Is there an argument for transparency of mind?

Bas van Fraassen

> What I mean is: is there a non-trivial sense in which, if something is in the scope of what we can believe then *we are clear on* whether we believe that or not, and if we have gradated belief (subjective probability) then *we are clear on* how likely it seems to us.

The Moore paradox.

1. 'A and I do not believe that A'

This could be true, and is in fact true for many sentences A. So what is paradoxical about it? It seems to be something that, although possibly true, will express what the speaker believes about himself only if there is something wrong with his state of belief. So consider:

2. 'I believe that A and that I do not believe that A'

Given the first part, I had better start believing that I do believe that A ... So if 1. expresses my state of opinion, it expresses a defective state -- with a defect that is 'internal': to see it we need not know how this opinion relates to facts that need to be checked independently. So it is a case of *incoherence*.

The ***Principle*** that appears to be operative in the preceding argument is this: if B expresses a state of opinion which is coherent, then 'I (fully) believe that B' also expresses (ceteris paribus) a coherent state of opinion. [1]

"Implication" (of a sort: consequence on pain of incoherence)

The conclusion has a consequence that seems strange. In ordinary logic, we note that if (P & ~ Q) is inconsistent, then P implies Q. So from the incoherence of 1., can we conclude something like:

A *implies(?)* I believe that A
I believe that A *implies(?)* A

in any sense of "implies"? As pointed out at the beginning, 1. could be true. So the "implies" cannot mean "if the premise is a true statement so is the conclusion". It is

[1] Is this a good principle? I think so. Consider the alternative: there is a coherent opinion about certain subjects such that, if you believe that you have that opinion, then your opinions are incoherent. That is very paradoxical -- since that would seem to be a true opinion about yourself! But do not try to justify the principle simply on that basis: do not rely on an intuition to the effect that if something is true about you, then your belief that it is so must be coherent -- the Moore paradox refutes ***that!*** However, here we are looking at a case in which the entire content is about one's own subjective state.

rather that the premise brings the conclusion along with it if it enters a content of coherent belief. To put it another way:

> the premise brings the conclusion with it in the sense that given the premise, one cannot believe anything contrary to the conclusion *on pain of incoherence*

so we should call this "*consequence on pain of incoherence*".

Probabilistic Moore

Suppose now that P(A) = 3/4 expresses some of my state of opinion. Does it follow that $P(p_0(A) = 3/4) = 1$?

Trying to duplicate the above argument for this case we consider:

1*. $P(A) = 3/4$ and $P(p_0(A) = 3/4) = x$ $[x < 1]$

Could these two expressions of opinion jointly express my state of opinion, if it is coherent? In analogy with the above we look at:

2*. $P(p_0(A) = 3/4$ and $p_0(p_0(A) = 3/4) = x) = 1$

The content is a factual statement on which we should comment: if the first conjunct is true, then the second conjunct says that you have a mistaken opinion about your own state of opinion. So this is a defect internally detectable. So if 2* is correct then you believe yourself to have an incoherent state of opinion --

> this is a self-fulfilling opinion, for if what it is about is coherent, then it clashes with that, so that the state as a whole is incoherent.

So we conclude that 2*. expresses an incoherent opinion. From that we conclude that the opinion expressed by 1* is incoherent.

The ***Principle*** that appears to be operative in the preceding argument is the same as before:

> if B expresses a state of opinion which is coherent, then 'I (fully) believe that B' also expresses (ceteris paribus) a coherent state of opinion

provided in this context full belief is [or implies] subjective probability 1.

Now if all the above is accepted, then it follows that:

$P(A) = 3/4$ *implies* $P(p_0(A) = 3/4) = 1$ *on pain of incoherence*