Prisoners and Hats

Grade Levels

This activity is adaptable for grades 6-12.

Objectives and Topics

The Prisoners and Hats puzzle is a logic puzzle that involves reasoning about the actions of other people, thus drawing in aspects of game theory. There are variations, but each involves having a predetermined number of people with colored hats placed on their heads. The objective is for someone to guess the color of the hat placed on their own head without looking at it.

Materials and Resources

You will need enough hats for all participants; half of these hats need to be red while the other half need to be white. (You could substitute other items for these hats as long as you have them in two distinct colors. Suggestions include circlets made out of pipe cleaners or pieces of colored paper taped to the participants' backs.)

Introduction and Outline

Story: The jail is full, so the jailer comes up with the solution of giving the prisoners a puzzle – if they succeed they can go free; if they fail they are executed.

Puzzle variations are listed in order of increasing difficulty. It is suggested that the rules be read aloud first and the participants set up according to the puzzle variation before placing any hats on the participants' heads.

Rules for Variations #1-4

1. Prisoners cannot look at their own hat.
2. Prisoners cannot communicate to one another (this includes talking, gesturing, writing, and signaling via physical contact).
3. Prisoners must raise their hand when they are ready to answer.
4. **If anyone figures out their own hat color, everyone goes free. If anyone guesses the wrong color, everyone is executed.**

Variation #1 (Warm-Up)

There are 2 prisoners, 1 red hat and 1 one hat. Both prisoners are in the same room facing each other (so they can see each other's hat but not their own). Who knows the color of their own hat?

Variation #2

There are 4 prisoners, 3 red hats and 1 white hat. One prisoner is taken to a separate room where he cannot be seen or heard. The three remaining prisoners are in the same room together where they all can see each other's hats (but still not their own). Who knows the color of their own hat?

Note: If it is impossible to isolate one of the prisoners, use just 3 prisoners but explain that you are still choosing from among 3 red hats and 1 white hat.

Variation #3

There are 4 prisoners, 2 red hats and 2 white hats. One prisoner is taken to a separate room where he cannot be seen or heard. The three remaining prisoners are in the same room together where they are lined up facing the same direction. They can only see the hats in front of them. Who knows the color of their own hat?

Note: Same comment as above.

Variation #4

There are 3 prisoners, 2 red hats and 3 white hats. All three prisoners are in the same room, lined up facing the same direction. They can only see the hats in front of them. Who knows the color of their own hat?

Rules for Variation #5

1. Prisoners cannot look at their own hat.
2. Prisoners cannot communicate to one another (this includes talking, gesturing, writing, and signaling via physical contact) unless they are told to do so.
3. **If you figure out your own hat color, only you go free. If you guess the wrong color, you are executed.**

Variation #5

All students participate in this variation. There is an unknown number of red hats and an unknown number of white hats. Everyone is in the same room, lined up facing the same direction. They can only see the hats in front of them. Starting with the prisoner in the back of the line and moving forward, each prisoner must say either "red" or "white". If the word matches their hat color, they go free. If it doesn't, they are executed. Prisoners are given some time to formulate a plan together before they are lined up.

Note: It is possible for all but one of the prisoners to survive. This one person will have a 50/50 chance of survival.

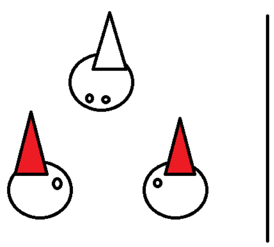
Possible Scenarios and Solutions

Variation #1

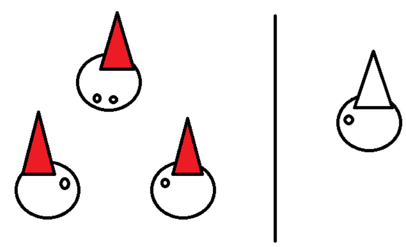
Either prisoner should state the color opposite of what he sees on the other prisoner's head.

Variation #2

There are two possible scenarios for this variation: *1*) the white hat is in the shared room, or *2*) the white hat is in the separate room.



Scenario 1

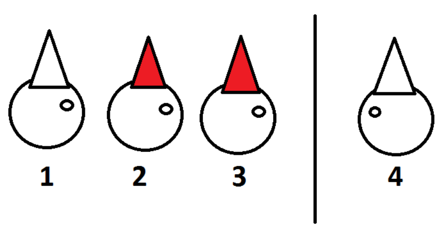


Scenario 2

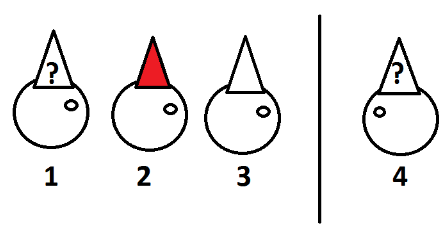
In Scenario 1, two of the prisoners in the shared room see that the single white hat is on someone else's head. Therefore either can conclude that he must have a red hat. In Scenario 2, all three prisoners in the shared room do not see the white hat, so all are uncertain. However after some time, since all seem uncertain, someone should deduce that everyone in the shared room is wearing a red hat.

Variation #3

There are several possible scenarios for this variation which could be boiled down to two: *1*) prisoners 2 and 3 are wearing hats of the same color, or *2*) prisoners 2 and 3 are wearing hats of opposite colors.



Prisoner 1 should answer

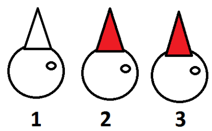


Prisoner 2 should answer

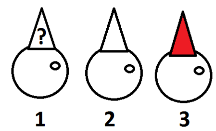
If prisoners 2 and 3 are wearing hats of the same color, then prisoner 1 (the person in back of the line) should conclude that he is wearing the opposite color. If prisoners 2 and 3 are wearing different colored hats, then prisoner 1 should remain silent as he is unsure. After some time, prisoner 2 should deduce that his hat and the hat in front of him are different colors; so he ought to say the color opposite of what he sees in front of him.

Variation #4

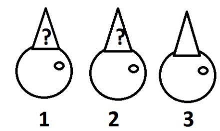
For this variation, the key is knowing how many red hats are in front of you.



Prisoner 1 should answer



Prisoner 2 should answer



Prisoner 3 should answer

If there are two red hats in the front of the line, then prisoner 1 can conclude with certainty that he is wearing a white hat. If there is only one red hat at the front of the line, then prisoner 1 remains quiet and prisoner 2 should conclude that he is wearing a white hat. If there are no red hats at the front of the line, then prisoners 1 and 2 both remain quiet, and prisoner 3 should conclude that he is wearing a white hat.

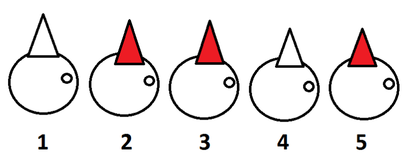
Variation #5

The prisoners agree that the first person (in the back) will follow a rule:

- If he sees an odd number of red hats, he will say "red".
- If he sees an even number of red hats, he will say "white".

By doing so, the subsequent prisoners will know the odd or even nature of the red hats, and thus be able to compare it with the number of red hats in front of them.

Example:



Prisoner 1 Follows the rule – he sees an odd number of red hats in front of him, so he says "red". He is incorrect, so he is executed.

Prisoner 2 Knows that including himself there is an odd number of red hats. He sees two red hats in front of him, thereby concluding he has a red hat. He survives.

Prisoner 3 Knows that at the start there was an odd number of red hats and now there is one less. Therefore, including himself there is an even number of red hats. He sees one red hat in front of him, so he concludes that he must also have a red hat. He survives.

Prisoner 4 Knows that at the start there was an odd number of red hats and now there are two less. Therefore, including himself there is an odd number of red hats. He sees one red hat in front of him, so he concludes that he must have a white hat. He survives.

Prisoner 5 Comes to the same conclusion as prisoner 4 – that there ought to be an odd number of red hats left. Since he is the last person, he states that he has a red hat. He survives.

Instructor Notes

- Have students close their eyes while the hats are being distributed.
- If there are more hats than students, hide the remaining hats so that the students will not be given extra clues as to which hats are being used.
- Be on the lookout for students signaling answers to each other. They are CHEATING.
- All solutions are merely suggestions.
- Variation #5 is difficult enough where you could explain the entire even/odd method and then just have your students try it out.