Anna — a dialogue about the problematic relation between consciousness and self-consciousness

Johan Gamper[1]

1 Subrosa KB

Abstract

In this dialogue, Anna and Alan talk about consciousness and its relation to self-consciousness. Could there be consciousness without self-consciousness? If not, what are the consequences? If it is possible, what are the consequences? How can we know of consciousness without self-consciousness?

Anna. Would you say that the imitation game defines human intelligence operationally?

Alan. Why not? Yes, it defines human intelligence operationally. What's your interest?

Anna. I'm curious about what's behind the operational definition.

Alan. The operational definition is just the tip of the iceberg. Almost everything is behind the operational definition. It's just a game. A set of rules *et voilà*. Come on now. What's the interest?!

Anna. Just a game? Okay. So the machine that passes the test just passes the test? No hidden agenda?

Alan. The game is a tool for navigation. It's a dark sea. Do you know what intelligence is?

Anna. What's the interest?

Alan. Okay. You're not interested. So you think that there must be something "behind", to make it interesting?

Anna. Stop it. Be kind for once. Why do you want to argue?

Alan. I think the soul is too deep to analyze. Let's be practical. Let's define what we know.

Anna. You're right. This is boring.

Alan. So surprise me. What's "behind"?

Anna. Consciousness.

Alan. You're right. An operational definition is parallel to consciousness. But if you cannot define consciousness, it's

meaningless? Right?

Anna. You think?

Alan. Alright. Stop it. Consciousness.

Anna. Okay. I'll help you. Can a machine be self-conscious?

Alan. Can anyone know?

Anna. Bridgman it!

Alan. Perhaps.

Anna. How would you know?

Alan. Operationally?

Anna. Go for it.

Alan. Okay. The machine is self-conscious if it passes the imitation game.

Anna. And consciousness? Can a machine be conscious? Without being self-conscious?

Alan. You couldn't ask it. But if it's not conscious about the questions the test is invalid? And if it's conscious about the questions it's self-conscious? So … you cannot define consciousness operationally with the imitation game. Only self-consciousness. … Okay. I'm getting it.

My test cannot capture consciousness. So how do the patterns match? Define self-consciousness with the imitation game and consciousness with some more basic measure. You cannot do it. However, you define consciousness it would be related to human organisms. That definition is useless for machines. The imitation game cannot discriminate between consciousness and self-consciousness. Does consciousness need some other foundation?

Anna.

It seems so. I work with self-consciousness since my clients cannot report what they don't know. But could they be conscious about things they cannot report?

Alan. Yes. That seems likely. Have you talked to Konrad? His ducks are like between your human subjects and my thinking machine. Do they see him? Who could tell?

Anna. If we take self-consciousness for granted it is easy to miss those who have conscious experiences without knowing it. Without having self-consciousness. When we say that the unconscious contents can be conscious we are in a context where the one being conscious or unconscious about his or her contents have the ability to be self-conscious.

Alan. My machines are in trouble. They may be conscious, but we wouldn't know if they don't tell us. And if they do, they

are self-conscious. Nothing wrong with that though.

Anna. Have I mentioned my cat? How can I tell if it sees me or not? Should we deny it consciousness because it lacks self-consciousness? Should we deny it consciousness if it lacks the ability to tell us it is conscious? I think it sees me.

Alan. I want a Victoria sandwich cake if I crack it.

Anna. I'm on.