Bayes’ Rule

Colin Jemmott

DSC96

Spring 2019

Bayesian reasoning is about how to revise our beliefs in the light of evidence.

Bayesian reasoning is about how to revise our beliefs in the light of evidence.

With Bayes' rule: more data = less wrong

Diseasitis

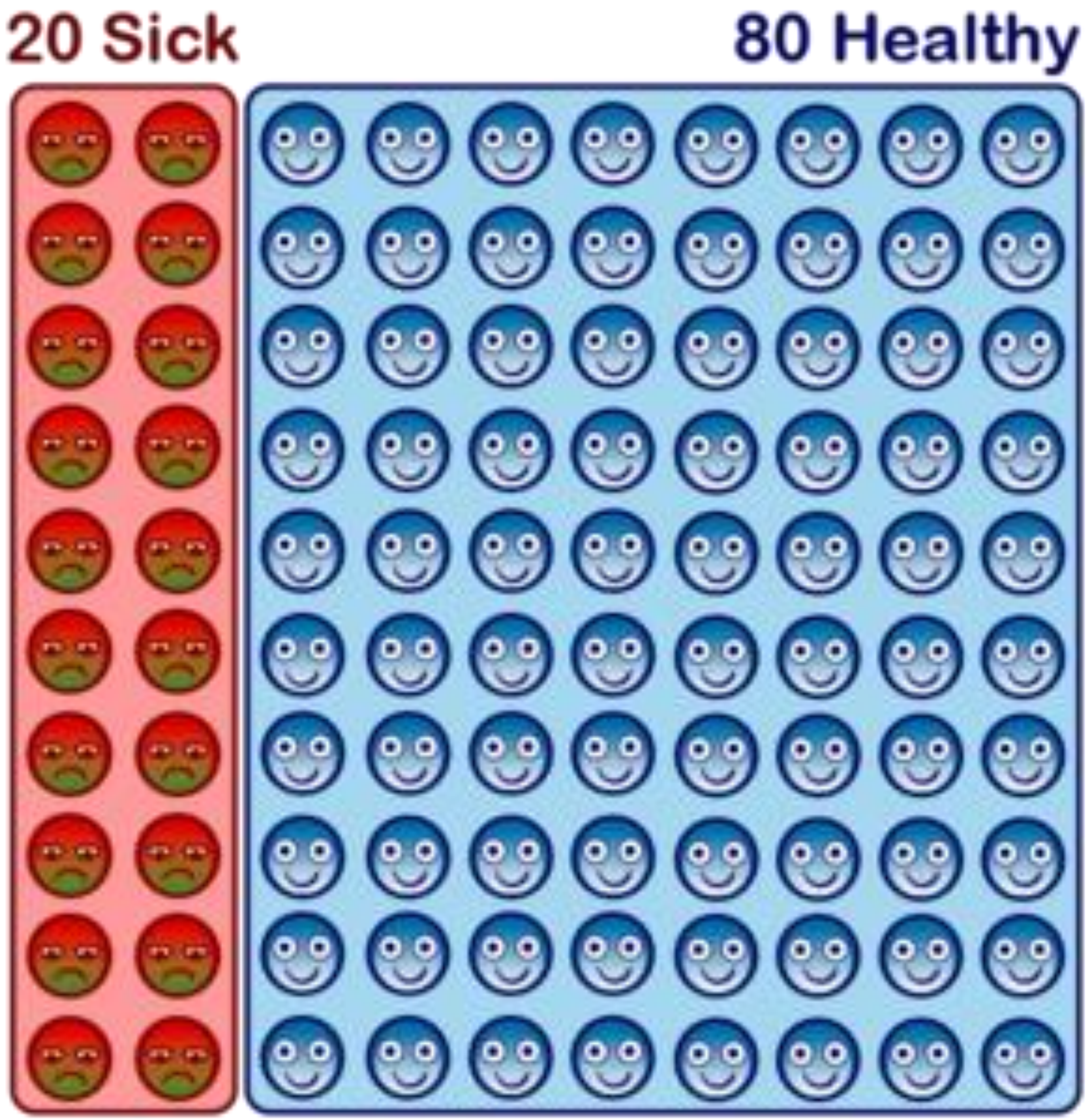
You are a school nurse testing for Diseasitis

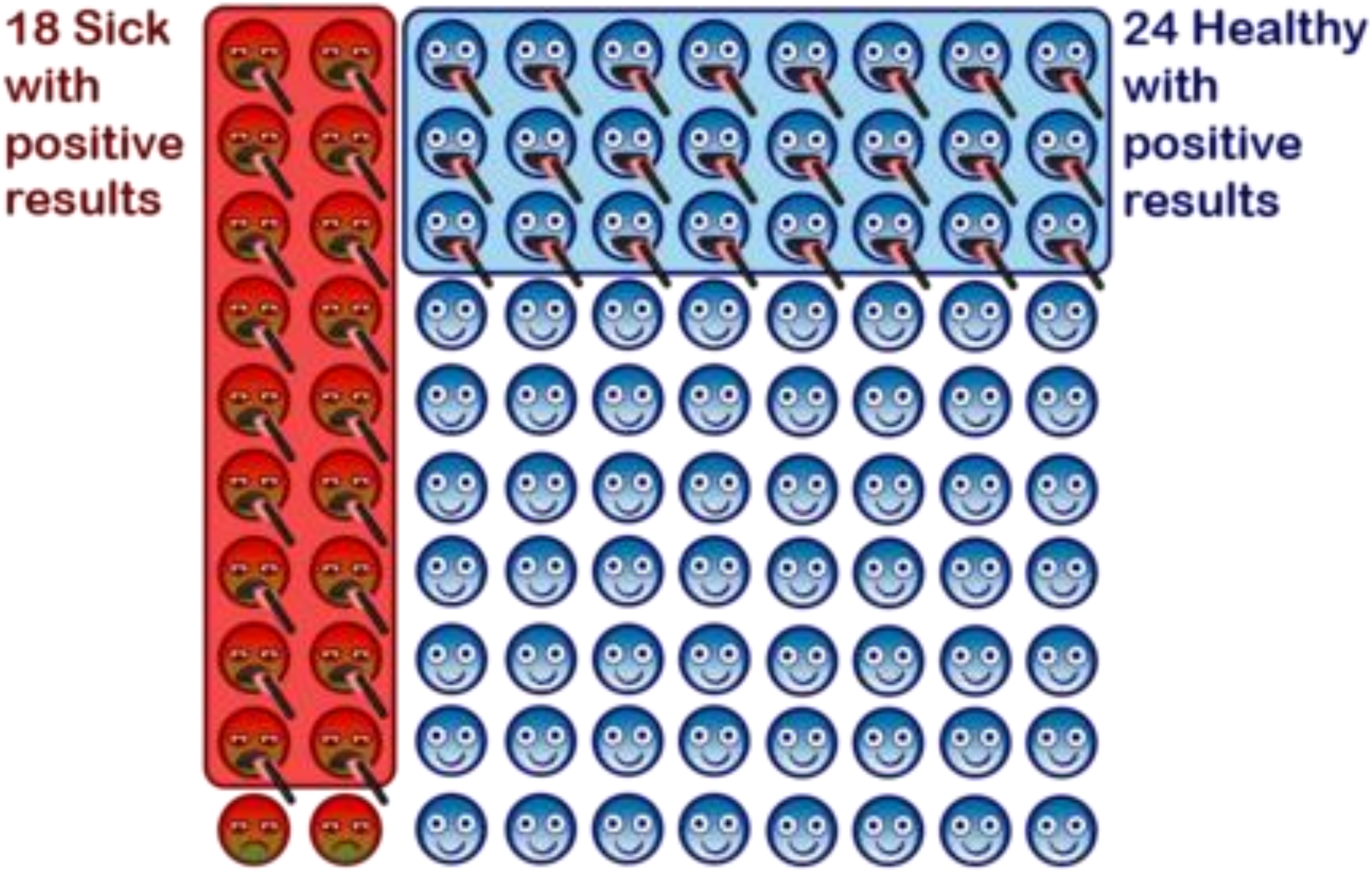
- ~20% of students have Diseasitis at this time of year

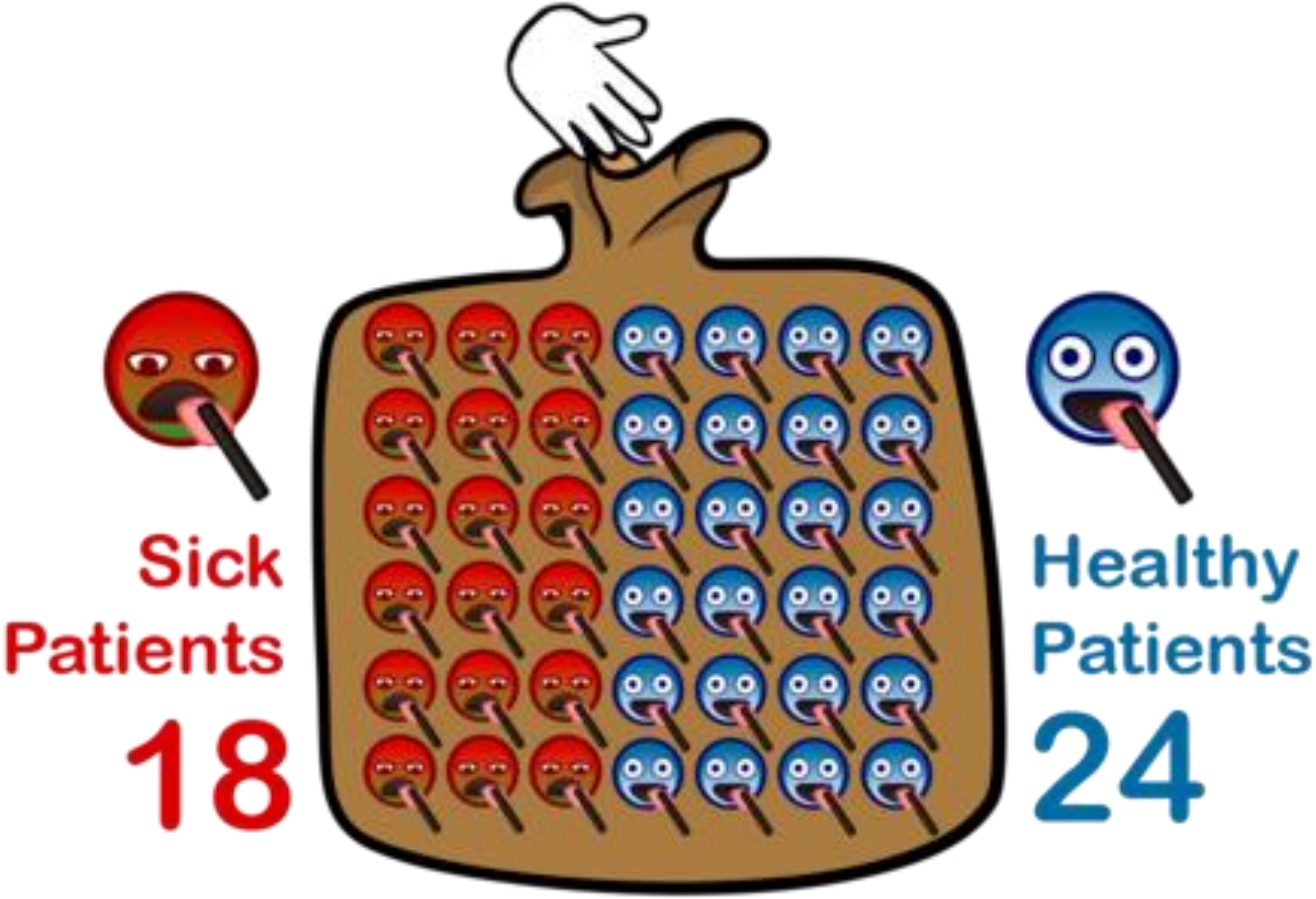
The test is a color-changing tongue depressor, which usually turns black if the student has Diseasitis.

- Among patients with Diseasitis, 90% turn the tongue depressor black.
- However, the tongue depressor is not perfect, and also turns black 30% of the time for healthy students.

One of your students comes into the office, takes the test, and turns the tongue depressor black. What is the probability that they have Diseasitis?


20 Sick
80 Healthy


18 Sick with positive results
24 Healthy with positive results


Sick Patients
18
Healthy Patients
24

Waterfall diagrams and relative odds

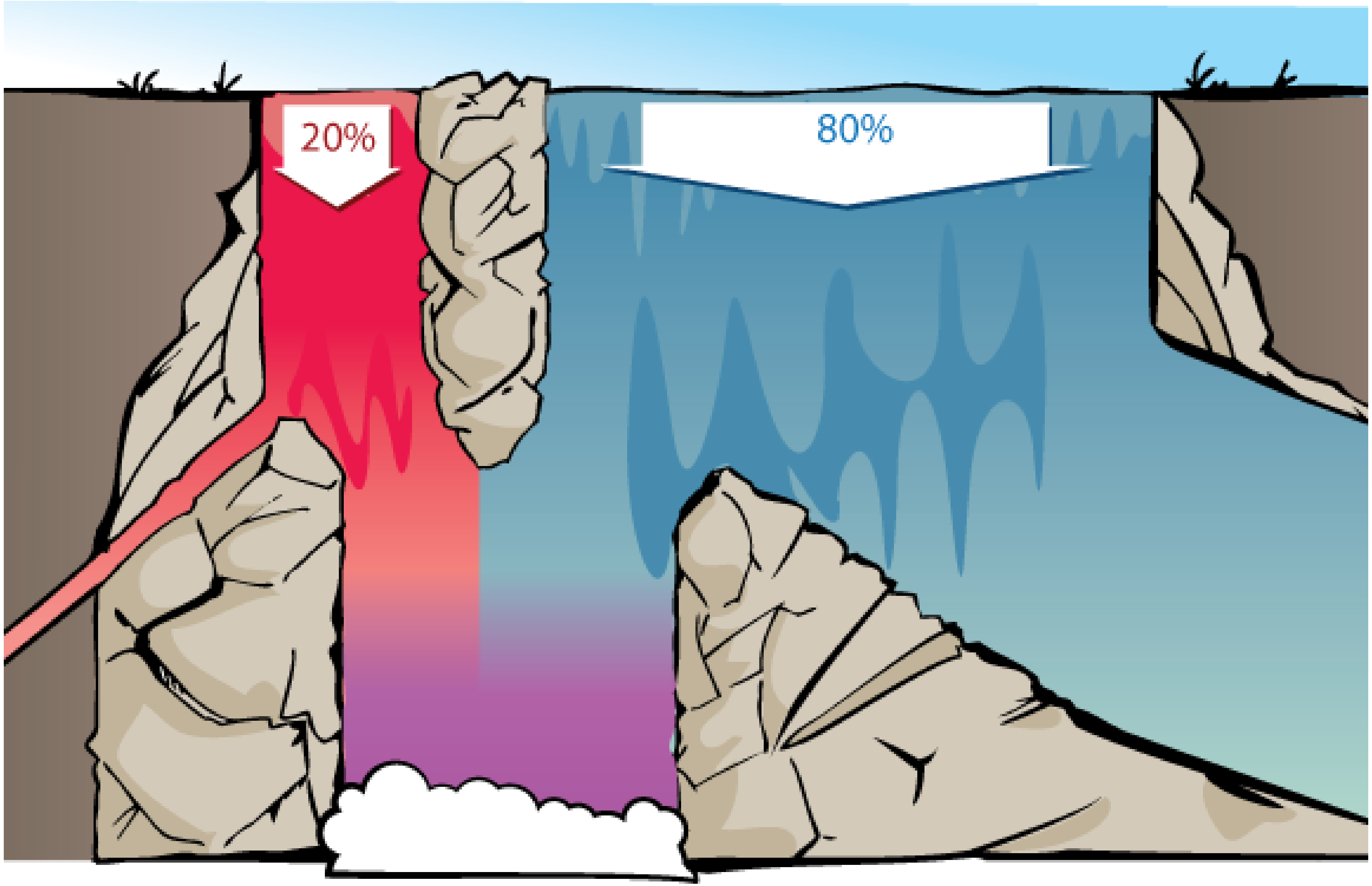
At the top of the waterfall:

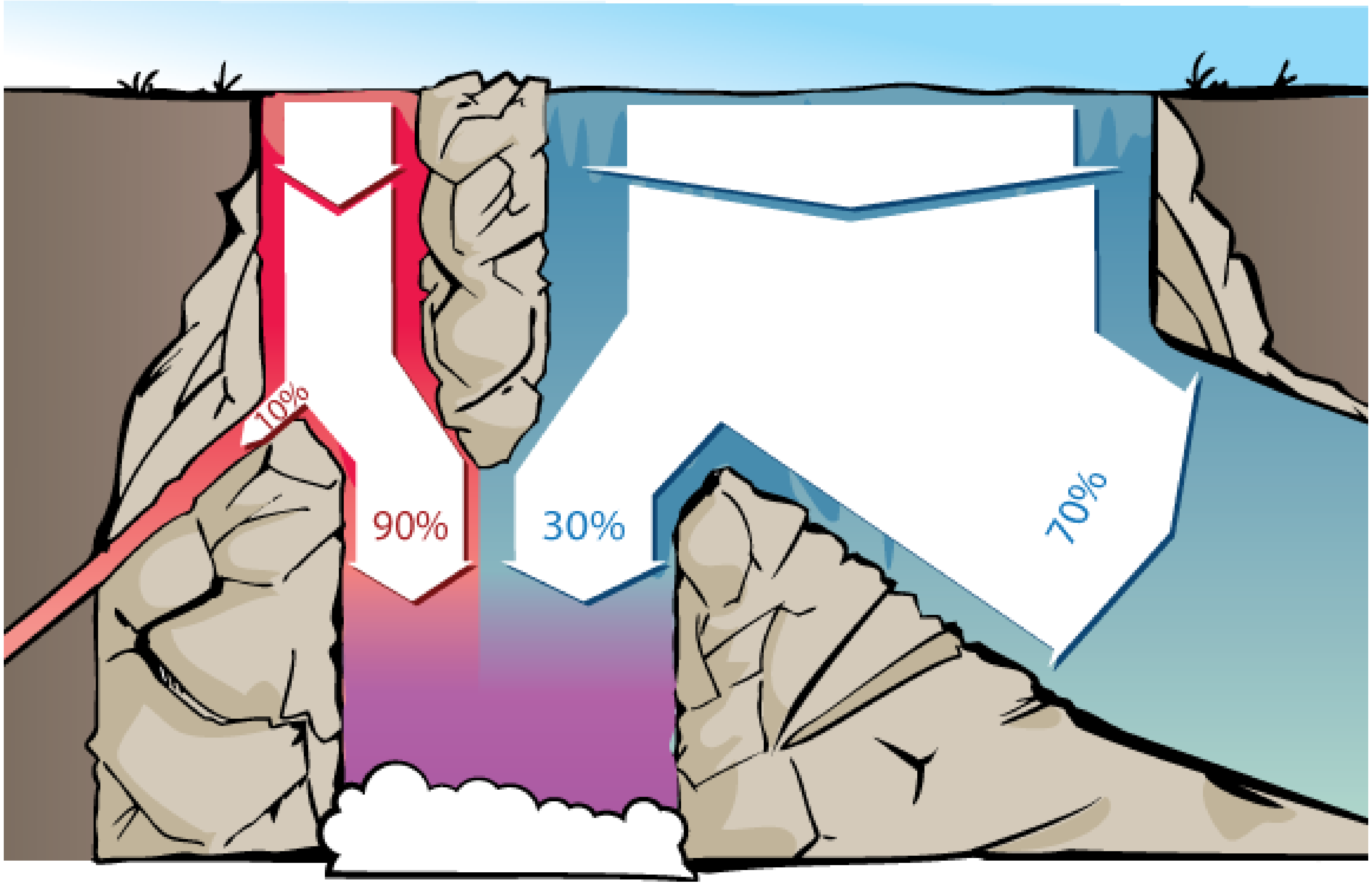
- 20 gallons/second of red water
- 80 gallons/second of blue water

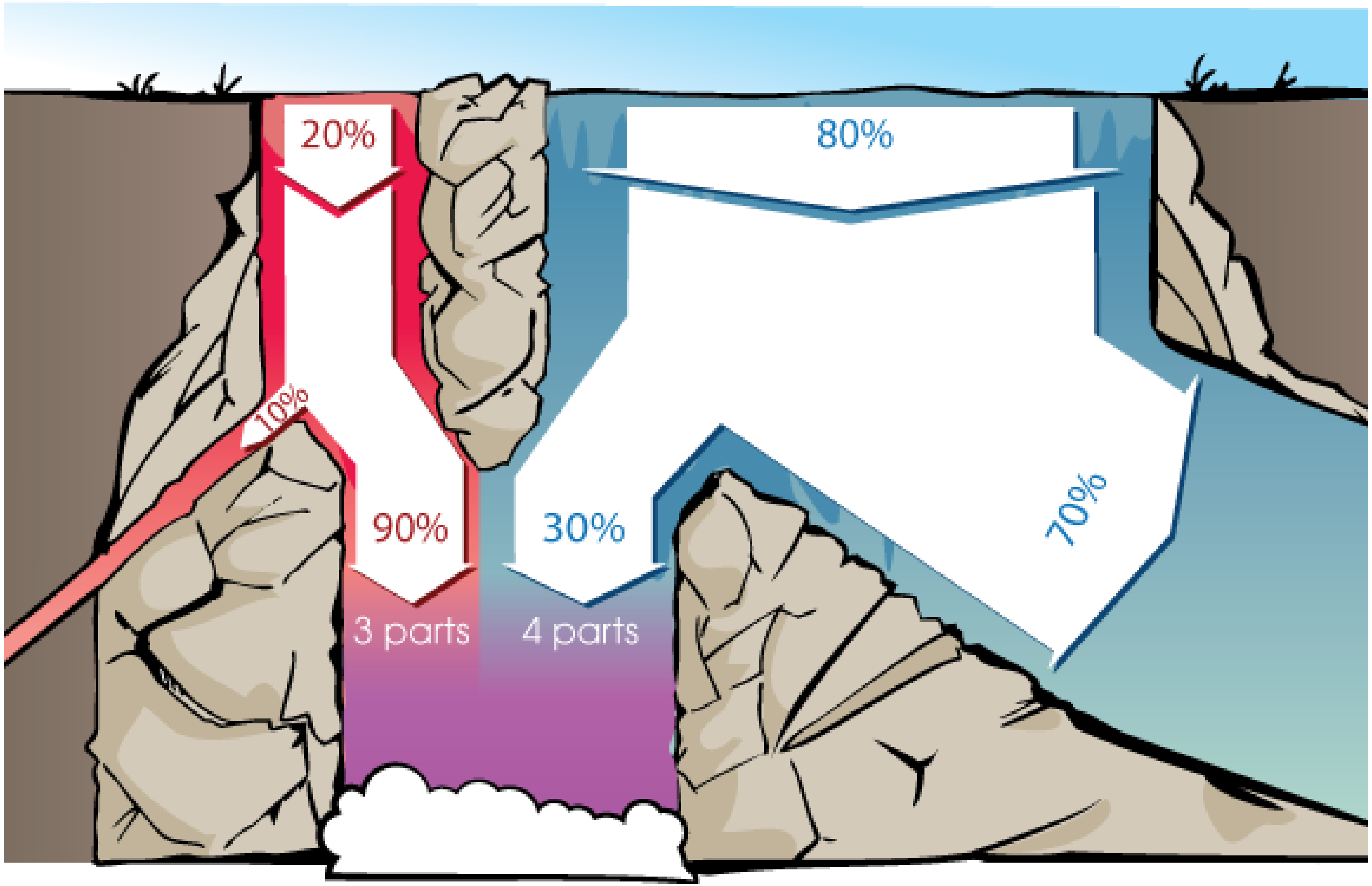
At the bottom:

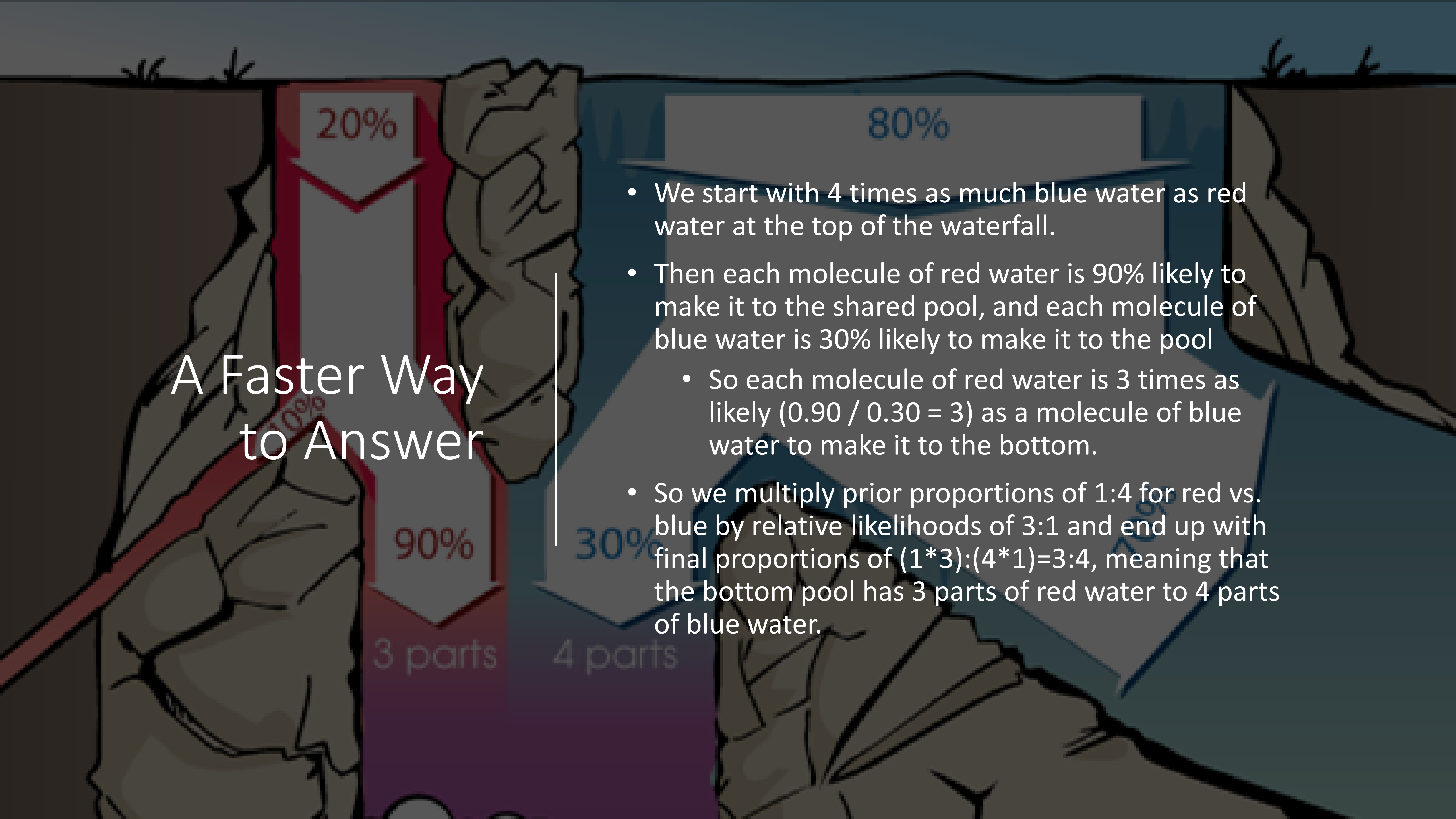
- 90% of the red water makes it to the bottom.
- 30% of the blue water makes it to the bottom.

- *Of the purplish water that makes it to the bottom of the pool, how much was originally from the red stream and how much was originally from the blue stream?*


20%
80%

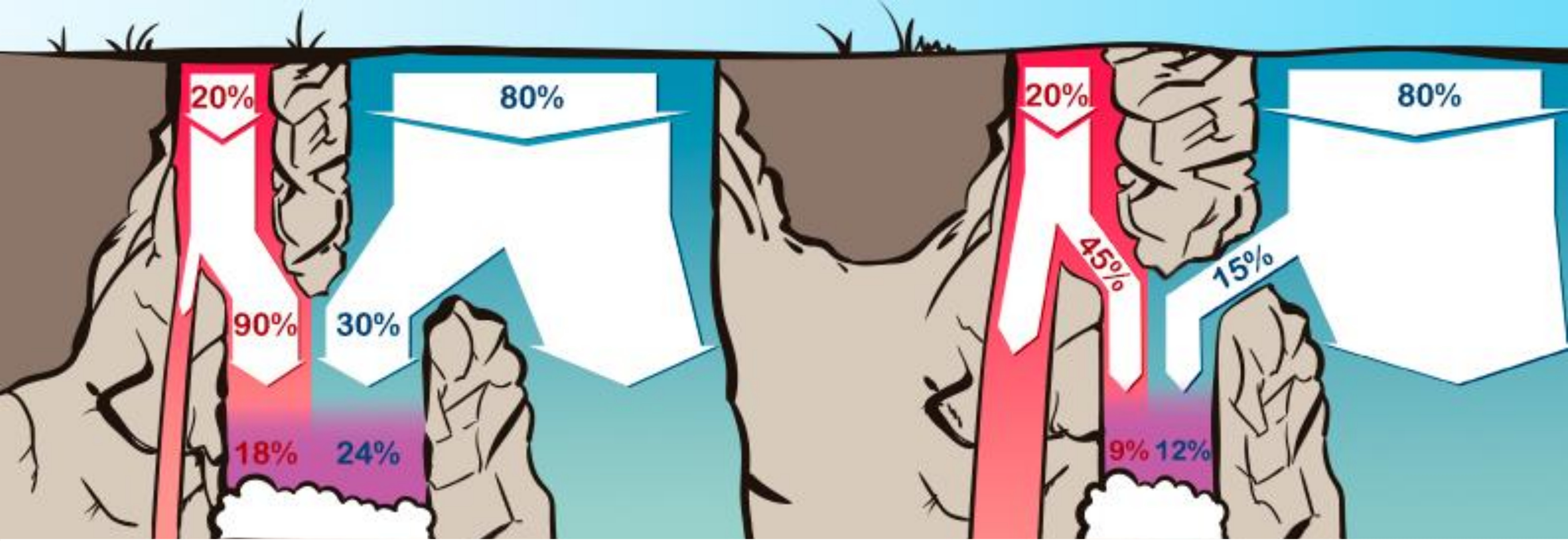

10%
90%
30%
70%


20%
80%
10%
90%
30%
70%
3 parts
4 parts



A Faster Way to Answer

- We start with 4 times as much blue water as red water at the top of the waterfall.
- Then each molecule of red water is 90% likely to make it to the shared pool, and each molecule of blue water is 30% likely to make it to the pool
  - So each molecule of red water is 3 times as likely (0.90 / 0.30 = 3) as a molecule of blue water to make it to the bottom.
- So we multiply prior proportions of 1:4 for red vs. blue by relative likelihoods of 3:1 and end up with final proportions of (1*3):(4*1)=3:4, meaning that the bottom pool has 3 parts of red water to 4 parts of blue water.



Prior odds

×

Relative likelihoods

=

Posterior odds

Probability

- *P(X)* will stand for the probability of *X*
- In other words, *X* is something that's either true or false in reality, but we are uncertain about it.
- *P(X)* is a way of expressing our degree of belief that *X* is true

A patient is, in fact, either sick or healthy; but if you don't know which of these is the case, the evidence might lead you to assign a 43% subjective probability that the patient is sick.

Conditional Probability

The Diseasitis involved some more complicated statements like:

- There is a 90% chance that a patient blackens the tongue depressor, *given* that they have Diseasitis.

In these cases we want to go from some fact that is *assumed* or *known* to be true (on the right), to some other proposition (on the left) whose new probability we want to ask about, taking into account that assumption.

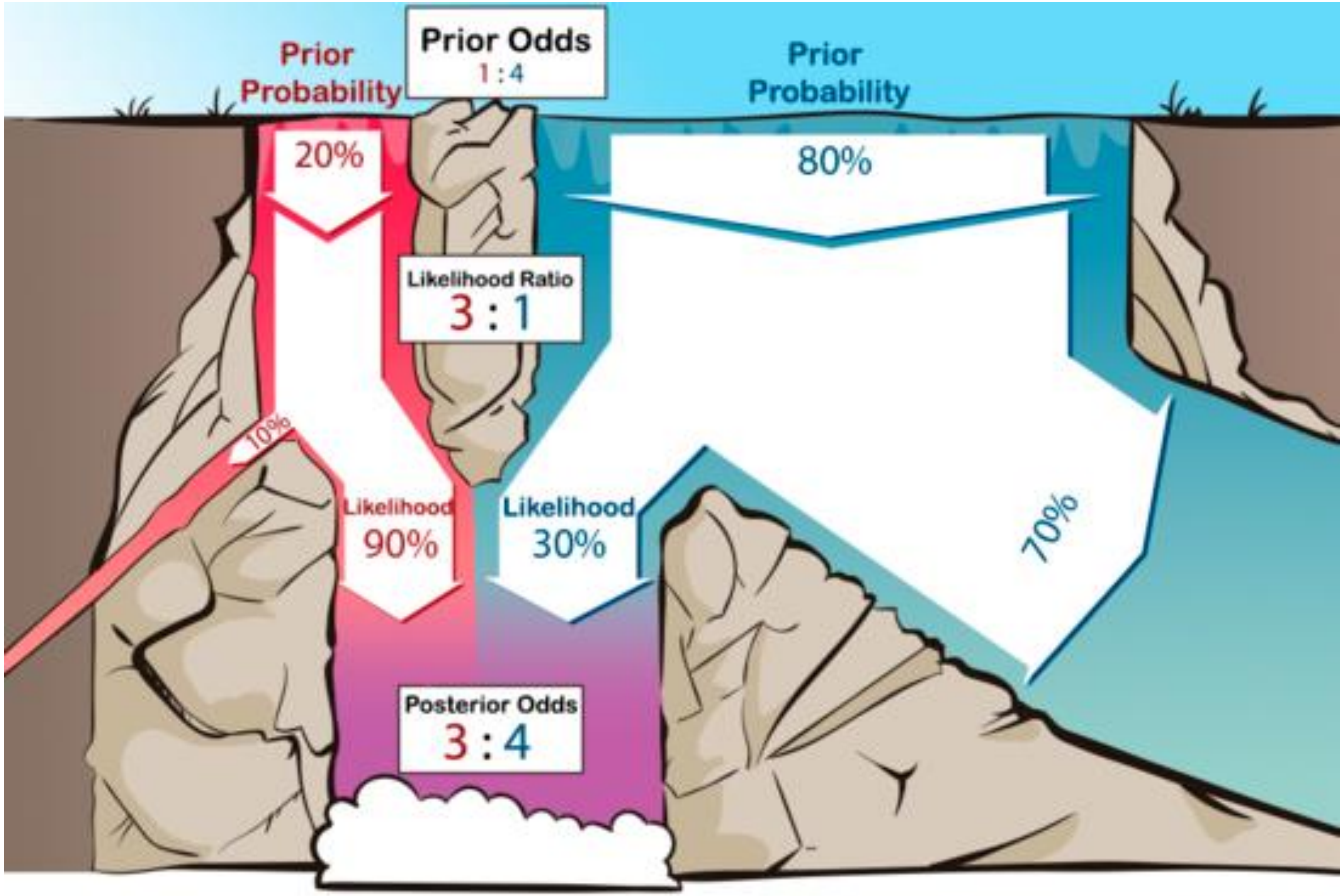
$$P(blackened|sick)=0.9$$

The standard notation for $P(X|Y)$ means "the probability of X, assuming Y to be true"

Bayes' Rule

$$\boxed{\text{Prior odds} \times \text{Relative likelihoods} = \text{Posterior odds}}$$

$$\frac{\mathbb{P}(sick)}{\mathbb{P}(healthy)} \times \frac{\mathbb{P}(blackened \mid sick)}{\mathbb{P}(blackened \mid healthy)} = \frac{\mathbb{P}(sick \mid blackened)}{\mathbb{P}(healthy \mid blackened)}$$



Posterior Probability = 3 / (3 + 4) = 3 / 7

Ok, but why do people talk about this:

Like they are in a cult?

Bayes' rule tells us how strong a piece of evidence has to be in order to support a given hypothesis.

Sparking Widgets

10% of widgets are bad and 90% are good. 4% of good widgets emit sparks, and 12% of bad widgets emit sparks. Can you calculate in your head what percentage of sparking widgets are bad?

The prior odds are 1 : 9 for bad widgets vs. good widgets.

12% of bad widgets and 4% of good widgets emit sparks, so that's a likelihood ratio of 3 : 1 for sparking (bad widgets are three times as likely to emit sparks).

posterior odds for bad vs. good sparking widgets. So 1/4 of sparking widgets are bad.

The evidence was weaker than the prior improbability of the claim.

Probability is the subjective experience of uncertainty.

Bookcase Aliens

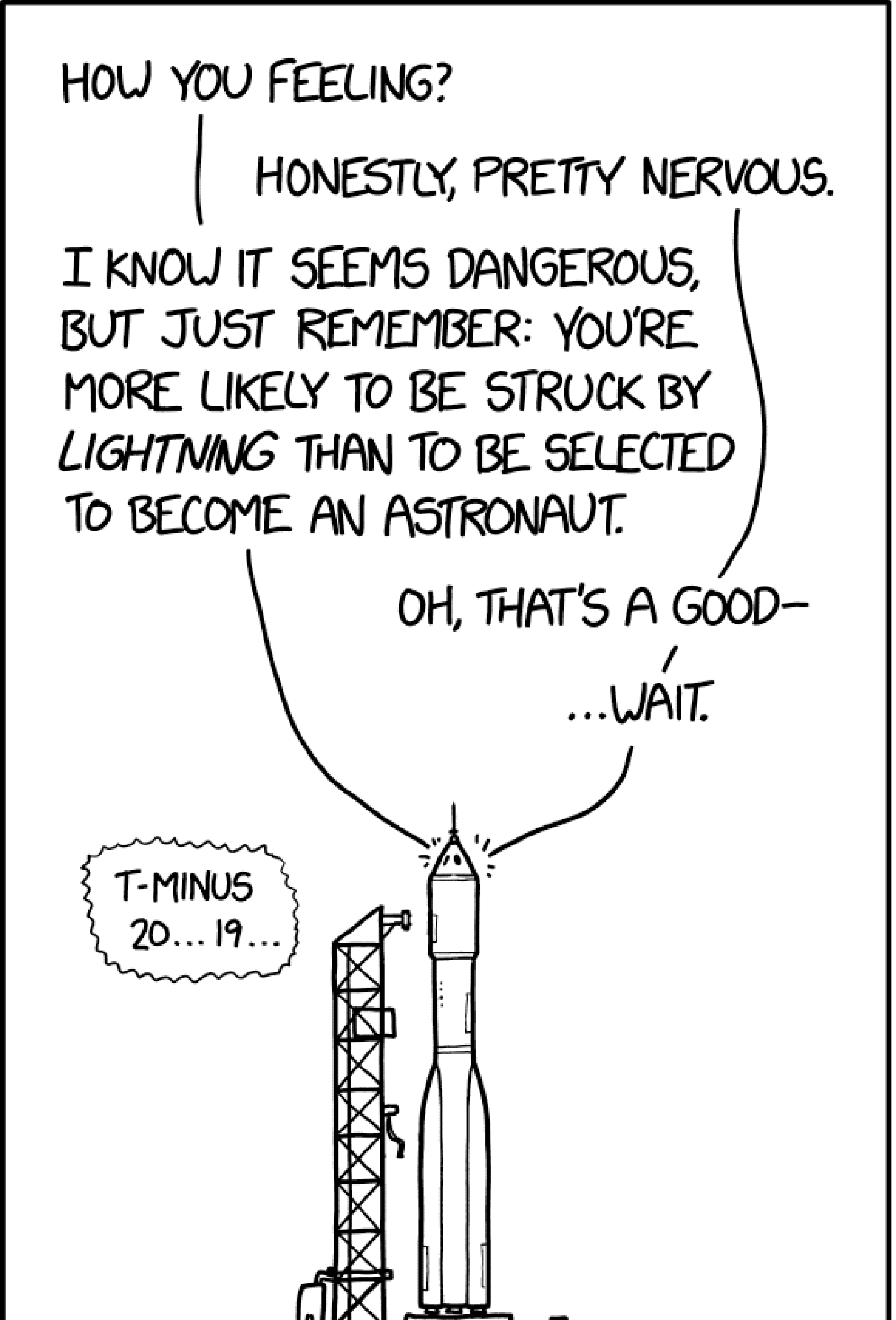
"Last week, I visited my friend's house, and there was a new bookcase there. If there were no bookcase aliens, I wouldn't have expected that my friend would get a new bookcase. But if there are Bookcase Aliens, then the probability of my finding a new bookcase there was much higher. Therefore, my observation, 'There is a new bookcase in my friend's house,' is strong evidence supporting the existence of Bookcase Aliens."

"There's a new bookcase in my friend's house", is indeed evidence favoring the Bookcase Aliens

- Depending on how long it's been since Bob last visited that house, there might be a 1% chance that there would be a new bookcase there
- On the other hand, the Bookcase Aliens hypothesis might assign 50% probability that the Bookcase Aliens would target this particular house among others

However, a reasonable prior on Bookcase Aliens would assign this a very low prior probability given our other, previous observations of the world.

- Let's be conservative and assign odds of just 1 : 1,000,000,000 against Bookcase Aliens.
- Then to raise our posterior belief in Bookcase Aliens to somewhere in the "pragmatically noticeable" range of 1 : 100, we'd need to see evidence with a cumulative likelihood ratio of 10,000,000 : 1 favoring the Bookcase Aliens.



Incremental Updating

- The observation of a new bookcase as being a single piece of evidence with a 50 : 1 likelihood ratio favoring Bookcase Aliens.
- The Bayesian doesn't toss the observation out the window because it's *insufficient* evidence; it just gets accumulated into the pool.
  - If you visit house after house, and see new bookcase after new bookcase, the Bayesian slowly, incrementally, begins to wonder if something strange is going on, rather than dismissing each observation as 'insufficient evidence' and then forgetting it.

Incremental Updating Example

You are visiting friends Andrew and Betty. They promised that one of them would pick you up from the airport when you arrive. You're not sure which one is in fact going to pick you up (50:50 prior), but you do know three things:

1. They have both a blue car and a red car. Andrew prefers to drive the blue car, Betty prefers to drive the red car, but the correlation is relatively weak. Andrew is 2x as likely to drive the blue car as Betty.
2. Betty tends to honk the horn at you to get your attention. Andrew does this too, but less often. Betty is 4x as likely to honk as Andrew.
3. Andrew tends to run a little late (more often than Betty). Betty is 2x as likely to have the car already at the airport when you arrive.

Let's say we see a blue car, already at the airport, which honks. Who is in the car? How sure are you?

Bayesian spam filters

- Start with 100 labeled emails:
  - 50 “Business”
  - 30 “Personal”
  - 20 “Spam”
  - The word "buy" has appeared in 10 of Business emails, 3 Personal emails, and 10 spam emails.
  - The word “hats" has appeared in 30 Business emails, 15 Personal emails, and 1 spam email.
- First, we assume that the frequencies in our data are representative of the 'true' frequencies
- Second, we make the *naïve Bayes* assumption that a spam email which contains the word "buy" is no more or less likely than any other spam email to contain the word “hats", and so on with the other categories.

Bayesian view of scientific virtues

Falsifiability: A good scientist should say what they do *not* expect to see if a theory is true.

Boldness: A good theory makes bold experimental predictions (that we wouldn't otherwise expect)

Precision: A good theory makes precise experimental predictions (that turn out correct)

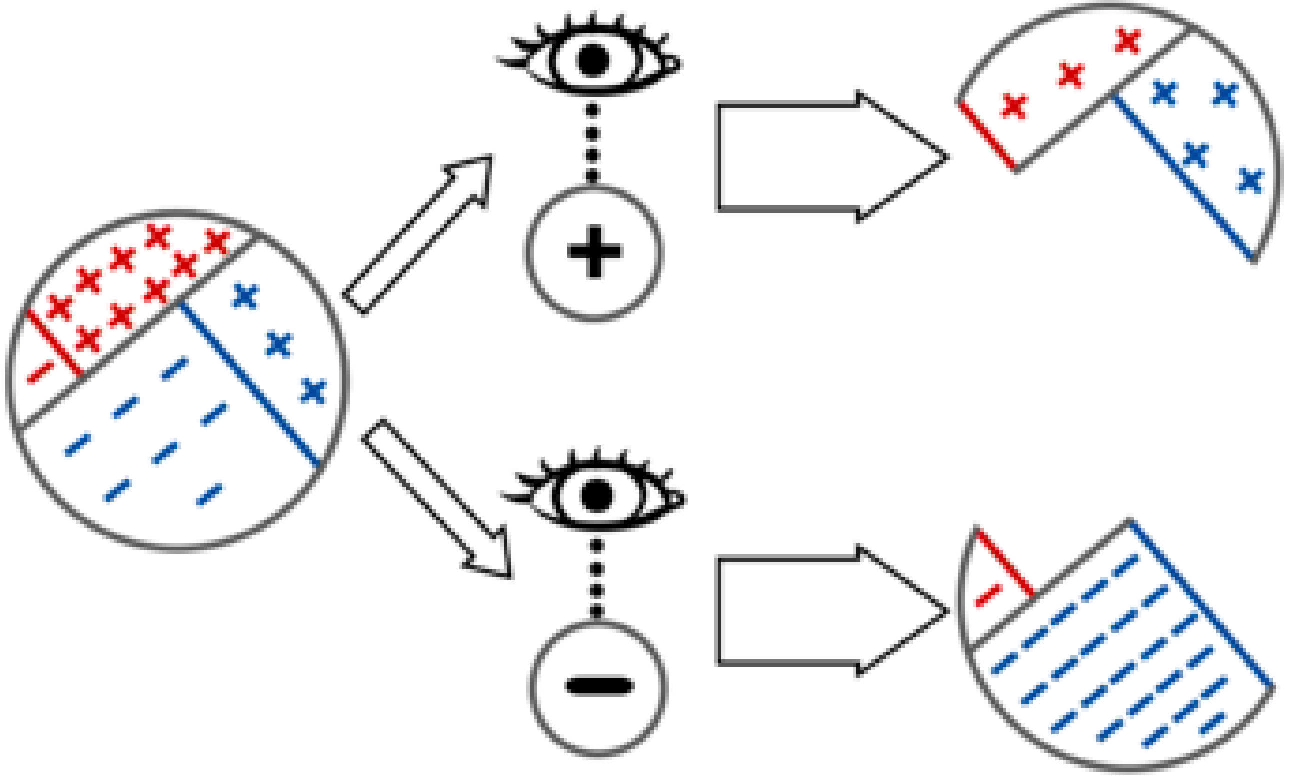
Experimentation: You find better theories by making observations, and then updating your beliefs.

Belief revision as probability elimination

We'll again start with the Diseasitis example:

- a population has a 20% prior prevalence of Diseasitis
- We use a test with a 90% true-positive rate
- And a 30% false-positive rate.

	Sick	*Healthy*
Test+	18%	24%
Test−	2%	56%



How to estimate probabilities

Prepare for revelation: What would you expect if you believed the answer to your question were about to be revealed to you?

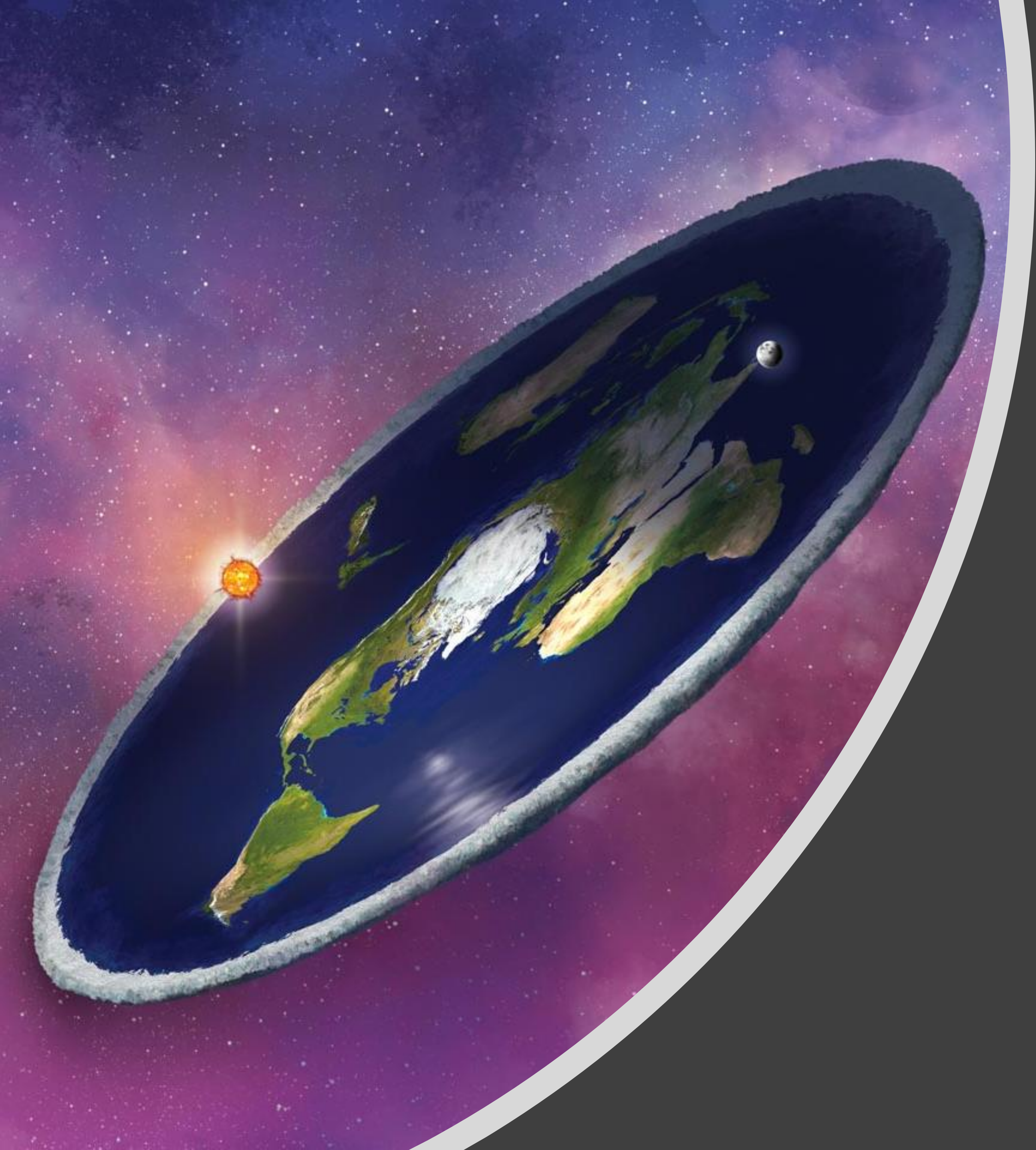
Bet on it: At what odds would you be willing to bet on a proposition?

Convert to a frequency: How many situations would it take before you expected an event to occur?

Find a reference class: How often have similar statements been true?

Make Multiple Statements: Imagine hypothetical evidence: How many statements could you make of about the same uncertainty as a given statement without being wrong once?

Imagine hypothetical evidence: How would your probabilities adjust given new evidence?



Make observations to update your beliefs.

ADVANCED NOTES

- map versus territory: But ignorance exists in the map, not in the territory. If I am ignorant about a phenomenon, that is a fact about my own state of mind, not a fact about the phenomenon itself.
- Rational thought produces beliefs which are themselves evidence.
- what is evidence?
- Calibration curves, priors and updating
- Bayes: strength of evidence as predictive accuracy
- Probability as extended logic