Reliabilism

In this paper I examine the reliabilist theory of justification. In particular, I examine the ways in which reliability is determined, and the merits of each way of determining reliability. I suggest that we should use different criteria for determining reliability, depending on the situation.

Briefly, reliabilist theories of justification hold that a belief's justification is a function of the reliability of the processes used to form that belief, where reliability refers to the tendency of a process to produce true beliefs in a person. Memory and perceptual processes which produce beliefs are counted as reliable, while hunches and wishful thinking are not counted as reliable belief-forming processes. Beliefs formed by perceptual and memory processes are said to be justified, whereas beliefs formed by hunches and wishful thinking are said not to be justified. Justification admits of degrees: beliefs can be more or less justified, depending on how reliable their belief-forming processes are. This is, of course, a simplified version of reliabilism, but it captures important elements of the theory and is suitable for our present purposes.

There are two readily apparent ways of determining reliability: either by looking at the tendency of processes to produce true beliefs in the population as a whole, or by looking at the tendency of processes to produce true beliefs in a given individual. I will call the former kind of reliability "general reliability", and the latter "individual reliability." For memory to count as a generally reliable process, memory must produce true beliefs in most of the people most of the time, and anyone who arrives at a belief that-p through a generally reliable memory

process has a justified belief that-p. In order for an individual to count memory as a reliable process, memory must produce true beliefs in that individual most of the time. This kind of reliability allows only the person for whom the process is reliable to count beliefs formed by that process as justified. I would also imagine that there are other ways of determining reliability, one of which is by looking at the tendency of a process to produce true beliefs in a certain group of people, and I will return to this idea later.

It seems that, depending on the case, it is desirable to favor one type of reliability over the others. Here are two cases where it makes sense to favor individual reliability over general reliability. Consider James, a man who has frequent hallucinations, and is aware that he has frequent hallucinations. Because of this, he comes to believe many things which are not true. For example, he believes that he was at school yesterday, when in fact he was never at school; rather, he stayed home all day. He uses a memory process today to form his belief today that he was at school yesterday. To use Alvin Goldman's terminology[1], James' input belief at t_1 is that he was at school yesterday, and his output belief at t_2 (which he arrives at through a memory process) is that he was at school yesterday.

Goldman stipulates that the process used to form the belief at t_1 "that I was at school yesterday" must be reliable. I am not quite sure which process James used to form his initial belief "that I was at school yesterday". Probably most of us use perceptual processes to form beliefs like "I am at school right now". We later input these beliefs into memory processes which then produce output

[1] Goldman, Alvin. *What is Justified Belief?* 347. Hereafter, cited, parenthetically, by page number.

beliefs like “I was at school yesterday”. So, James yesterday had perceptions as of being at school and formed the belief “I am at school right now.” His belief at t_1 was formed by perceptual processes, which Goldman tells us are generally reliable. Today, James input his belief from yesterday “I was at school” and through a memory process came to believe “I was at school yesterday”, and this is a justified belief. However, James does not form a true belief at t_2 about what he did yesterday, even though the process he used to form that belief (a memory process) is counted as generally reliable. Moreover, many of his beliefs which are formed by memory processes will not be true, because of his frequent hallucinations, which cause his input beliefs to be false. But, if the input beliefs are justified, and he uses a memory process to produce new beliefs, these new beliefs will count as justified, according to general reliability. But why should this be so? It is clear that memory processes are not reliable for James[2], at least not when the input beliefs are beliefs formed by perceptual processes. James’ beliefs which result from memory processes should not be counted as justified.

Conversely, consider the case of a person whose hunches are very accurate. Maybe this person is like Hercule Poirot, and his hunches produce true beliefs most of the time. However, if we appeal to general reliability, hunches are not counted among reliable belief-forming processes, so our Poirot’s beliefs which result from his hunches will not be justified. But this seems wrong. His hunches produce true beliefs most of the time, so why should they not be counted as reliable for him, just because hunches in general are not reliable?

[2] It is certainly true that James’ memory processes are conditionally reliable. However, it seems that, in James’ case, conditional reliability of his memory processes is not enough to confer justification on beliefs formed by those memory processes, because of the large number of false input beliefs.

There are a number of reasons to favor individual reliability over general reliability for these cases. It allows us to say, in the hallucination case, that beliefs formed by memory processes are not justified. Indeed, it seems that, for James, memory processes produce false beliefs much of the time, so beliefs resulting from memory processes should not be justified for James. We can also say that our Poirot has justified beliefs that result from his hunches. And, intuitively, this seems right, because his hunches tend to produce true beliefs, so why should we not count those beliefs as justified? With individual reliability, we are also able to easily deal with the problem of the case of wishful thinking in world *W*, which Goldman discusses. (349) For individuals in *W* wishful thinking is a reliable process, so beliefs formed by wishful thinking are justified for those in *W*. We need not say that wishful thinking is always and everywhere reliable, only that it is reliable for individuals in *W* (and anyone else for whom wishful thinking produces true beliefs). Another benefit of this approach is that reliabilism will "get it right" more often. If processes are determined to be reliable on a person-by-person basis, then any given reliable process (in relation to the appropriate person) will produce more justified true beliefs than if processes are determined to be reliable by testing their tendency to produce true beliefs in the population as a whole.

Let us now consider a case similar to James' where it seems that we should use general reliability rather than individual reliability. This is the case which Feldman considers in *Epistemology* (94), of Brian and a certain Brain in a vat: "Brain is a normal person with accurate and well-justified beliefs about the world

around him. Brain is Brian's mental duplicate. Brain has experiences just like Brian's. And Brain's beliefs are analogous of Brian's." However, most of Brain's beliefs are false, because he is not actually experiencing any of the things which he believes his is experiencing. Feldman suggests that reliabilism runs into trouble here, because it implies that Brain's beliefs are unjustified, when in fact they are justified. And, if we use individual reliability to determine Brain's justification, the result *will* be that Brain's beliefs are unjustified. However, if we use general reliability to determine the justification of Brain's beliefs, then his beliefs will be justified. His beliefs result from perceptual and memory processes, just like Brian's beliefs, and perceptual and memory processes are both counted as generally reliable, so Brain's beliefs are justified.

What is the difference, then, between Brain's situation and James' situation? Why should we use general reliability for Brain and individual reliability for James? One reason is that James is aware that he has frequent hallucinations, and aware that this affects the beliefs he forms. It would be very strange for James (or anyone else) to think that most of his memory beliefs are justified, and for him to say, "Since memory processes produce justified beliefs in most people, my beliefs which result from memory processes are justified," when he clearly knows that his beliefs which result from memory processes are not justified. While we cannot say that believing processes to be reliable makes them reliable, we can perhaps say that when a person knows or suspects that one of his processes which is held to be generally reliable tends to produce false beliefs, then he should use individual reliability to gauge the justification of his beliefs that result

from that process. Since Brain is not aware of his situation, he has no reason to doubt the reliability of his processes, and so we use general reliability to determine the justification of his beliefs.

Let us turn now to the case where reliability of a process is determined by looking at the tendency of the process to produce true beliefs not in the population at large, but in a certain group of people. Imagine there is a family whose members have hunches which are almost always correct. Hunches tend to produce true beliefs in members of this family. The mother can always tell, for example, which child made a mess in the kitchen. She has a hunch that it was Susie who made the mess and then forms the belief that Susie made the mess.[3] These beliefs are always or almost always true for the mother, and for the other members of the family. It seems reasonable to say, then, that hunches are reliable processes for this family. Now, this family happens to have a son whose hunches are usually wrong. However, the son does not know this, and thinks that he is like the rest of his family. He forms beliefs which result from his hunches, but the beliefs usually turn out to be false. It seems reasonable to say that the son's beliefs which result from his hunches are still justified, even though they are mostly false, because they are justified for the rest of his family, and he has no reason to doubt that he is like the other members of his family, with regard to hunches. The son is like Brain; their beliefs are justified because the processes used to form those beliefs are reliable for a larger population. In the

[3] This is only an example, of course, and if you think that all mothers can tell which of their children is responsible for making a mess, then pick another example.

son's case, the process is reliable for his family, and in Brain's case the process is reliable for the population at large.

This gives us another answer to the problem of wishful thinking in world *W*. Wishful thinking is a reliable process for individuals in *W*, because it is reliable for the group of people that lives on *W*. Even if, for a particular individual on *W*, wishful thinking did not produce true beliefs, that individual could still count as justified beliefs which resulted from wishful thinking, because wishful thinking is a reliable process for the group of people which that individual is a part of. This individual is like Brain and the son. We need not say that wishful thinking is reliable in this world, or that because it is not reliable in this world it is not reliable in *W*, because this clearly seems false. Wishful thinking *is* reliable in *W*.

It seems, then, that there are at least three different ways of determining reliability and, depending on the situation, it is preferable to use a one kind of reliability rather than another. It may be objected that this is a problem, that a reliabilist theory of justification cannot use more than one kind of reliability when determining justification. I do not see why this must be the case. Why should we limit ourselves to one kind of reliability, when this sort of limiting will "get it right" less often than if we use different kinds of reliability? Probably more extensive criteria then I have suggested must be given for determining when to use which types of reliability, but I am sure that such criteria can be given.