Abstract
In this paper I show that one version of motivational internalism follows from the so-called ideal observer theory. Let us call the version of the ideal observer theory used in this essay (IOT). According to (IOT), it is necessarily the case that it ought to be that $A$ if and only if every ideal observer wants it to be the case that $A$. We shall call the version of motivational internalism that follows from (IOT) (moral) conditional belief motivational internalism (CBMI). According to (CBMI), it is necessarily the case that, for every $x$: if $x$ were an ideal observer, it would be the case that $x$ believes that it ought to be that $A$ only if $x$ wants it to be the case that $A$. Given that it is necessarily true that no ideal observer has any false beliefs, (IOT) entails (CBMI). Or, so I shall argue.

Keywords: the ideal observer theory, motivational internalism, practical reason

1 Introduction
In this paper I show that one version of motivational internalism follows from the so-called ideal observer theory. The classical modern formulation of the ideal observer theory is [16]. According to Roderick Firth,

If it is possible to formulate a satisfactory absolutist and dispositional analysis of ethical statements, it must be possible [...] to express the meaning of statements of the form “x is right” in terms of other statements which have the form: “Any ideal observer would react to x in such and such a way under such and such conditions.” [16, 329]

The main idea seems to be that an act is right if and only if (iff) it would be approved by any ideal observer, and more generally that an act has
a moral property iff any ideal observer would react in such and such a
way if such and such conditions were realized.

According to Firth such an analysis is absolutist, dispositional, objectivist, relational, and empirical. An ideal observer is omniscient with respect to all non-ethical facts, omnipercipient, disinterested, dispassionate, consistent, and in other respects normal. The ideal observer theory used in this essay is inspired by, but not identical to, Firth’s. Our version, which we will call (IOT), entails two important theses (O) and (B).

(O) Necessarily, it ought to be that $A$ if and only if every ideal observer wants it to be the case that $A$.

(B) Necessarily, for every $x$: if $x$ is an ideal observer, then if $x$ believes that $A$ then $A$.

Firth thinks that we should assume that an ideal observer knows all (non-ethical) facts whatsoever. (B) is a slightly different, and in one sense weaker, condition. According to (B) it is necessarily true that no ideal observer has any false beliefs. But we don’t have to assume that an ideal observer knows everything. (B) is a reasonable condition since desires based on false beliefs might be irrational. Furthermore, we don’t have to restrict (B) to beliefs about non-normative facts, since normative facts, according to the ideal observer theory, just are (constituted by) facts about what ideal observers want. And it seems quite reasonable to assume that an ideal observer’s beliefs about what ideal observers want are true. We could, in fact, restrict (B) as Firth does. But then our argument would become slightly more complex. So, I will not labour the details here.

We have stipulated that (B) follows from (IOT). It is worth noting that this is not a definition that is peculiar only to our version of the ideal observer theory. Here is a quote from Michael Smith, for instance:

My suggestion is that to be fully rational an agent must not be suffering from the effects of any physical or emotional disturbance, she must have no false beliefs, she must have all relevant true beliefs, and she must have a systematically justifiable set of desires, that is, a set of desires that is maximally coherent and unified. [48, 263] [My italics.]

Furthermore, if it is necessary that (i) every ideal observer believes what she knows, and (ii) it is necessary that no ideal observer believe that $A$ and believe that not-$A$, then (B) follows from the proposition that it
is necessary that every ideal observer knows everything. So, (B) is also entailed by every version of the ideal observer theory that assumes that every ideal observer is omniscient and satisfies conditions (i) and (ii), which seem very reasonable.

The following considerations show why it is plausible to assume that (IOT) entails (B). Suppose that this is not the case. Then ideal observers can have false beliefs. And if some ideal observers have false beliefs, it is very unlikely that all will want and accept the same things. This entails that it is difficult to formulate a reasonable deontic logic (see section 2), and may lead to implausible relativistic conclusions. Furthermore, if ideal observers can have false beliefs, we can derive intuitively implausible normative results from (IOT). Suppose there is just one ideal observer and this ideal observer wrongly believes that a new vaccine has no bad side effects. Furthermore, suppose that this is supported by all evidence. Therefore, the ideal observer wants the doctors to administer this vaccine to all children in the world. According to the current version of the ideal observer theory, it follows that the doctors should administer the vaccine. However, the ideal observer’s belief is false; the vaccine has severe side effects and administering it to all children in the world would have catastrophic consequences. So, according to our intuitions, the doctors should not administer the vaccine. Accordingly, if the ideal observer theory does not entail (B), or some very similar principle, it has intuitively incorrect normative consequences in this case and doesn’t seem to be a reasonable theory.

For our purposes in this essay, it doesn’t matter if (IOT) also entails other propositions. We don’t have to specify the necessary and sufficient conditions for the theory in order to establish our results.³

Let us now turn to a definition of “internalism”. There are many forms of this philosophical thesis: existence, judgement and belief internalism, conditional and unconditional internalism, reasons and motivational internalism etc.⁴

One of the most important distinctions between different kinds of internalism is the distinction between unconditional (unrestricted, indefeasible) and conditional (restricted, defeasible) versions. Here is an example of an unconditional belief motivational internalism.

(UMI) Necessarily, for every \( x \): if \( x \) believes that it ought to be the case that \( A \), then \( x \) wants it to be the case that \( A \).

Depending on how we read the “necessity-operator” in (UMI) we obtain several forms of (UMI) with different strength. Internalism has traditionally been interpreted as some kind of conceptual or analytic truth,
but we might also formulate versions of (UMI) where the necessity is some sort of metaphysical, nomological or historical necessity. Finally, it is also possible to drop the “necessity-operator” entirely.

If we replace “believes” with “judges” in (UMI), we obtain a similar kind of unconditional judgment motivational internalism. By replacing “wants” or “wants it to be the case that” in (UMI) with some other expression, e.g. “is motivated to see to it that” or “has a reason to see to it that”, it is possible to formulate many other versions of internalism.

As [7, 126] points out, it is regularly assumed in contemporary metaethics that theories of this kind are too strong, since “it seems possible to conceive of someone who makes a moral judgement but fails to be motivated accordingly because she suffers from, e.g., apathy, depression, exhaustion, or emotional disturbance”. The amoralist is an individual who acknowledges moral obligations, but remains utterly unmoved by them. And the amoralist seems to be a conceivable person. Unconditional forms of internalism have a hard time explaining the phenomenon of weakness of will.

To avoid problems of this kind, internalists have often defended some version of conditional internalism of the following form instead:

\[(CMI) \text{Necessarily, for every } x: \text{ if } x \text{ has the property } P, \text{ then if } x \text{ believes that it ought to be the case that } A, \text{ then } x \text{ wants it to be the case that } A.\]

Again, there are many different interpretations of (CMI). “Necessarily” can stand for conceptual, metaphysical, nomological, historical necessity etc. We have not used a counterfactual in (CMI), but there are also counterfactual versions of this schema. According to [7], there are three broad kinds of specification of the property \(P\) in theories of this sort: the property \(P\) can stand for the property of being psychologically normal, the property of being practically rational or the property of being morally perceptive.

The kind of internalism that is entailed by (IOT) is a (moral) conditional belief motivational internalism, abbreviated (CBMI). This version is most similar to the second kind of theories mentioned by [7], where “the property \(P\)” stands for “the property of being practically rational”. However, since it is not obvious that it is not possible to be completely rational and still have false beliefs, the current version is perhaps best classified as a separate kind. Here is our official definition of (CBMI):

\[(CBMI) \text{Necessarily, for every } x: \text{ if } x \text{ were an ideal observer, it would be the case that } x \text{ believes that it ought to be} \]
that $A$ only if $x$ wants it to be the case that $A$.\footnote{10} (CBMI) connects beliefs about norms with motivation in ideal observers. So, (CBMI) entails that if you believe that you ought to do something then you also want to do it, given that you are an ideal observer. Consequently, according to (CBMI) it is possible that you believe that you ought to do something without being motivated to do it, but only if you are not an ideal observer. (CBMI) seems to be a plausible form of internalism. It doesn’t exclude weakness of the will or the possibility of amoralists. So, it avoids the problems with unconditional versions of internalism. Yet there is a necessary connection between moral beliefs and wants according to the theory. (CBMI) is, in a sense, a quite weak kind of internalism, since the concept of an ideal observer is the concept of a highly idealized individual. But I don’t think it is so weak as to become uninteresting.

In [8], Richard Brandt mentions seven reasons that make the ideal observer theory attractive.

(1) [… The] theory enables us to regard as really relevant to ethics all the facts which on reflection we take to be relevant; (2) […] it enables us to explain the heterogeneous- ness of the actions which we regard as right or wrong; (3) […] it explains how ethical disagreement is possible even when there is agreement about the nature of the act being appraised; (4) […] it explains why our feelings and attitudes—and especially our sympathies—are (and properly are) engaged in ethical reflection, and why moral philosophers have thought that moral experience is distinctively a union of cognition and emotion; (5) […] it enables us to hold that moral opinions are subject to objective criticism and are correct or incorrect; […] (6) it explains why we value the advice of knowledgeable, impartial, and consistent persons at times of moral decision, and why we reject previous moral opinions of our own which we think reflect self-interest, inconsistency, or lack of information; and […] (7) it enjoys advantages over the emotive theory such as the capacity to give a satisfactory analysis of “ethical relevance”, and the ability to explain why ethical judgments do not always correspond with favorable or unfavorable attitudes on the part of the judge. [8, 407]

Such considerations also support (IOT). I will not discuss any arguments for the ideal observer theory or internalism in this paper.\footnote{11} But both
of these theses are interesting and have been accepted, in one form or another, by many philosophers. So, it is well worth spelling out the details of the argument from (IOT) to (CBMI). Even though it is plausible to assume that most ideal observer theorists have been internalists, few people have explicitly investigated the connections between these two theories. This is one reason why the argument in this paper is interesting. One notable exception is Smith. Smith has argued for motivational internalism in several places (e.g. [47] and [48]). His argument is, however, quite different from the one introduced in this essay. His versions of the ideal observer theory and motivational internalism are also different from the versions used in this paper. So, it is plausible to investigate the alternative argument presented in this essay and to continue to explore the connections between the ideal observer theory and motivational internalism.

In section 2 I will briefly discuss two objections against (IOT), since it might seem to be the case that this version of the ideal observer theory can be easily dismissed on some purely formal grounds. Section 3 includes our main argument from (IOT) to (CBMI) and section 4 contains a short conclusion.

2 Two arguments against (IOT)

In [22] Jonathan Harrison mentions some potential problems with Firth’s ideal observer theory. I will consider two of his “technical” arguments that might seem to be particularly troubling for (IOT). Here is the first argument.

Unless there is a God, and He is an ideal observer in Professor Firth’s sense, […] it is quite certain that nothing answers to the description of an ideal observer which Professor Firth has given. Whether something answers to this description or not, Professor Firth quite rightly wants to make the truth of ethical judgments independent of the actual existence of an ideal observer. Hence Professor Firth’s analysis of moral judgments is attended with the difficulties involved by universal propositions about non-existent classes. […] Of universal propositions whose truth is meant to be independent of the nullity of the class of entities they are about I do not think any satisfactory […] analysis has been given. ‘(x) x is an ideal observer materially implies x approves of A’ will not do. For since not-p implies p materially implies
q, all such propositions about null classes will be true. If ‘x is an ideal observer’ is false for all values of ‘x,’ then ‘(x) x is an ideal observer materially implies x will approve of A’ will be true, simply for that reason, and so will ‘(x) x is an ideal observer materially implies x will disapprove of A.’ In this case if there are no ideal observers, any ethical statement we care to make will be true, for all such statements will involve universal statements about a null class, and on this analysis of such statements, all of them are true. [22, 256-257]

This kind of reasoning is also a potential problem for (IOT). If the quantifier expression “every” in (O) is interpreted in the standard way and there are no ideal observers, then everything is obligatory. This is clearly an absurd conclusion. However, if this expression is interpreted as a “possibilist” quantifier whose range consists of all possible individuals, this is no longer true. According to this interpretation, it is not the case that the proposition that all ideal observers want it to be the case that A follows from the proposition that no ideal observers exist. So, this argument seems to be conclusive only if we can rule out a possibilist interpretation of our quantifiers, and this is not a simple task. Of course, the idea that we can quantify over non-existent things (if there are any), is controversial. However, I think there are some interesting arguments for this view, and I haven’t seen any plausible arguments that refute it. Furthermore, one can now find quite sophisticated versions of this theory in the philosophical literature.

Another possible response for a defender of the ideal observer theory is to insist that there are ideal observers, but that these are purely hypothetical, non-concrete individuals, like planes without friction or ideal gases seem to be. Even a so-called actualist can avoid Harrison’s first argument in the way we have indicated, if it is reasonable to postulate such abstract or non-concrete ideal observers.

The following quote summarizes the second argument:

‘A is right’ and ‘A is wrong’ are contradictories. All actions are either right or wrong, and no action can be both. . . . Yet, on Professor Firth’s definition of ‘right,’ right and wrong are not contradictories but . . . contraries. If ‘A is right’ means ‘All ideal observers approve of A,’ ‘A is wrong’ presumably means [ . . . ] ‘All ideal observers fail to approve of A,’ and though these two statements cannot both be true, they can both be false, and would both be false if some ideal observers approved of A, while others did not. [23, 261]
A similar argument can be directed against (IOT). If it is permitted that $A$ iff every ideal observer accepts $A$, and it is wrong that $A$ just in case every ideal observer fails to accept $A$, then “it is permitted that $A$” and “it is wrong that $A$” are contraries rather than contradictories. However, there are other more plausible definitions of these concepts. Consider the following definitions. (F) Necessarily, it is forbidden that $A$ iff every ideal observer wants it to be the case that not $A$. (P) Necessarily, it is permitted that $A$ iff some ideal observer accepts that it is the case that $A$. Assume that it is necessary that for every ideal observer $x$: $x$ wants it to be the case that $A$ iff it is not the case that $x$ accepts that it is not the case that $A$. Furthermore, suppose that it is necessary that for every ideal observer $x$, if $A$ is necessarily equivalent with $B$, than $x$ wants (accepts) that $A$ iff $x$ wants (accepts) that $B$. Then it follows that it is necessary that it is forbidden that $A$ iff it is not permitted that $A$.

A potential problem with this response is that the definition (P) might seem implausible. The following equivalence is perhaps more reasonable. (P') Necessarily, it is permitted that $A$ iff every ideal observer accepts that it is the case that $A$. If we assume that it is necessary that all ideal observers want (and hence also accept) the same things and that there is an (existing or only possible) ideal observer, it follows that it is necessary that every ideal observer wants (accepts) that $A$ iff some ideal observer wants (accepts) that $A$. Given these assumptions (P) and (P') become equivalent and all the usual relationships between our normative concepts are forthcoming.

Of course, this response assumes that all ideal observers want the same things. And this assumption is controversial. But I believe that Firth would be happy with this response. He thinks that if $x$ is right, and if there were more than one ideal observer, all the ideal observers would have the same ethically significant experience with respect to $x$. And this position doesn’t seem implausible to me, even though I cannot defend it here. In [17] we read the following:

[T]he analysis [of statements of the form “$x$ is right”] is not intended to imply either that there exist any ideal observers or that there do not. But it does imply that if $x$ is right, and if there were more than one ideal observer, all the ideal observers would have the same ethically significant experience with respect to $x$. This fact could be brought out by formulating the analysis to read: “If there were any ideal observers, they would all have such and such an experience with respect to $x$. [8, 414]
So, even though I think Harrison’s technical arguments against Firth’s theory are quite clever, they don’t appear conclusive. Consequently, there seem to be no quick way of dismissing (IOT) as an unreasonable form of the ideal observer theory.

3 The main argument: the (IOT) argument for (motivational) internalism

We are now in a position to prove that the ideal observer theory entails motivational internalism, or more precisely that (CBMI) follows from (IOT). To prove this we assume that (IOT) is true and that (CBMI) is false and derive a contradiction. We will call this argument “the (IOT) argument for (motivational) internalism” or simply “the (IOT) argument”.

1. (IOT) is true in our possible world @. [Assumption]
2. (CBMI) is false in @. [Assumption]
3. (O) is true in @. [From 1]
4. (B) is true in @. [From 1]
5. “Necessarily, for every $x$: if $x$ is an ideal observer, then if $x$ believes that it ought to be that $A$ then it ought to be that $A$” is true in @. [From 4]
6. “For every $x$: if $x$ were an ideal observer, it would be the case that $x$ believes that it ought to be that $A$ only if $x$ wants it to be the case that $A$” is false in some possible world $w_1$. [From 2]
7. “If $c$ were an ideal observer, it would be the case that $c$ believes that it ought to be that $A$ only if $c$ wants it to be the case that $A$” is false in $w_1$. [From 6]
8. There is a possible world $w_2$ that is as close to $w_1$ as possible but in which “$c$ is an ideal observer” is true and “If $c$ believes that it ought to be that $A$, then $c$ wants it to be the case that $A$” is false. [From 7]
9. “$c$ believes that it ought to be that $A$” is true in $w_2$. [From 8]
10. “$c$ wants it to be that $A$” is false in $w_2$. [From 8]
11. “For every $x$: if $x$ is an ideal observer, then if $x$ believes that it ought to be that $A$ then it ought to be that $A$” is true in $w_2$. [From 5]

12. “If $c$ is an ideal observer, then if $c$ believes that it ought to be that $A$ then it ought to be that $A$” is true in $w_2$. [From 11]

13. “If $c$ believes that it ought to be that $A$, then it ought to be that $A$” is true in $w_2$. [From 8, 12]

14. “It ought to be that $A$” is true in $w_2$. [From 9, 13]

15. “It ought to be that $A$ if and only if every ideal observer wants it to be the case that $A$” is true in $w_2$. [From 3]

16. “Every ideal observer wants it to be the case that $A$” is true in $w_2$. [From 14, 15]

17. “If $c$ is an ideal observer, then $c$ wants it to be the case that $A$” is true in $w_2$. [From 16]

18. “$c$ wants it to be that $A$” is true in $w_2$. [From 8, 17]

But 18 contradicts 10. Hence the assumptions cannot be true. The argument is valid and only uses ordinary semantic assumptions made in propositional, predicate, modal and counterfactual logic. It follows that (IOT) entails (CBMI).

4 Conclusion

In this paper I have tried to prove that one version of motivational internalism follows from the so-called ideal observer theory. If it is necessarily true that it ought to be that $A$ if and only if every ideal observer wants it to be the case that $A$, and it is necessarily true that for every $x$: if $x$ is an ideal observer, then if $x$ believes that $A$ then $A$, then it is necessarily true that for every $x$: if $x$ were an ideal observer, it would be the case that $x$ believes that it ought to be that $A$ only if $x$ wants it to be the case that $A$. So, if we have good reason to believe that the ideal observer theory is true, we have good reason to believe that motivational internalism is true. This is an interesting result, since both the ideal observer theory and motivational internalism are attractive theories that have been defended by many philosophers.
Conversely, if we have good independent reasons to believe that motivational internalism is true, the argument in this paper can be interpreted as an abductive argument for the ideal observer theory. For this theory entails motivational internalism. Then motivational internalism provides evidence for the ideal observer theory. Such an argument is, of course, highly fallible, and if we use the argument in this way the merits of the ideal observer theory should be compared to the merits of other metaethical theories that entail some plausible version of internalism. Nevertheless, this way of interpreting the argument seems quite interesting to me and worth further discussion.

Obviously, it also follows from our results that if motivational internalism isn’t true, the ideal observer theory isn’t true either. And if it is reasonable to believe that motivational internalism is false, it is reasonable to believe that the ideal observer theory is false. As an anonymous referee pointed out, it would be interesting to know whether there is any ideal observer theorist that isn’t an internalist, since a position of this kind seems to be incompatible with the results in this paper. However, since very few persons that defend the ideal observer theory have said anything about the connections between this theory and motivational internalism, and I am not aware of anyone that holds such a position, I am afraid we can only speculate about this.

Are there any other immediate metaethical implications of our results in this paper? Traditionally, internalism has often been used in various arguments against different versions of cognitivism. It is pretty clear, however, that the internalism discussed in this paper doesn’t have any implications of this sort. The ideal observer theory is almost always interpreted as a cognitivist theory, and (IOT) is certainly intended to be a theory of this kind. Since (IOT) entails (CBMI), it follows that (CBMI) is compatible with (IOT), given that (IOT) itself is consistent, as it seems to be. So, we cannot conclude that internalism entails non-cognitivism. On the contrary, the results prove that there is at least one interesting form of motivational internalism that is compatible with at least one interesting form of cognitivism. In light of the historically close allegiance between internalism and non-cognitivism, this is an interesting result. Whether there are any other significant implications of the results in this paper is a topic for another day!
Acknowledgements

I would like to thank all anonymous referees for interesting comments on earlier versions of this paper.

Notes

1 Other philosophers that have been influenced by or defended some kind of ideal observer theory include Richard Brandt, Thomas Carson, Michael Smith, Charles Taliaferro and Jason Kawall. It has also been suggested that Francis Hutcheson, David Hume, Adam Smith and R. M. Hare can be viewed as offering some kind of ideal observer theory (see [16], [10] and [30]). For a general introduction, see [30]. Contributions to the discussion of this theory or topics related to its truth-value include, for instance, [1], [2], [3], [6], [8], [9], [10], [12], [13], [14], [15], [16], [17], [18], [19], [22], [23], [47], [48], [26], [27], [28], [29], [39], [51], [55], [63].

2 (O) entails that it ought to be the case that A only if every ideal observer wants it to be the case that A. This is a kind of existence internalism. However, our motivational internalism (CBMI) doesn’t follow from (O) in itself. We also need (B), or some similar principle, to prove this.

3 For the record, I don’t think that (O) and (B) are sufficient. An adequate ideal observer theory would have to include more conditions, e.g. we should also assume that every ideal observer has knowledge of, or at least true beliefs about, all (relevant) non-moral facts. If the purpose of the paper were to develop a fully reasonable ideal observer theory much more would also have to be said about what we mean by an “ideal observer”. Fortunately, this is not needed for our present purposes.

4 The literature on internalism is too vast to quote in full. I simply refer the reader to [7] and [59] for an overview and many references to the relevant literature. For more information on some different versions of internalism, see e.g. [45, 144–145] and [53].

5 For some arguments against internalism, see e.g. [11, chapter 3], [35], [37], [43], [52], and [54].

6 For a similar classification, see [37].

7 [5]. See also [4].

8 [31], [47], [48], [58].

9 [34], [56], [60]. See also [20], [38], [45, Chapter 5], [46].

10 The precis formulations of (O), (B) and (CBMI) are important. In (O) and (B) we don’t use counterfactuals. This is motivated by the fact that we want to avoid the so-called “conditional fallacy” mentioned in [24] (see also [59] and for a possible response [57]) and the problem with the so-called “example model” of motivational internalism that Smith discusses in chapter 1 in [48, 18-20]. Another reason is that it is easier to formulate a plausible deontic logic with (O) (and (B)). But, we use a counterfactual in (CBMI). One reason for this is that I don’t think any actual human being is an ideal observer. In fact, I believe that it is probably historically (but not logically) impossible for an actual human being to be an ideal observer. However, we want to be able to speak about what actual
people *would* want, if they *were* ideal observers. (CBMI) emphasizes this point. Nevertheless, the difference between a counterfactual version of (CBMI) and a version of (CBMI) with a material implication should not be exaggerated, since if the initial necessity operator is strong enough the two versions will be equivalent. The same goes for (B), but not for (O). Furthermore, if the necessity in (B) is sufficiently strong, (B) supports the corresponding contingent counterfactual. For in most plausible counterfactual systems it follows that if it is necessary that \( A \) implies \( C \), then if it were the case that \( A \) then it would be the case that \( C \). Unfortunately, it would take us to far from our main topic to discuss these reasons in more detail in the present paper. Note that the necessity in (O), (B) and (CBMI) are usually interpreted as some kind of conceptual or analytic necessity, if not otherwise noted. However, our main argument goes through for any S5-like necessity concept. So, nothing *essential* hinges on this choice.

11 [30] mentions some arguments for and against the ideal observer theory. The following papers, among others, include critique of Firth’s particular version: [8], [9], [19], [22], [23] and [39]. [17], [18] and [55] contain some responses. For an introduction to some arguments for and against internalism, see [7] and [59].

12 “Possibilism” is used in many senses. According to one popular interpretation, call it (P), it is the thesis that there are things that do not exist. When we speak of a “possibilist” interpretation of the quantifiers, we mean that they range over absolutely everything, including merely possible objects (if there are any), not only existing things. Given this interpretation, (P) is contingent. Obviously, if some ideal observer doesn’t exist, (P) is true.

13 It might be argued that [41] has shown that “possibilism” is absurd, and therefore, indirectly, that the “possibilist” answer to Harrison’s objection is absurd. Quine’s arguments are certainly intriguing, but they don’t seem conclusive. For some possible responses, see e.g. [44] and [40, chapter 5].

14 For some introductions to possibilism, actualism and related topics, and many relevant references, see e.g. [36], [42] and [62]. For an interesting recent defence of “possibilism”, see e.g. [61]. It should be noted that Williamson doesn’t like the expression “possibilism”. Nevertheless, I think he can be interpreted as a kind of possibilist, since he allows unrestricted quantification over absolutely everything, including non-existing things (if there are any). An actualist that accepts the ideas in [33], for instance, can also avoid Harrison’s objection, given that she accepts the existence of at least one non-concrete ideal observer. Furthermore, every actualist that believes that God exists and that God is an ideal observer, can also avoid Harrison’s first argument. The same is true of an actualist polytheist who believes that all gods are ideal observers. Whether such actualist positions are plausible or not is another question. I think there are some good arguments for possibilism. However, even if the possibilist answer to this problem should turn out to be implausible, these “actualist” moves show that it might be possible to avoid Harrison’s first argument. Unfortunately, it would take us to far from our main topic to discuss various arguments for and against possibilism and a possibilist interpretation of the quantifiers in this paper.

15 E.g. all of the following equivalences are necessary. It is permitted that \( A \) iff it is not forbidden that \( A \). It is forbidden that \( A \) iff it is not permitted that \( A \). It is obligatory that \( A \) iff it is forbidden that not \( A \). It is forbidden that \( A \) iff it is obligatory that not \( A \). It is permitted that \( A \) iff it is not obligatory that not \( A \). It is obligatory that \( A \) iff it is not permitted that not \( A \).
16 [12] introduces a version of the ideal observer theory that doesn’t assume that all ideal observers want the same things.

17 [8], [12] and [13] argue that there might be more disagreement among ideal observers than Firth accepts, while [55] defends Firth’s view.

18 The propositional logic used in the proof is standard and need no further comments. The modal steps require an S5-like necessity operator. It might be interesting to note that if we don’t use a counterfactual in (CBMI), but an ordinary material implication instead, the modified steps in the argument go through in every normal modal logic. The counterfactual steps are plausible in almost any counterfactual logic, for instance in Stalnaker’s and Lewis’s favourite systems, see e.g. [49], [50], and [32]. Step 7 in the (IOT) argument is perfectly reasonable for a possibilist and also for an actualist that allows quantification over (abstract or non-concrete) ideal observers. For a possibilist, step 7 doesn’t entail that c exists. Step 7 is reasonable in almost every quantified modal logic with classical quantification theory and a non-empty, constant (fixed) domain, see e.g. [21], for instance system Q1, [25, 141-169] and [33]. The same is true of step 12 and step 17.

Daniel Rönnedal
Stockholm University
Department of Philosophy
106 91 Stockholm, Sweden
<daniel.ronnedal@philosophy.su.se>

References


