The Blue-Eyed Islanders Logic Puzzle

This is an old-ish logical puzzle. I'm putting this here to memorialize my take on the solution and in particular on the information conveyed by the "foreigner." The formulation of the puzzle below is adapted from Terry Tao's blog (modified to be somewhat less gruesome). There is also an XKCD formulation.

Statement of Puzzle

- A tribe resides on a little-visited tropical island.
- The tribe comprises 100 people, most with brown eyes (as is nearly universal in that part of the world), but a few with blue eyes.
- Their religion forbids them to know their own eye color, or even to discuss the topic in any way.
    - Thus, each islander can and does see the eye colors of all other residents, but has no way of discovering his or her own (there are no reflective surfaces, or it is forbidden to gaze directly into a reflective surface).
- If a resident does discover his or her own eye color, then their religion compels them to leave the island forever the next day on a ferry, carrying a large sign saying, "I found out my eye color. Farewell, dear friends." However, this has never happened — yet.
- All the tribespeople are highly logical and devout, and they all know that each other is also highly logical and devout (and they all know that they all know that each other is highly logical and devout, and so forth).
    - "Highly logical" means that any conclusion that can logically deduced, from the information and observations available to an islander, will instantly be known to that islander. (Think Mr. Spock or Mr. Data.)
- Of the 100 islanders, it happens that 10 of them have blue eyes and 90 of them have brown eyes, although the islanders are of course not fully aware of these statistics (each of them can only see 99 of the 100 tribespeople).
- One day, a blue-eyed foreigner visits the island and wins the complete trust of the tribe.
- On the evening of his departure from the island, he addresses the entire tribe to thank them for their hospitality.
- However, not knowing the customs perfectly, the foreigner makes the mistake of mentioning eye color in his address, remarking "how interesting it is to see another blue-eyed person like myself in this region of the world." Then he leaves the island forever on his waiting boat.

The puzzle question is:

What effect, if anything, does the foreigner's announcement have on the tribe?

Possible Solutions

Argument 1. The foreigner's address has no effect, because his comments do not tell the tribe anything that they do not already know (everyone in the tribe can already see that there are several blue-eyed people in their tribe).

Argument 2. Several days after the address, all the blue-eyed people line up to leave on the ferry, never to return.

The logic for Argument 2 is as follows, expressed in a fairly wordy way:

- If there were only 1 blue-eyed person (Alice), she would know (by personal observation) that there are 99 brown-eyed people on the island.
    - Which means that Alice would suddenly know that she must be the blue-eyed person mentioned by the foreigner, and so she has to leave the next day.
- If there were 2 blue-eyed people (Alice and Bob), each would know that there are 98 brown-eyed people on the island, and would also know that there is at least one other blue-eyed person (Alice can see Bob and vice-versa).
    - But Alice doesn't know if she has blue or brown eyes, and Bob doesn't know if he has blue or brown eyes. So for all they know, the tribe might be in the "1 blue-eyed person" condition.
    - But on the first day after the address, Bob doesn't leave, so Alice knows that cannot be true — otherwise Bob would have left.
    - So Alice knows now that she must have blue eyes as well. Reversing the logic, Bob now knows that he must be blue-eyed.
    - So they both leave on the next ferry (day 2 after the address), never to see their families or friends again. Everyone left says, "Whew, wasn't me!"[1]
- If there were 3 blue-eyed people (Alice, Bob, Carl) , the situation continues with the same logic.
    - Each sees 97 brown-eyed people and 2 blue-eyed people, so each of the three knows that there are either 2 or 3 blue-eyed people on the island.
    - No one leaves on day 2, so as the ferry pulls out, Alice, Bob, and Carl all deduce they are blue-eyed, and so leave on day 3.
- Et cetera.

My (Bob Cox's) Point of View

The flow of action in Argument 2 is fairly irrefutable. But it takes a long time to play out. The question that came to my mind is:

> What information did the address from the foreigner add to the information that people on the island already have? That is, what new piece of knowledge about the system did they get ***immediately*** (without having to wait for behavior to follow to make the further deductions that lead to mass migration)?

---

[1] If the islanders also know ahead of time that there can only be brown or blue eyes in their tribe, then on the day following the mass exit of the blueies, the brownies all have to leave — since they now know they must be in the other eye color case. For sale, one tropical island!

To address this question, I'll start with

The case of 1 blue-eyed person:

(1) This individual knows everyone else on the island has brown eyes, but did not know that there are any blue-eyed people on the island. Therefore, he/she gets a direct piece of information about the eye-color distribution that he didn't have before. Adios.

The case of 2 blue-eyed people:

(2) Every person on the island knows that there are blue-eyed people on the island, so there is no obvious direct information added. But, consider what each person's mental model of what other people's information is about the eye-color distribution *before* the foreigner's arrival. In particular, suppose Alice and Bob are the 2 blue-eyed freaks.

- Alice sees Bob's blue eyes, so she knows that there are either 1 or 2 blue-eyed people on the island. Bob, of course, is in the same state of information.
- Each brownie knows that there are either 2 or 3 blueies present.
- That is, each person on the island has a personal mental model of the eye-color distribution that has some uncertainty ("1 or 2?" for the blueies, "2 or 3" for the brownies). Of course, they never talk about this.
- But what is Alice's mental model of what Bob can know? (Again, before the foreigner's speech.)
    - Since Alice doesn't know the color of her own eyes, for all she knows Bob doesn't see any blue eyes at all, or he sees hers only (since all 98 other pairs of eyes are brown).
    - Since Bob doesn't know the color of his own eyes, before the outsider speaks, Alice's mental model of what Bob knows is "0 or 1 or 2" blueies on the island — unlike her own mental model, which is "1 or 2".
    - That is, Alice doesn't know beforehand that Bob knows there are blue eyes on the island, even though she knows that everyone else on the island knows about Bob's blue eyes.
    - When the outsider speaks, now she knows that Bob knows. This is the new information she gets (and Bob also, by symmetry) — the removal of some uncertainty about what someone else can know/deduce. She now knows that Bob's mental model does not include "0". Now she has to wait and see what Bob does about it.
        - Note that all the brownies in this 2 blue-eyed people scenario start in the position of one of the 3 blueies in the next case, as of course each of them don't know his/her own eye color and so they are in the situation of seeing 2 blue-eyed people on the island, 97 brown-eyed people, and wondering "am I number 3? Should I start packing?"

The case of 3 blue-eyed people (this is the confusing part when expressed in words):

(3) Alice, Bob, and Carl are blue-eyed, so each of them sees 2 other blueies. Before the speech, they each have their own mental model of "2 or 3" blueies on the island. All the brownies have the mental model of "3 or 4" blueies.

- Alice knows that Bob can see Carl's eyes, so Bob must know there is at least 1 blueie on the island.
- Since Alice doesn't know her own eye-color and Bob doesn't know his own (and Alice knows that), Alice's mental model of Bob's possible knowledge is “1 or 2 or 3” blueies.
- However, now consider Alice's mental model of what Bob's mental model of Carl's knowledge is.
    - As far as she knows, Bob sees only Carl's blue eyes (since Bob doesn't see his own and Alice's might be brown from *her* point of view).
    - In this counterfactual case (which she doesn't know is false), Bob might think that Carl doesn't see any blue eyes at all. Therefore, Alice's mental model of Bob's mental model of Carl's knowledge is “0 or 1 or 2 or 3” blueies (it stops at 3 because she knows Carl can see 97 pairs of brown eyes).
    - Thus, when the outsider speaks, Alice's mental model of Bob's mental model of Carl's level of knowledge changes — that is, Alice is less uncertain about what Bob thinks about what Carl knows — because Carl can no longer include “0” in his mental model. Alice can no longer think that Bob might think Carl doesn’t know about blue eyes on the island — after all, the foreigner (in whom they all trust) has just told them about it.
    - This reduction in uncertainty about what a second party can know about what a third party can know is the extra information that is conveyed/deduced instantly when the foreigner speaks — information being defined as data that reduces uncertainty.

If you understand the case **(3)** explanation above, then the larger cases are straightforward, if a little hard to say without feeling like you're babbling: what Alice thinks about what Bob thinks about what Carl thinks about what Doris knows about eye colors, etc. Then, as each day passes and the actions (or lack of actions) occur, then the uncertainty about what each person knows about what other people know unwinds, until suddenly the ferry gets filled when that knowledge hits home.

Therefore, the truly logical thing to do when an islander hears the pronouncement of the foreigner is to go out and drink himself/herself into a stupor. Perhaps that is why their religion also bans alcohol.