

C. Winifred Jackson

University of Arizona Science of Consciousness Conference, September 2020

Poster Presentation (Category 1.0 Philosophy)

Public Link to Presentation (Virtual Presentation, Due to COVID):

<https://youtu.be/nk1zu7v5yaA>

TRANSCRIPT:

Hi, my name is Fred Jackson. I am an undergraduate Liberal Studies major currently focusing on Interdisciplinary Ethics and Philosophy of Mind at Oregon State University and today I am going to talk about robots, pie, and being a good person.

The relationship between consciousness and ethics is a complex and much larger topic than this presentation covers. In this presentation I will specifically discuss the first responsibility of consciousness, identifying who is affected by consciousness and who the interested parties are in matters of ethics by considering the following:

If we are conscious, we take action. If we take actions, we agree that there is meaning to those actions. If our actions have meaning we must be able to evaluate and prioritize those meanings; we must have ethics. Therefore, if we are conscious, we must have ethics. How do we then proceed to outline and evaluate those ethics? My question today is not whether or not we have experiences and consciousness, but rather if we are conscious beings what does that tell us about being a good person? Are there inherent ethical responsibilities in consciousness? If so, what are they?

First, let's talk about pie. Let's imagine that, someday in the not terribly distant future, you are walking down the street and you find the world's most adorable pie shop. The sign catches your eye, of course, as do the beautifully displayed pieces of edible craftsmanship in the window, but ultimately it's the smell that draws you closer. Apple pie, just like your Nana used to make. You'd recognize that smell anywhere and as you press your face up to the glass you remember how she used to make them for you after the first day of school every year. It's been far too long since you've had a slice. Your stomach rumbles and that settles it. You walk into the shop, pay them whatever it is they are asking, and get one perfect piece of fall themed heaven. Sitting down to eat it, you can almost taste it before your first fork full and when you DO manage to take that bite you realize this is the second best pie you've ever enjoyed. No one makes apple pie better than Nana.

We call this an experience. Pie is always an experience. You are perceiving the world around you through sight, smell, touch, taste, and sound. You are processing that sensory information, integrating it with your memories, determining a course of action, and having feelings...nostalgia, anticipation, satisfaction. Perhaps, as Nagel might suggest, this is what it is like to be a person.

Now let's talk about robots. Specifically, artificial general intelligence. At its most basic level, Artificial General Intelligence is a machine that can learn, understand, and perform any intellectual task a human might do. More than Alexa ordering your groceries for you, this is Star Trek's Data. While they don't exist today, someday in the not terribly distant future, there could be an AGI that can complete the same actions you do. They could visually identify the pie shop, utilize mechanized olfactory senses to determine a particular smell, decide upon a course of action based on this input, walk into the shop and purchase a slice of almost-as-good-as-Nana's apple pie.

Would this robot be having an experience? At current technology levels, we couldn't prove that. It's important to note, however, that at current technology levels we can't prove you had an experience either. The existence of subjective experience, or consciousness, is still a matter of debate and there are a lot of people working on figuring that out right now.

Today however we are talking about the ethical implications of consciousness, so let's go back to the pie shop. This time, let's put you and the robot next to each other. You both see the sign: World's Most Adorable Pie Shop. That information is processed through a series of electrical signals which help you both understand: pie lives here, and it's for sale. As you step closer, particles from the air enter your olfactory receptacles and a further series of electrical signals indicate from a database of previously concluded examples that this pie is likely to contain apples. This is the beginning of a decision tree, in which both you and the robot determine based on a series of factors whether or not to purchase the pastry containing apples. You both decide in the positive, of course, and proceed to purchase the second-best pie ever made. You place it into your food receptacles, and the mechanisms which enable you to process organic matter act with minimal input based once again on a series of electrical impulses.

"But wait!" you may say, "I have feelings about pie!" To which I say-of course you do. Its pie. However, no matter how many machines we hook you up to, we cannot see your feelings. The technology available today is amazing. We can watch you processing data. We can see that you have received sensory input. We can see that you are activating memory recall. We can see that your body is releasing endorphins. But there is no machine which will tell us about your Nana's apple pie. There is no advancement of modern technology which can point to any part of you and say, "this is the feeling of apple pie on the first day of school". We can only observe your physical processes and the results.

So now is the part where I point out that this is exactly as true for sufficiently advanced artificial general intelligence. You and the robot share what I call "the similar unknown". In the presence of observably similar actions taken by another being, we must assume the existence of that which we can experience but not prove within ourselves. Human beings already recognize the similar unknown-that's how we know that our pet cat is probably hungry and our pet rock is probably not. Dr. Fluffbottom will produce audible sounds designed to get attention and physically indicate their preference to be fed, just as a small non-verbal human would. We recognize the similar action as an indication of similar experience: hunger. Rocky, while decorative in her own

way, provides no such indicator.

We recognize the similar unknown primarily through the expression of preference. Cats prefer (often loudly and at 4am) to be fed. Infants prefer their mother's voices. You prefer apple pie, and I prefer pecan. This is how we know that you, and I, and sufficiently advanced Artificial General Intelligence, and baby humans, and Dr. Fluffenbottom share a similar unknown-the implication of subjective experience and perhaps consciousness-and Rocky does not.

It is my assertion that this process, the identification of other beings which share the similar unknown and therefore share the recognizable possibility of subjective experience and consciousness, is the first ethical responsibility of consciousness. According to Sartre, the ability to act is followed by actions themselves (existence precedes essence), and it is my argument that the existence of action necessarily creates the need for meaning. We exist. We eat pie. By eating pie, we agree there is a reason to eat pie-because we are hungry, because we miss our Nanas, because pie is delicious and makes us happy, supporting my assertion that essence precedes ethics. In order to determine what is ethical and what is not, whether we should or should not eat pie, we are required to first determine those affected by the action-who may experience pie and the consequences of the decision to eat it? Existence creates action which necessarily creates meaning and a need for ethics-ethics require us to attempt to identify the presence of objective experience.

Of course, the ethical responsibility to identify beings within whom exists the similar unknown is but the first step in what will likely be a comprehensive study. I am only beginning to explore the intersection of philosophy of mind and normative ethics. If we experience something, we act on that experience. If we're hungry, we eat. When we act, we agree there is a meaning to those actions. We eat apple pie because it is delicious, or comforting, or because we rely on the calories to continue physiological function, or because it is our biological impulse. If our actions have meaning, we must be able to evaluate that meaning in order to govern our own behavior. We must have ethics or else how do we decide to eat pie rather than lima beans or nothing at all? The answer lies in part in philosophy of mind-in how our biology and our mind interact to create experience and, perhaps, consciousness. The answer also lies in part in ethics-what behavior do we value and how do we act to reflect those values? At the crossroads where the two meet, I believe we can find answers that go beyond asking what it means to exist or identifying the functional behavior of moral goodness and move into a space where we can ask ourselves "what are the ethics of existence?".

The universe is a vast and ever-expanding place, filled with mysterious, immeasurable experiences. We cannot know what lies within the subjective experience of the person, genetic or artificial, who sits across from us in life's pie shop. We do know, however, that when we laugh or cry, sleep or dance, walk the city or eat apple pie we are expressing our innermost experiences, our preferences, our own personal and deeply complex universe, in which we feel. Without being able to measure it we know what it is to be happy or disheartened, tired or energetic, aimless or hungry and our behaviors share the evidence of those experiences with the world. It is therefore our responsibility to recognize the similar unknown within others, whether

we call that experience or consciousness or something else, and imagine them as complexly as we know ourselves to be-to assume that the same capacity for joy and sadness and need and fulfillment exist in those creatures which show preference, and to include all conscious beings in our quest to fill the universe with ethical behavior and good desserts.

Acknowledgments

Thank you so much for allowing me the opportunity to present. None of this would have been possible without the dedication of Mike Freeborn, who believed in me from the time I was a young human and continues to create the future by inspiring young minds. I must also acknowledge that I will never finish thanking Matthew Stockton, Faculty Department Chair of Philosophy at Portland Community College, in whose classes I discovered a passion for philosophy and realized-to my everlasting benefit-what a terrible attorney I would make. My sincerest respect and gratitude goes to Dr. Stephen Jolin, Ph.D., Philosophy Department, Portland Community College, who was there to inspire me as I took my first steps toward the exploration of the point at which ethics, consciousness, and artificial general intelligence collide. It of course would have been impossible to finish this project without the professional wisdom of **Dr. Ryan Hickerson, PhD.** Professor of Philosophy at Western Oregon University and the wonderful encouragement and engaging conversation of Dr. Linda Marie Richards, Professor of History at Oregon State University both of whom challenged me and inspired me toward my best thinking. I cannot finish this project without thanking my beloved friends, family, and informal reviewers-especially Burton Jackson, Matthew Fox, Joel Smith, Matthew Stout, Janet Littrell, and Vladimir Gerencir for their endless patience, enthusiastic debates, and undying faith. Finally, and always most important, this work is dedicated to my children who have allowed my experiments in applied ethics for many years and endured every philosophical lecture I've ever given with grace and the bare minimum amount of eye rolling. The three of you are the entire reason being a good person has any real meaning to me at all.

If you would like to contact me about my presentation, my ongoing work on the ethical implications of consciousness, or anything else I can be reached via email at c.winifredjackson@gmail.com. Thank you and stay safe.