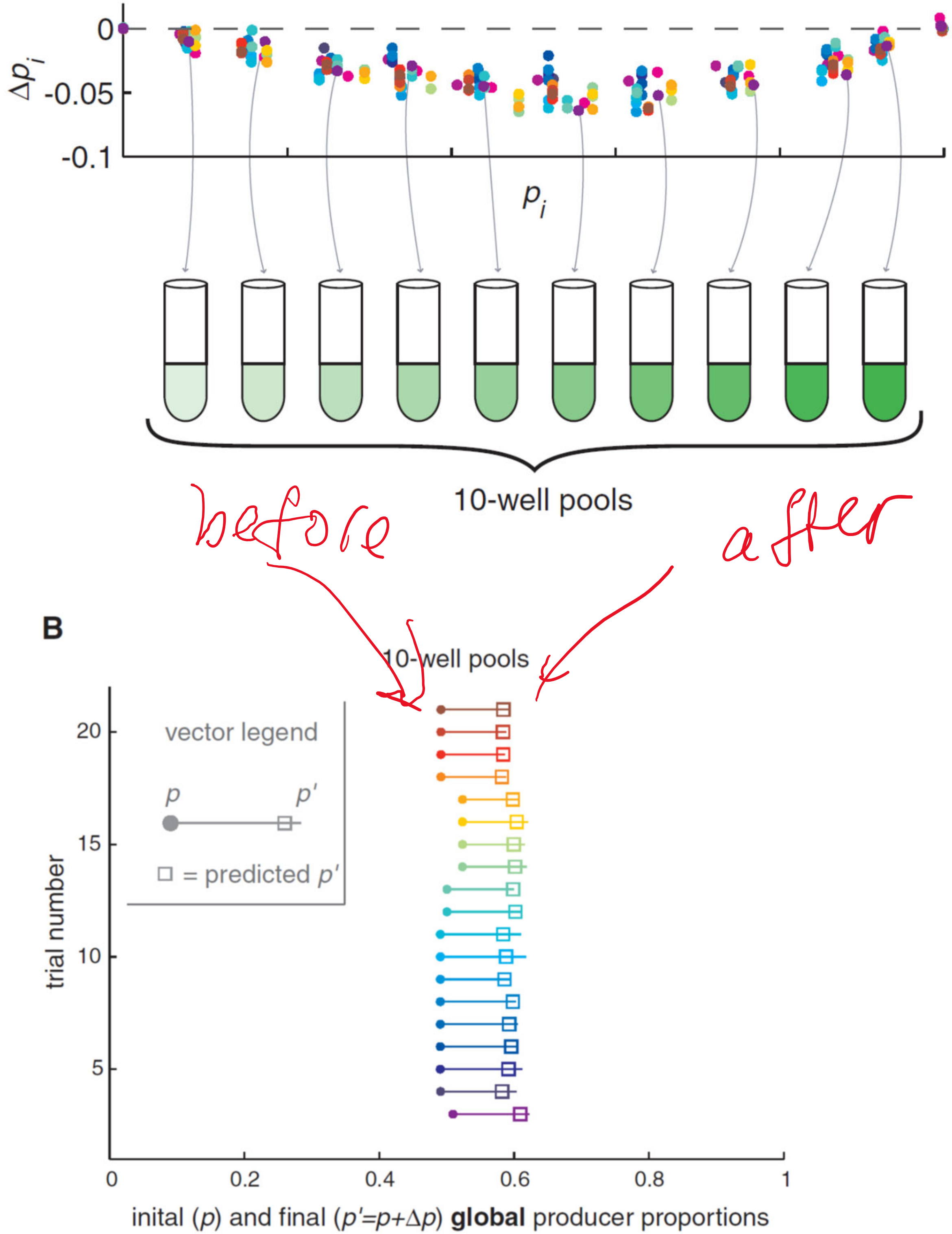
John S. Chuang,* Olivier Rivoire, Stanislas Leibler

The maintenance of "public" or "common good" producers is a major question in the evolution of cooperation. Because nonproducers benefit from the shared resource without bearing its cost of production, they may proliferate faster than producers. We established a synthetic microbial system consisting of two *Escherichia coli* strains of common-good producers and nonproducers. Depending on the population structure, which was varied by forming groups with different initial compositions, an apparently paradoxical situation could be attained in which nonproducers grew faster within each group, yet producers increased overall. We show that a simple way to generate the variance required for this effect is through stochastic fluctuations via population bottlenecks. The synthetic approach described here thus provides a way to study generic mechanisms of natural selection.

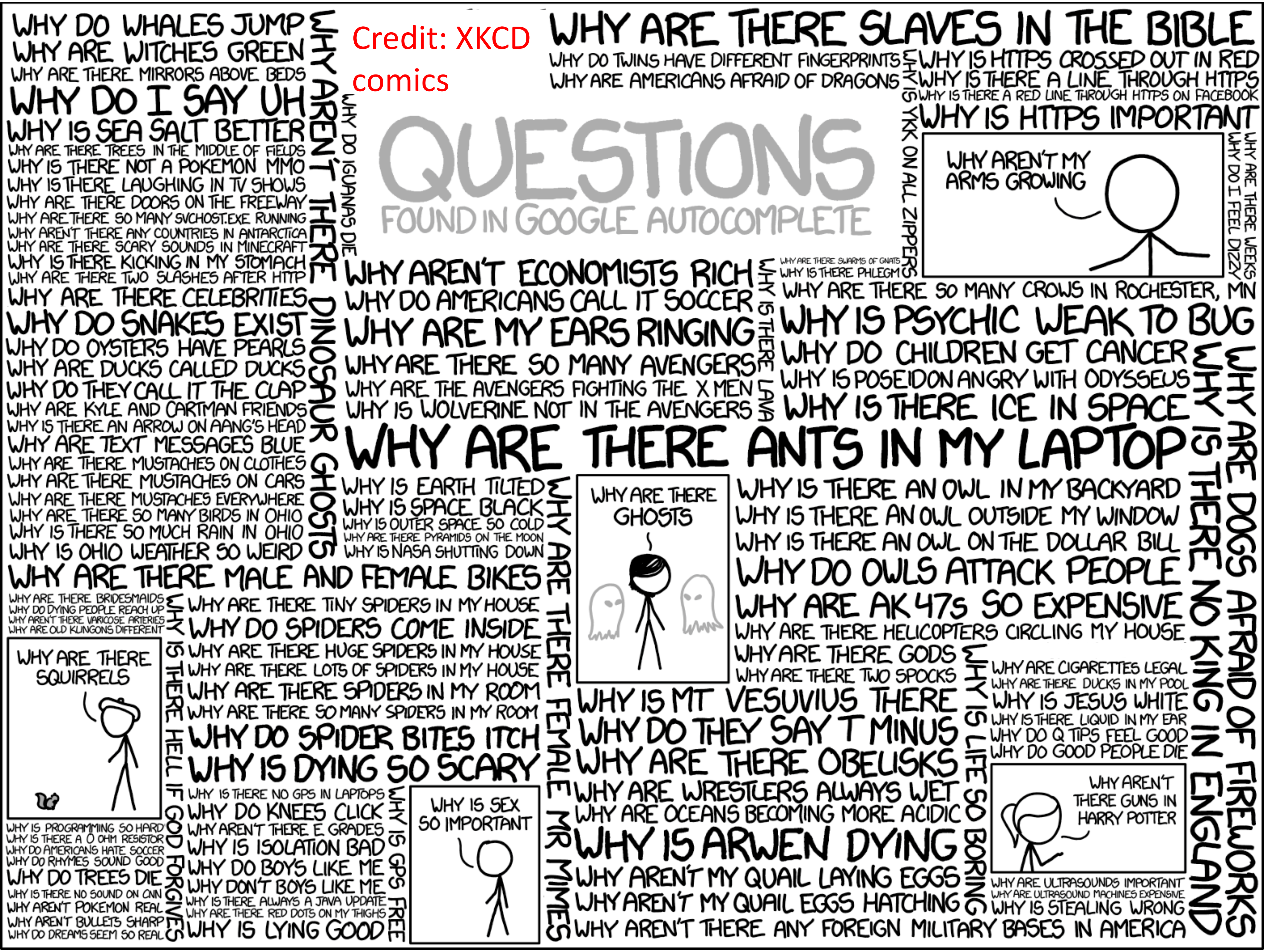
- The common good was a membrane-permeable Rhl autoinducer molecule rewired to activate antibiotic (chloramphenicol; Cm) resistance gene expression.



Fraction of altruists in each of individual test tubes dropped

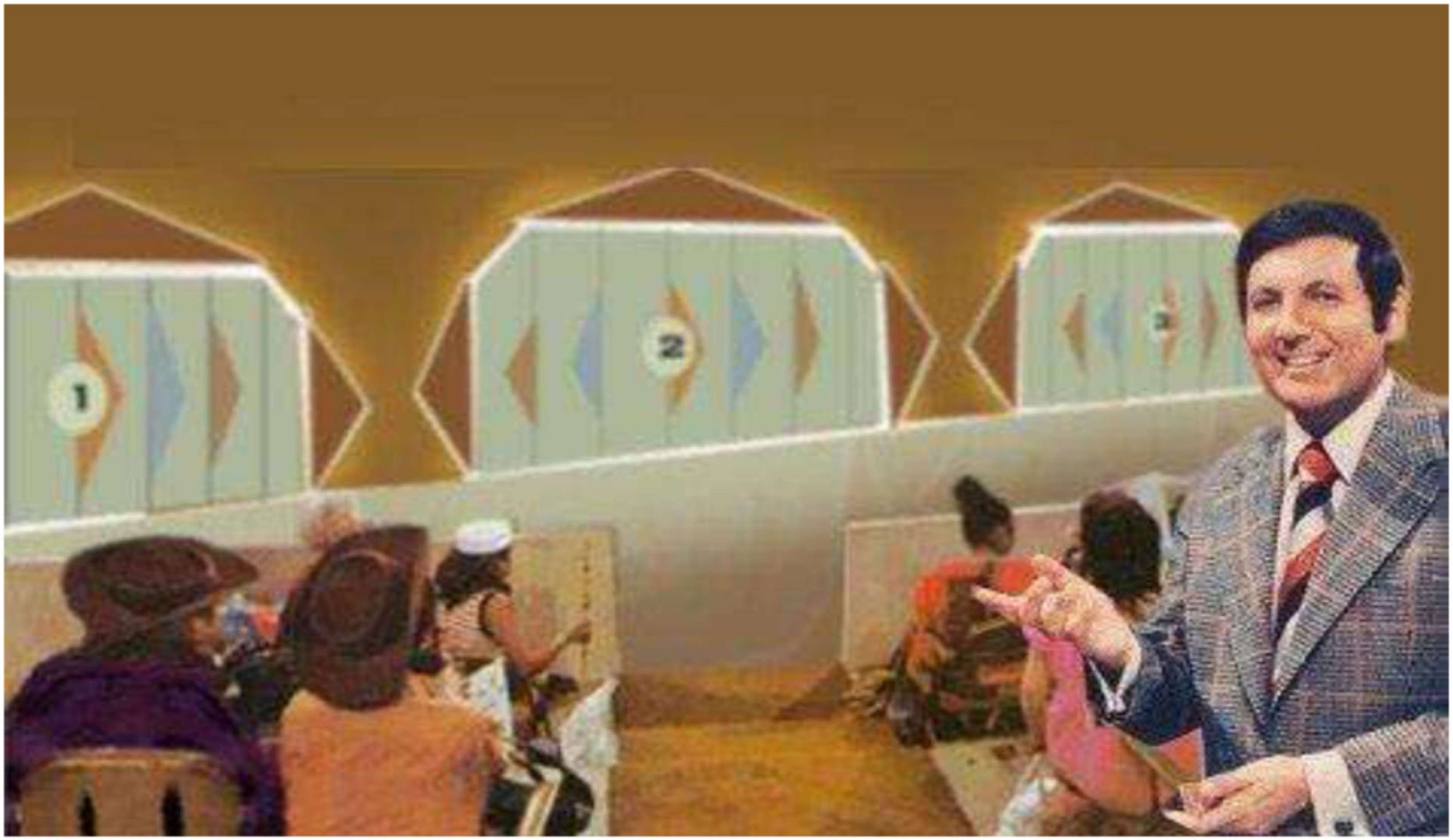


Yet the overall fraction of altruists in all test tubes combined increased



Credit: XKCD comics

Monty Hall problem



Monty Hall OC, OM (born Monte Halparin)

August 25, 1921 – September 30, 2017

was a Canadian-American game show host, producer, and philanthropist

Game show “Let’s Make a Deal” aired 1963-now

Monty Hall problem

- In *Make a Deal* there are three closed doors. Behind a random one of these doors is a car; behind the other two are goats. The contestant does not know where the car is, but Monty Hall does.
- After the contestant picks a door Monty always opens one of the remaining doors, one he knows does not hide the car. If the contestant has already chosen the door with the car behind, Monty is equally likely to open either of the two remaining doors.
- After Monty has shown a goat behind the door that he opens, the contestant is always given the option to switch doors.
- What is the probability of winning the car under the switching and non-switching strategies?

Monty Hall problem. What strategy gives you a better chance to win the car?

A. Better to switch doors
B. Better not to switch doors
C. Switching does not matter
D. The answer depends on the phase of the moon
E. I don't know

Get your i-clickers

Monty Hall problem.
What strategy
gives you a better chance
to win the car?

A. Better to switch doors
B. Better not to switch doors
C. Switching does not matter
D. The answer depends on the phase of the moon
E. I don't know

Get your i-clickers

Don’t feel bad if you guessed wrong

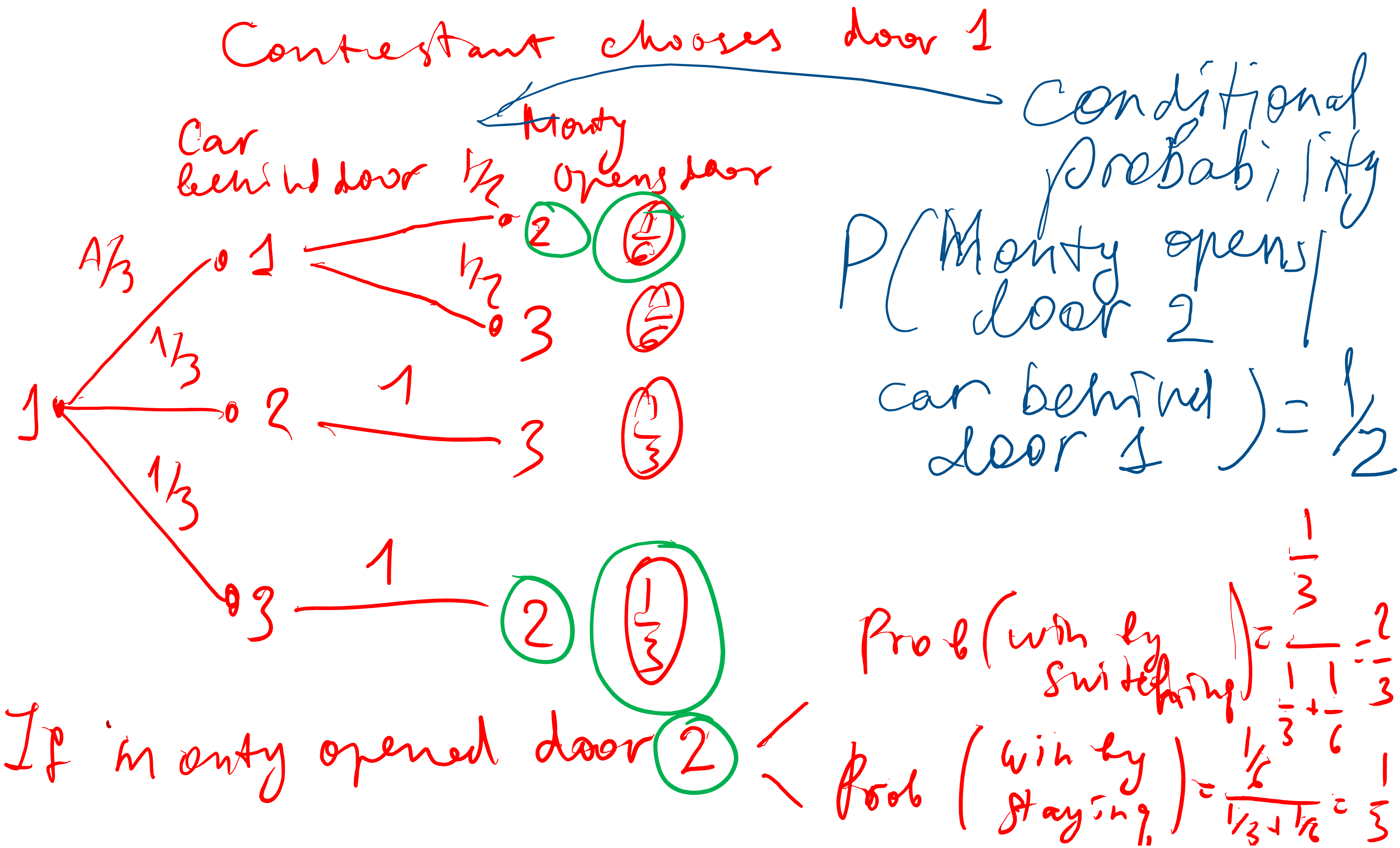
- When first presented with the Monty Hall problem an overwhelming majority of people assume that each door has an equal probability and conclude that switching does not matter
- Out of 228 subjects in one study, only 13% chose to switch
- Paul Erdős, one of the most prolific mathematicians in history, remained unconvinced until he was shown a computer simulation confirming the predicted result
- Pigeons repeatedly exposed to the problem show that they rapidly learn always to switch, unlike humans

Solution #1 (intuitive)

- With Prob=1/3 you guess the car door right: you loose the car if you switch
- With Prob=2/3 you got it wrong and picked a goat door. Then Monty opens another goat door. What is left is the car door. You win the car if you switch!

Solution #2.
Tree & conditional probabilities

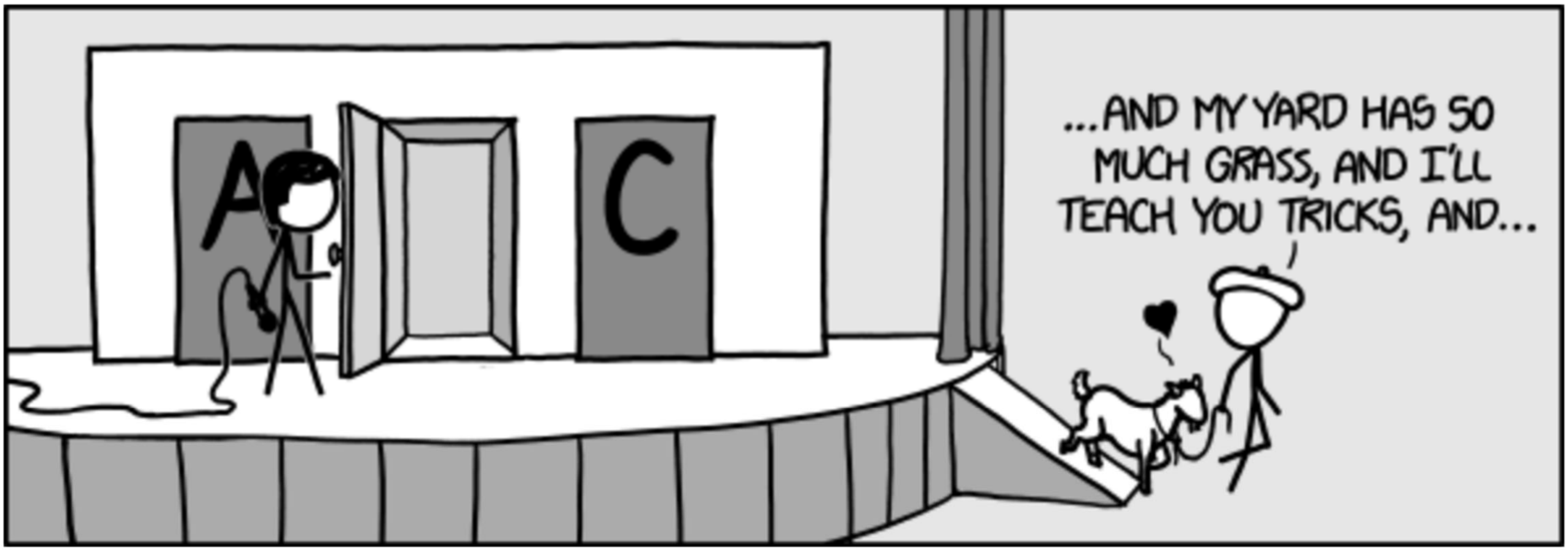
Solution #2.
Tree & conditional probabilities



A more detailed tree diagram

- Shinyapp website https://dacalderon.shinyapps.io/montyhall/

Thanks to my former BIOE 505 student,
Alejandra Zeballos Castro, for bringing it to my attention



comic credits: xkcd

Let’s check the theory by playing the dame

Go to https://dacalderon.shinyapps.io/montyhall/

- Right side will play “switch the door” strategy
- Left side of the auditorium will play “do not switch the door” strategy
- Play at least 30 rounds (more is better)
- In the end we will **add up the numbers from all tables**

```
switchers
 played:  270

 won:  176

 fraction won:  0.6519

non-switchers

 played:  100

 won:  31

 fraction won:  0.3100
```