KEEP THIS TAB OPEN THROUGHOUT THE STUDY

Important: This material is the intellectual property of Stanford researchers. Please do not distribute it in any way.

Background to the Study

We all make judgments about probabilities. You might choose one job rather than another because of your judgment that you will probably be happier in that job. Or you might take some medication because of your judgment that it is probably safe.

This study is about probabilities.

However, humans are susceptible to various *cognitive biases*—that is, errors in their judgment.

Note that you will be asked questions about this and another problem later, and your ability to give correct answers in this study will determine whether you receive the bonus payment (if you are completing this study for payment).

Let us now consider the hypothetical scenario.

The Story of the Prisoners

Imagine that you and three other people—Alison, Billy and Carly—are in prison. Three of you will be imprisoned for life, and one of you will be set free. A lottery was used to

randomly determine who will be set free. So each of you have an equal chance of being set free at the beginning of this story.

The prison warden knows who will be set free, and you ask him if he can tell you who it is. He says that he can tell you the names of ***two*** prisoners who will ***not*** be set free, but he cannot tell you whether you will be set free or not. We will also suppose he cannot lie about who will be set free.

He then tells you that Billy and Carly will ***not*** be set free. Consequently, either you or Alison will be set free.

Now, once the warden has given you this testimony—that is, his statement about who will ***not*** be set free—which of the following is true: you are more likely to be set free, Alison is more likely to be set free, or both of you are equally likely to be set free?

At this point, an intuitive answer is that you and Alison are equally likely to be set free. After all, only two options remain, and you both started off with an equal probability of being set free. This answer, however, is incorrect. It results from a cognitive bias—an error in human judgment.

Surprisingly, the correct answer is that **Alison is more likely to be set free**.

This is partly because we need to first consider the *prior probabilities* of who will be set free—that is, the probability of being set free *prior* to receiving some evidence. In this case, the evidence is the warden's testimony that Bill and Carly will ***not*** be set free.

At the beginning of our story, then, there are four outcomes:

- Outcome 1 = You will be set free
- Outcome 2 = Alison will be set free
- Outcome 3 = Billy will be set free
- Outcome 4 = Carly will be set free

Remember that which prisoner will be set free is determined by a random lottery, so each person initially has an equal prior probability of being set free. For example, the prior probability that you will be set free is 1/4 or 25%, and it is the same with the other outcomes.

The second kind of probability that we need to consider is the probability of the testimony given the various outcomes. Recall that the testimony was this:

> Warden's testimony = the warden's statement that Billy and Carly will ***not*** be set free

Also recall that the warden said he cannot tell you whether you will be set free, and he can tell you the names of *only* two people who would ***not*** be set free. He then gave you his truthful testimony that Billy and Carly will ***not*** be set free.

For example, consider the outcome where Billy will be set free. If Billy was to be set free, then the warden would not have truthfully told you that Billy and Carly would not be set free. This is because we have supposed that the warden cannot lie. So there is a 0% probability that the warden would give you his testimony if Billy was to be set free.

But if Alison was to be set free, then there is a 100% probability that he would tell you that Billy and Carly will ***not*** be set free. And the reason for this is that he would **have** to tell you that Billy and Carly will ***not*** be set free: because he cannot lie, he would not say that Alison would ***not*** be set free if she was to actually be set free, and because he cannot tell you your fate, he cannot tell you that you will ***not*** be set free.

Let us consider the probability that he would give you the same testimony if ***you*** were to be set free. So now imagine that you will be set free. Then, the warden could have given you one of any three combinations of names about who will ***not*** be set free. He could have said:

1) that Alison and Billy will ***not*** be set free

2) that Alison and Carly will ***not*** be set free, or

3) that Billy and Carly will ***not*** be set free

Since there are three combinations which the warden could have said, the probability that the warden would tell you that Billy and Carly will ***not*** be set free is 1/3 if you were to be set free.

Ultimately, then, because the outcomes have an equal prior probability, and because the warden's testimony has a probability of 100% if Alison was to be set free but only a probability of 1/3 if you were to be set free, the probability that Alison will be set free given the warden's testimony is then 3/4, or 75%—but not 1/2 or 50% as many believe.