

Filosofia Unisinos
Unisinos Journal of Philosophy
23(3): 1-16, 2022 | e23304

Unisinos – doi: 10.4013/fsu.2022.233.04

Article

Access and Unknowable Obligations

Acesso e obrigações desconhecidas

Oleh Bondar

<https://orcid.org/0000-0003-2657-6510>

Shandong University, Jinan, Shandong, China. Email: olegbondarb581@gmail.com.

ABSTRACT

The article is devoted to the question of whether unknowable obligations are possible. According to the popular view (known as Access), an act is obligatory only if its agent can know that this act is obligatory. Sorensen (1995) argues against Access, and Sider and Wieland both defend Access by suggesting (S) – *For any obligation O, individual X must refrain from making O unknowable*. We consider various difficulties concerning the possibility of (S) – an infinite regress, self-referential nature of (S), unprovability of (S). It is argued that (S), from a certain perspective, is itself the unknowable obligation.

Keywords: Obligations, access, regress, self-reference.

RESUMO

O artigo é dedicado à questão de saber se as obrigações desconhecidas são possíveis. De acordo com a visão popular (conhecida como Access), um ato é obrigatório somente se seu agente puder saber que esse ato é obrigatório. Sorensen (1995) argumenta contra Access, e Sider e Wieland defendem Access sugerindo (S) – *Para qualquer obrigação O, o indivíduo X deve abster-se de fazer O sem conhecimento*. Consideramos várias dificuldades em relação à possibilidade de (S) – um regresso ao infinito, natureza autorreferencial de (S), improbabilidade de (S). Argumenta-se que (S), de uma certa perspectiva, é si a obrigação desconhecida.

Palavras-chave: Obrigações, acesso, regresso, autorreferência.



Introduction

Imagine the following situation. I participate in the presidential elections. I intend to vote for one of the candidates (say Tom) because I share his political views. Suppose now that Tom is a corrupted politician, but I do not know this fact. Of course, I agree that if Tom is a corrupted politician, then I have the obligation not to vote for Tom because I have the obligation not to vote for corrupted politicians. Then it turns out that I have the obligation not to vote for Tom, but at the same time, I do not know that I have the obligation not to vote for Tom (because I do not know that Tom is a corrupted politician). The question is: is a certain obligation an obligation (for me) if I do not know that I have this obligation? In other words, are there such things as *unknowable obligations*?

At first glance, it seems that the fact that a person X has the obligation O always presupposes X's knowledge that he has O. If O (the obligation not to vote for Tom) is an *obligation*, even if X does not know that he has such an obligation (and, as a result, he votes for Tom), then based on the fact that violation of obligation is blameworthy¹, X should be blamed for voting for Tom even if X didn't know that he had O. But that seems to be extremely counterintuitive. In fact, it would be plausible to conclude that X should not be blamed for his violation of O. Therefore, those who deny that unknowable obligations are obligations, accept Access²

(Access) An act is obligatory only if its agent can know that this act is obligatory.

Sorensen (1995) argues against Access. By Access, if I cannot know that I have an obligation, then I do not have that obligation. This conclusion, as Sorensen expresses himself, "dumbs down ethics" (1995, p. 254). Moreover, we can imagine a certain "epistemological Robinson" – a person who cannot know that he has a moral obligation. Then, by Access, we can conclude that he has no moral obligations. This is highly implausible. Finally, it follows from Access that our obligations are epistemologically dependent – a certain obligation is an obligation only insofar we know about that obligation. So by eliminating the source of my knowledge that I have a certain obligation, I eliminate that obligation, since I no longer have a possibility to know if I have that obligation. Sorensen (1995) gives the following example:

If all of one's obligations were knowable, one could mute the call of duty by diminishing one's cognitive capacity. After all, if I am not obliged to do something, then I am permitted to refrain from doing it. Thus if I can preclude the possibility of knowing that I am obliged to donate some of my inheritance to charity (by burning rather than reading the only copy of a will), then I can ensure that I am permitted to keep the entire inheritance. Simple incuriosity would considerably narrow one's obligations (Sorensen, 1995, p. 254).

By Access, I have the obligation to donate my inheritance to charity only if I can know that I am obliged to do that. If the will is the only source from which I can know that I am obliged to donate for charity, then if I burn the will (without reading it), I eliminate the only way to know if I have the obligation to donate for charity. Hence, I no longer have the obligation to donate for charity (after I burned the will), since I have no way to find out if I'm obliged to donate for charity. Thus, it follows from Access that I can *make* my obligation unknowable. This is a dubious consequence of Access. In the example above, if I have an obligation to donate for charity, then I must donate for charity³. It is not within my power to

¹ See Sher (2009) about the relation between knowledge (information) and responsibility.

² Sorensen introduces the term (Access) in Sorensen (1995).

³ I will not touch here the distinction between primary and secondary obligations (as well as discussions about incompatible obligations) that can serve as a possible counterexample against the claim that if I have an obligation to do O, then I must do O. See Zimmerman (1996)

change the fact that I am obliged to respect my parents – and thus to donate for charity. However, Access seems to assert otherwise – it is possible for me to violate my moral obligation to donate for charity by keeping myself ignorant about the source of that moral obligation. Thus, if I can refrain from making the obligation knowable, I can refrain from fulfilling the obligation.

1. In Defence of Access

In response to Sorensen, Sider (1995) defends Access by introducing the following principle:

(S) For any obligation O , individual X must refrain from making O unknowable.

But (S) faces a problem. Suppose that X is a person who wants to avoid fulfilling his obligation⁴. Now, from X 's perspective we can argue that (S) generates an infinite regress of obligations, for if (S) is the obligation grounding the primary obligation, O , X can say that (S) is unknowable to him. Thus, to make (S) knowable (to prevent X 's making (or keeping) (S) unknowable), the defender of (S) must introduce another obligation:

(SS) X must refrain from making (S) unknowable.

The result will be the same: X can always answer that (SS) is unknowable to him, so to make (SS) knowable, the defender of (S) must introduce another obligation – (SSS) – such that by (SSS), X is obliged to refrain from making (SS) unknowable, and so on endlessly. Wieland (2014b, p.58) introduces this form of the argument against Access as follows⁵:

- (1) I can have the obligation O only if I can know that I have O
- (2) I can have O only if I have another obligation O^* , such that O^* is the obligation to refrain from making O unknowable
- (3) Even if I don't know whether I have O , I always have O^*
- (4) I always have O^*
- (5) If I always have O^* , I always have another obligation – O^{**} - such that O^{**} must prevent me from making O^* unknowable, I also have the obligation O^{***} such that O^{***} must prevent me from making O^{**} unknowable, and so on
- (6) I have an infinite number of obligations
- (7) I have an infinite number of obligations only if I can know all of them
- (8) I can't know an infinite number of obligations
- (9) Thus, Access is false

This is how the argument goes. For any obligation O_n , I always have the obligation O_{n+1} – the obligation to refrain from making O_n unknowable. Thus, I have an infinity of obligations. By Access, I can have the obligation only if I can know that I have the obligation. I have therefore infinity of obligations, and I can have an infinity of obligations only if I can know infinity of obligations.

There are two crucial points concerning (1-9). Let us look closer at (6). Is it possible to have an infinite number of obligations? At least we can argue that the claim *I have an infinity of obligations* is not contradictory. Suppose, for example, that I cannot have an infinity of obligations. If Access is true, then for a certain obligation O_n I no longer have the obligation O_{n+1} – the obligation to refrain from making O_n unknowable. But, by the argument above, I always have O_{n+1} if I have O_n . However, should we consider

⁴ Wieland (2014a, 2014b) calls this person *shirker*.

⁵ My presentation of this argument is slightly different from Wieland's, but in general, they are identical

the obligation O_{n+1} as some kind of additional knowledge about the primary obligation (say *not to burn the will*)? I agree with Wieland (2014a) that if we consider O_1 as the obligation *to refrain from making* O (where O is our primary obligation) *unknowable*, we are on the same, epistemic level" as in the case with our primary obligation O - O_1 does not require us to improve our epistemic position, namely to improve our knowledge to learn something new about O . Similarly, if I have O_{n+1} , this obligation doesn't tell me that I must learn something new compared with my obligation O_n - I must refrain from making O_n unknowable, but I am not obliged to learn something new about O_n . If so, it is possible for me to have an infinity of obligations.

We can now answer to the second crucial point - (8). By (8), we cannot have an infinity of obligations because we cannot know an infinity of obligations. But, as I pointed above, this statement is refutable. The defender of this position asserts that our cognitive abilities are limited, and so we cannot know an infinity of obligation. Thus, by Access, we do not have an infinity of obligations. This argument can be summarized as follows:

- (a) Suppose I have an infinity of obligations
- (b) I can have an infinity of obligations only if I can know infinity of obligations
- (c) My cognitive abilities are limited, so I cannot know an infinity of obligations
- (d) If I cannot know infinity of obligations, I do not have an infinity of obligations
- (e) I do not have an infinity of obligation
- (f) Thus, Access is false

As we saw above, the correctness of this argument depends on what we mean by the word "know". If we say that, for the obligation O^* , the word "know" means "to refrain from making the primary obligation O unknowable", the statement "We have an infinity of obligations" is compatible with Access, and also with the claim that our cognitive abilities are limited.

Nevertheless, the opponent of Access can always say that, by Access, I can have the obligation O only if can I know another obligation, O^* , such that knowledge about this secondary obligation differs from knowledge about our primary obligation, and thus an infinite regress reappears. Consider:

- (10) For any obligation O , I can know O only if I can know O^* such that O^* is about O

An important consequence of (10), as one might argue, is this:

- (11) For any obligation O , it is necessary for me to have knowledge⁶ about O^* in order to have knowledge about O . Thus, if O is the obligation, it is necessary for me to know O^*

Given (10) and (11), consider:

- (12) I have the obligation to refrain from burning the will
- (13) I have the obligation to refrain from making (12) unknowable
- (14) I have the obligation to refrain from making (13) unknowable

Given (12 -14), the opponent of Access can argue that we have a new regress. By Access, I can have (12) only if I can have (13). From (11), my knowledge of (13) is a necessary condition to have (12), and my knowledge of (14) is a necessary condition to have (13). What is (13)? By (11), if I have the obligation O ,

⁶ Here the phrase *to have knowledge* does not mean *to have actual knowledge*. Rather, this phrase expresses the *possibility* to have knowledge.

I necessarily have O^* (such that O^* is about O), and O^* asserts that I can know O only if I can know that I can know O (that is if I can know O^*)⁷. Thus, if I *actually* have (12), then by (13) I know (12) only if I know that I know (12) (that is, if I know (13) such that (13) is about (12)). But “to know something” is not the same as “to know that you know something”. Thus, the argument will be as follows:

- (15) Suppose I actually have (12)
- (16) I have (12) only if I know (12)
- (17) I know (12) only if I know (13)⁸
- (18) If I know (13), then I know that I know (12)
- (19) [I know X] differs significantly from [I know that I know X]
- (20) [I know (12)] differs significantly from [I know (13)] (that is, [I know that I know (12)])
- (21) (13) is irreducible to (12)
- (22) If (13) is irreducible to (12), I have an infinity of *essentially different* obligations
- (23) I can have an infinity of obligations only if I can know infinity of obligations
- (24) I can't know infinity of *essentially different* obligations
- (25) Thus, Access is false

This argument rests on two key premises. Firstly, the proponent of (15-25) asserts that my knowledge about (12) differs *essentially* from my knowledge about (13) – that is, from the fact that I know that I know (12). Secondly, the proponent of (15-25) accepts (11) – the statement (17) logically follows from (11). Thus, (15-25) is true only if (13) differs essentially from (12) (that is, (21) is true), and if (11) is true. Regarding (15-25), (11) asserts that *I know that I know (12)* is a necessary condition for me to know (12) – I can know (12) only if I know that I know (12) (i.e., if I know (13)). Is it true? Consider:

- (26) $O_1 = I \text{ know that } O$
- (27) $O_2 = I \text{ know that I know that } O$
- (28) $O_3 = I \text{ know that I know that I know that } O$

Thus, the proponent of (15-25) asserts that I can have O_1 only if I can have O_2 . This is false. I can have propositional knowledge (knowledge of an object, fact, etc.) without having reflexive knowledge (knowledge of my knowledge of an object, fact, etc.). Therefore, (11) has false instances – it is *not necessary* for me to have O_2 in order to have O_1 . Does this fact entail the falsity of Access? No. Otherwise, I could say that X is obliged not to burn the will even if X does not know that he is obliged to refrain from making the fact that X is obliged not to burn the will unknowable. But in fact, X is not obliged not to burn the will, if he doesn't have

⁷ Note that O^* , taken separately from the context of our problem with (Access), does not assert that I can know O only if I can know O^* . O^* there is just an assertion that X must refrain from making O unknowable. However, we consider this problem in the context of shirker-style reasoning (i.e. from the standpoint of X). By Access, in order to have O , O should be knowable to me. From (S), I must refrain from making O unknowable, and thus I have O^* . But a shirker now within her rights to interpret the whole context of our problem as follows – if I must refrain from making O unknowable because I have an obligation O^* , and I have O such that O^* is about O , then I can know that I have O only if I can know that I have O^* . Given that, a shirker introduces the sentence (10). Thus, in the context of shirker-style reasoning, I cannot know O if I cannot know O^* . Here we need to distinguish between two perspectives – between shirker-style and non-shirker-style (here non-shirker is a person who do not try to make her obligation unknowable) types of reasoning. A shirker asserts that, necessarily, I can know O only if I can know O^* . But I argue that this shirker's assertion is false. From my point of view, it is not true that I can know that I have O only if I can know that I have O^* . So, my argument against shirkers here is as follows. A shirker thinks that Access implies (11). Given (11), a shirker concludes that it is impossible for her to get the knowledge that she has O because it is impossible for her to get the knowledge that she has O^* . So, a shirker concludes that if (11) is an instance of (Access), then (Access) is false since (11) generates an infinity of obligations. Contrary to X, I argue that (11) is indeed false, but this falsity does not entail that Access is false (see my argument under (26-28)). Thus, I argue that both (10) and (11) are false – (10) and (11) both follow from Access only if we accept a shirker's point of view. But this point of view is false.

⁸ The difference between (17) and (13) is that (13) expresses that I can know O only if I can know O^* . Suppose that I have O . (17) then is about the fact of my *actual* having O – I have O , only if I have O^* .

the obligation to refrain from making his primary obligation (not to burn the will) unknowable. Suppose I say to X: "You must refrain from voting for corrupted politicians (say Tom) even if you do not know that you have the obligation to refrain from voting for corrupted politicians because it is not necessary for you to know that you have the obligation to refrain from voting for corrupted politicians in order to know that you must refrain from voting for Tom". X would say that this claim is absurd, and we must agree with him. What the fact that it is not necessary for me to have O2 in order to have O1 really shows is that (13) does not differ *essentially* from (12). We can say that there is no epistemic improvement of our knowledge in the case of (13) comparing with (12). If I assert that I have (13), then I rather assert that:

(29) If I have the obligation to refrain from burning the will, I am obliged not to downgrade my current epistemic level concerning my obligation to refrain from burning the will

But I do not want to say that (13) entails (30),

(30) If I have the obligation (12) then I have the obligation (13) to consider something, (12), as the obligation of a certain epistemic level

because it would follow from (30) that I must accept

(31) If I have the obligation (13), I have the obligation to distinguish between (12) and (13)

But if (31) were true, then given the fact that I have (12) if I have (13), (32) would be true:

(32) If I have the obligation (12), (12) obliges me to distinguish between (12) and (13),

And this statement is clearly false – if I have the obligation to refrain from burning the will, the fact of my having this obligation does not oblige me to distinguish between different levels of obligations. I can have the obligation (12) without knowing a theory of obligations (that is, without having theoretical knowledge about the nature of obligation). Otherwise, I would have an infinity of distinctions within my primary obligation:

(33) If I have the obligation (12), and (12) obliges me to distinguish between (12) and (13), (12) obliges me to distinguish between (12) and (14)... and so on, and (12) obliges me to distinguish between (*distinction of (12) and (13)*) and (14)... and so on

In fact, I do not need to make such distinctions to have an obligation. For suppose otherwise. If X wants to avoid his obligation to refrain from burning the will, he could argue that, by (Access), he can have this obligation only if he can know that he has this obligation. But I can know that I have a certain obligation, as X would say, only if I can know what does the word *obligation* means. Thus, as X argues

(34) For any obligation O, I can have O only if I can have a clear meaning of the term *obligation*

Now, suppose that X argues as follows. For any obligation O, I do not have O, because it is impossible for me to understand what does the term *obligation* means. Suppose I want to understand the meaning of *obligation*. Define O (perhaps circularly) as an entity that possesses some obligatory-making property P. Suppose further that X is a Predicate Nominalist, in the Armstrong's (1974, p. 191) sense), and he wants to clarify what does it mean for O to have P as follows

(35) O has P = def. O falls under the predicate P

Now X argues that (35) does not still provide him with a clear understanding of what does it mean for O to have P. X tries to define properties in terms of predicates, but (35) does not eliminate the concept of property because *falling under the predicate P* is itself a property (say P1). X, again, tries to define the property *falling under the predicate P* in terms of predicates, and immediately falls into the infinite regress described by Armstrong (1974):

- (36) O has P = def. O falls under the predicate P1 of falling under the predicate of P
 (37) O has P = def. O falls under the predicate P2 of falling under the predicate P1 of falling under the predicate P

Given this result, X could argue that it is impossible for him to understand how O could have P. If so, X can never have a clear meaning of the term *obligation*. Finally, X concludes that, by (Access), if he cannot know the meaning of *obligations*, X doesn't have obligations.

One might reply that (as was shown above) it is not necessary for me to understand a theoretical meaning of the term *obligation* in order to understand the term *obligation*. However, this is not a trivial problem – even if the defender of Access can block X's inference by claiming that this analysis does not essentially affect the problem with Access because it proves too much, we can formulate a similar argument against the possibility of Access. In the case of X's analysis, X should always introduce new definitions of *O has P* to understand the meaning of *has P*". Within this regress, X must introduce (36) to understand (35), (37) to understand (36), and so on. But suppose that X has a clear understanding of his obligation not to burn the will (and, of course, X understands the meaning of *obligation*), but he wants to make himself unknowable about his obligation. In order to prove to X that he must refrain from making O (the obligation to refrain from burning the will) unknowable, I must introduce to X, by Access, that X has the obligation to refrain from making O unknowable. Doing that, we face a similar regress as with the case of X's analysis in (35-37). Wieland (2014b, p. 63) introduces this regress as follows:

- (38) I have to demonstrate that X has O
 (39) I first have to demonstrate that X has the obligation to refrain from making O unknowable
 (40) I first have to demonstrate that X has the obligation to refrain from making (39) unknowable
 (41) I first have to demonstrate that X has the obligation to refrain from making (40) unknowable

The argument (38-41) entails that I can never prove to X that he has the obligation O. Every time I try to prove to X that he has the obligation, I have to introduce more and more obligations. And every time I introduce a new obligation (that is, I provide proof that X has the obligation), X is within his right to ask me for a new proof. Hence, even if we can reject the argument (1-9), we are not in a position to reject (38-41), because even if we know that X does not actually have an infinity of obligations of the form O*, O**, O***... if X can have O, it does not prevent us from having an infinity of *explanations*⁹. For example, suppose I can know O. Then, a shirker will assert that if I can know O, then necessarily, I have O*, O**, etc. But in fact I can know O even if I am not in a position to understand, for example, a relation between O** and O*. The fact of my having obligation does not entail that I have full and consistent theoretical explanation of the fact of my having this obligation.

⁹ Wieland (2014b, p.60): "...if I learn the utilitarian principle that I ought always to maximize utility, then, in a certain sense, I have thereby learned many obligations, namely the obligation to maximize utility at t1, to maximize utility at t2, and so on". Could this proposal serve as a solution to the problem with an infinite regress of explanations? Wieland argues that a universal generalization, "Maximize utility!", grounds plurality of its instances. But Rosen (2017) notes that "universal generalizations are partly grounded in their instances", so Wieland's proposal is circular, and thus cannot be true.

2. (S), Self-reference, and Unknowability

Wieland (2014b, p. 64) suggests the following solution (call this solution (WA)). Suppose X has the obligation to refrain from burning the will. X wants to make (or keep) this obligation unknowable. X will not be able to refrain from this obligation, because he has O1 and O2:

- (O1) X should refrain from burning the will
- (O2) X should refrain from making (or keeping) O1 or O2 unknowable

The argument runs as follows. The obligation O1 is either knowable to X, or unknowable to X. Suppose that O1 is unknowable to X. By O2, X is obliged to make O1 knowable. Thus, X cannot refrain from making O1 knowable. However, X can say that O2 is unknowable to him, and so O1 is unknowable to him. But O2 asserts that X has the obligation to refrain from making (or keeping) O2 unknowable. Thus, if X asserts that O1 is unknowable to him, it is true (by O2) that X should refrain from making O2 unknowable. Then, by (O2), it is true that X must make O2 knowable, so O2 entails that X should make O1 knowable even if X does not know that X has the obligation to refrain from making O1 unknowable.

Suppose otherwise – O1 is knowable to X. X wants to make O1 unknowable. Then, X knows that he should refrain from burning the will, but X wants to make this obligation unknowable. But if O1 is currently knowable for X, X knows that he must refrain from burning the will – that is, X must refrain from making the obligation to refrain from burning the will unknowable. Thus, if O2 is in power, X is obliged to refrain from making the obligation not to burn the will (O1) unknowable. X, as before, can argue that O2 is unknowable to him, and if O2 is unknowable to him, then O2 is not in power. But O2 is the obligation to refrain from making (or keeping) O1 or O2 unknowable. Thus, if X argues that O2 is unknowable to him, he does not have O2. But then, X has O1, and thus cannot make (or keep) O1 unknowable. However, the obligation to refrain from making (or keeping) O1 unknowable is a part of O2, so X cannot make O1 unknowable. Could X say that he has not O1 because O2 is not knowable? No, because X cannot refrain from making (or keeping) O2 knowable – doing this, he violates his obligation O2. Thus, as Wieland argues, X cannot evade O1 by claiming that O1 (or O2) is unknowable to him.

But isn't it true that O2 is a self-referential obligation – the obligation that asserts something about itself? Wieland seems to agree with this diagnosis: "...[X] should make O2 itself knowable, and refrain from making O2 unknowable. In this respect, O2 is self-referential..."¹⁰ (Wieland, 2014b, p. 64). Why should we introduce such a suspicious „entity“ as a self-referential obligation? The answer is this. O2 guarantees us that X will not be able to evade O1 or O2 by claiming that O1 and O2 are unknowable to him. Therefore, O2 plays a decisive explanatory role in the argument. But there is another question – is O2 true? As I understand it, there are two possible ways of defining (O2). If (O2) is just a conjunctive sum of obligation concerning O1 and O2 (the obligation to refrain from making and keeping O1 unknowable and the obligation to refrain from making and keeping O2 unknowable), I find the argument above implausible – if X asserts that he doesn't have O1 because O2 is not knowable to him (O2 is the obligation to make O1 knowable), then the defender of (WA) simply asserts that X is not allowed to evade O2 (that is, to refrain from making O2 knowable) because O2 is the obligation asserting that X must make this

¹⁰ Wieland (2014b, pp. 64-65): "It is important to emphasize that Access is in force. This means that if O1 or O2 is not knowable to [X], then [X] does not have that obligation. Yet, given the special nature of O2, it can be shown that such a situation will not occur". I don't think that this proposal can solve our problem. At least from Fine (2010) we know that the sentences of the form *If p, then q* are in general not true. If X asks why he is not allowed to evade O1, the defender of (WA) will reply that X is not allowed to evade O1 because X has O2. Thus, O2-or-O1 grounds O1, and X has O2-or-O1 because X has O2. But given Fine's analysis, O2-or-O1 must be grounded in both O1 and O2. Does O1 ground O2-or-O1? No, because the proponent of (WA) presupposes that O2-or-O1 grounds O1. If O1 were a ground for O2-or-O1, then the whole argument would be circular.

obligation, O2, knowable! But this is wrong – such obligations do not exist. For suppose that O2 is such an obligation. O2 then asserts that X has O1. The defender of (WA) would then argue that X has O1 because X has O2. But if we asked why X has O2, the defender of (WA) would answer that X has O2 because O2 asserts that X has O2! But we can't assert that O2 is true because O2 asserts that O2 is true. It is clear that (WA) should not be understood in this way. Thus, (O2), as I understand it, should be treated as a disjunctive self-referential obligation. Again – if (O2) is true?

Logic indeed allows us to formulate sentences like „If O2, then either O2 or not-O2“, or „O2's being the case grounds O2-or-O1's being the case“¹¹. In these cases, O2 entails something or grounds something. But in the case of (O2), O2 is the sentence about O1 and O2, - O2 is a part of itself. We have

(42) O2 is [O2 V O1]¹²

Then, replacing O2 with O2 V O1, we have

(43) O2 is [O2 V O1] V O1]]

(44) O2 is [O2 V O1] V O1]] V O1]]] V O1]]]]...]

Again, we face an infinite regress. However, the defender of (WA) could resist this argument by claiming that this regress is not *vicious* due to the nature of disjunction. However, even if this regress is not vicious and thus cannot serve as a counterexample to (O1-O2), we can formulate the following argument which is an instance of (O1-O2)

(A) I should refrain from making (or keeping) (S)¹³ unknowable

(S) I should refrain from making (or keeping) (S) or (A) unknowable

The sentence (A) is true. I can have an obligation O only if I can know that I have O – that is if I have Sider's (S). (S) then will be the obligation of the special sort – *the obligation about all obligations*. Of course, I can have (A) only if I can know that I have (S), so by (S), I must refrain from making (A) unknowable. Could I evade my obligation to refrain from making (or keeping) (A) unknowable? The defender of (WA) will respond – no, I can't evade (A) because I have (S). For suppose I assert that (S) is unknowable to me. But if (S) is unknowable to me, then (A) is knowable to me, so I can't evade (A). (A) asserts that I'm obliged to know (S). (S), by definition, is the obligation about every obligation, so (S) is about (A) and (S). Consider now the following case:

(C1) A is unknowable to me

In the case of (C1), (S) is knowable to me, because (A) is not knowable to me – it follows from the disjunctive nature of (S). But (A) is about (S) (about the obligation to refrain from making any obligation unknowable). Thus, if (A) is unknowable to me, then I know (from (S)) that I should refrain from making (or keeping) (S) unknowable, and I don't know (from (A)) that I should refrain from making (or keeping) (S) unknowable. Contradiction!

It is not surprising that we got such a result. (O1-O2) and (A-S) both contain self-reference – O2 says about itself, and (S) says something about (S), which in turn says about (A) such that (A) is about (S). We

¹¹ This is controversial. Many philosophers believe that disjunct's being true ground the whole disjunction, while others (Fine, 2010, Litland, 2015) suggest the view that disjunctions are always grounded in their *true disjuncts*

¹² For helpful comments, see Berker (2019)

¹³ Recall that, by Sider's (S), if X has O, X must refrain from making O unknowable

are within our rights to assert that (O2) - *X should refrain from making (or keeping) O1 or O2 unknowable* – is not consistent. If so, then (S) - *I should refrain from making (or keeping) (S) or (A) unknowable* – is not consistent too. But in the case of (A-S) we are faced with another problem - the problem with *the obligation about all obligations*. Recall that (S) asserts – if you have O, you must refrain from making O unknowable. Thus, for any O, if you have O, you have (S) such that (S) is about O. But (S) is itself O – the obligation to refrain from making your obligations unknowable is itself an obligation. So, (S) asserts that you must refrain from making (S) unknowable – you can have (S) if you can know (S). But (O1-O2) asserts

(45) X can have O without knowing that X has O

By letting O be (S), we have

(46) X can have (S) without knowing that X has (S)

(S) is the obligation about *all* obligations, so (S) is about (S). (S) is also an obligation to refrain from making (S) unknowable, so we have (S) only if we can know (S)

(47) X can have (S) if X can know that X has (S),

(47), together with (46), implies

(48) X can know that X has (S) without knowing that X has (S)¹⁴

(48) is false – it is possible that I could *have* O without knowing that I have O, but it is impossible that I could *know* that I have O without knowing that I have O. Does it mean that (S) is, in fact, self-defeating (because self-referential)? Suppose there is some full, consistent logical system of obligations Q containing the rule about all obligations (S). Is Q able to prove its own consistency? No, because this possibility would contradict Gödel's second incompleteness theorem. From the perspective of (38-41) X can reply that if the obligation is not proven as an obligation, it cannot be considered as an obligation. But this answer is too strong – we do not need to be ethical instructors in Sorensen's sense to recognize our obligations. Moreover, as we could see, there are *unprovable* obligations. But it is not the case that a certain obligation should necessarily be theoretically proven in order to be considered as an obligation. Therefore, the answer to the question "*Are there unknowable obligations?*" will be as follows – *yes, (S) is an unknowable obligation*. If unprovability implies unknowability then, by (38-41), (S) is unknowable. However, we can always defend Access and (S) by claiming that (S) is not unknowable for the case of (1-9). Obligations are not logical truths. If I am obliged to respect people, it would be wrong to assert that

¹⁴ An anonymous referee of the journal offers a suggestion that (48) must be formulated rather as follows: (48*) "It is possible for X to know that X has (S) without X's actual knowing that X has (S)". I agree that (48*) is true for any ordinary obligation like to *donate for charity if the will says that I must donate for charity (call this O)*. However, there is a crucial disanalogy between O and (S). O does not assert that I must make O knowable or refrain from making O unknowable. It is possible then that I can know O without having actual knowledge that I have O. But consider the case of (S). (S), unlike O, is a self-referential obligation (because (S) is about any obligation, including (S)) – an obligation saying of itself. Suppose now that (S) applies to (S). We have now an obligation of the following sort

(S) Refrain from making me unknowable / make me knowable!

Here an obligation to refrain from making (S) unknowable is (S) itself. (S) is then at the same time both, an obligation and the source of knowledge that I have this obligation. Compare now O and (S). Suppose that (S) does not exist. Even if (S) does not exist, I can know that I have (S) without having actual knowledge that I have O. But in the case of (S), I cannot say that I can know that I have an obligation to refrain from making this very obligation unknowable, if there is no such obligation, or this obligation is not actually knowable to me, because only this obligation, (S), and nothing else, says me that I have this very obligation. Thus, I believe that a case of (S) is something different from any case of other obligations.

I am obliged to respect people because I can logically prove this truth. I am obliged to respect people because this obligation is an ethical rule. It would be ridiculous to ask about the truth conditions for the obligation to respect people. But (S) is both an obligation and a truth condition¹⁵ for an obligation. This is how (S) differs from other obligations. This complexity of (S) is the source of the contradiction. But if we consider (S) as an ethical attitude or rule (like the obligation to respect people), the contradiction will disappear¹⁶, and we can legitimately refer to (S) without violating Access.

In agreement with Sider and Wieland, I argue that the argument from infinite regress of *obligations* poses no difficulty for (S) because none of its sub-arguments (an impossibility of having an infinity of obligations, an epistemic regress, an infinite regress in defining "obligations") is valid. But contrary to Wieland, I argue that we actually have an infinity of *explanations*, and thus we have an explanatory regress.

In the section 3, I argue that Wieland's argument (O1-O2) does not provide us with a definitive solution to the problem with (S). The source of the problem is that (S) itself a self-referential universal obligation – the obligation about *all* obligations, and thus we have a good reason to think that (S) cannot be proven. However, if my argument is correct, (S) does not need to be proven. The whole issue rests on the assumption that there is a loophole between (S) and its knowability. But there is no such loophole. If I'm obliged to respect people, I know that, and so I know that I'm not allowed to evade this obligation. Surely, I can evade my obligation by making (S) unknowable, but it follows from the nature of ethical obligations that I'm not allowed to do that. In fact, (S) is an essential part of our ethical obligations, so (S) is rather an ethical than a logical statement.

3. Objections and Replies

It is now time to clarify some questionable points of my arguments and reply to criticism.

I have argued that a shirker can introduce an argument, according to which Access entails that in order to be able to know O, it must be possible to know O*. A shirker formulates this argument by introducing (10) and (11). Then, a shirker argues that there is a substantial difference between O and O*. If so, as a shirker thinks, then it follows from Access that I must know infinity of substantially different obligations. It is impossible, and so a shirker concludes that Access is false. In this respect, I proposed an argument that this shirker's conclusion is not sound. Following X, we can know O only if we can know O* (that is, only if we can know that we can know O). But I have provided an argument that, for any agent X, it is not the case that X can know O only if she can know O*. Thus, it is important to clarify that such sentences as (10) and (11) are true *only from shirker's point of view*. Being formulated from a non-shirker's point of view, it is not the case that, for any obligation O, O can be knowable only if the next-order obligation, O*, can be knowable. Thus, it is true, following shirker, that O* asserts that I can know O only if I can know that I can know O (that is if I can know O*). But this is not true from the point of view of the opponent of a shirker, - from this point of view, O* is just an obligation not to make O unknowable. This obligation, O*, does not say of itself that, necessarily, X can know O only if X can know O*.

Let's move on to the argument (15-25). This argument may seem suspicious because it relies on the suspicious move from *can know* (in the case of (1-9)) to *know*. Note however that goal of (15-25) is different from the goal of (1-9). By formulating (1-9), a shirker tries to show that Access is false because my possibility of having O entails the existence of an infinity of obligations. Thus, contrary to our assumption that we can know O, as a shirker will be happy to assert, we cannot know O. Following (15-25), however, the argument is a little bit different. A shirker now argues as follows: suppose that I *actually*

¹⁵ This truth-condition is *knowability*.

¹⁶ More about this in the last part of this article.

know O , and *that I know O* is possible only if I know O^* . But, given (15-25), a shirker asserts that she cannot know O^* , because it would entail that O^{**} (an obligation to refrain from making O^* unknowable) must be knowable.

Consider how a shirker can argue for a premise of (15-25). A shirker here accepts that she actually knows that she has O . Then, she infers that if she knows that she has O , then she must know O^* . Finally, she concludes that since she cannot know O^* , she cannot know O . Suppose that a shirker says:

It seems that I have an obligation to pay rent (O). As a shirker, I don't want to pay rent, so I try to make O unknowable. But you are talking now that I can't make O unknowable because, for any obligation, I must refrain from making O unknowable (we call this principle (S)). So, I have an obligation O^ . Okay, let us suppose that your argument is sound. You say that I have O^* , thus I know now that I have O^* . But I still do not understand how O^* could be knowable to me. If (S) is in power (O^* is an instance of (S)), then I must refrain from making O^* unknowable, so I must have O^{**} , and from the fact that O^* is knowable, I conclude that O^{**} must be knowable to me too. It is false that O^* and O^{**} are essentially the same obligations, for in order to understand what is O^* , I must, in particular, understand the relation between O and O^* . But in order to understand O^{**} , I need to understand, in particular, the relation between O^{**} and O , and the relation of O^{**} to this relation. This entails an infinite regress of relations. From this, I conclude that your argument, according to which, for every obligation O^* , if I know O^{**} such that O^{**} is about O^* , then my knowledge about O^* and my knowledge about O^{**} are at the same epistemic level, is false.*

So, in this passage (which is a representation of (15-25) our shirker tries to say that even if she has actual knowledge (that is, not only a possibility to have this knowledge) that she has an obligation (she was told that she actually has this obligation), her *having an actual knowledge* that she has an obligation, as it follows from this argument, is not sufficient to state that she *can get* such knowledge. Thus, I take (15-25) to be a strengthened version of (1-9). By (15-25), it is not the case that Access is false because there is an infinite number of obligations, but because every single obligation O^* , O^{**} is essentially different from other obligations, and our attempt to understand what is (S), as a shirker asserts, entails an infinite regress. In Sorensen's original example, X can find out that she is obligated to donate for charity just by reading the will (of course, if the will asserts that X must donate for charity). By reading the will, X gets actual knowledge that she has an obligation to donate to charity. But in (15-25), a shirker tries to argue that a situation with instances of (S) (i.e. with O^* , O^{**} ...) is essentially different from the situation with her reading the will. She asserts that even if someone told her that she has O^* (for instance, an obligation to refrain from making her obligation to pay rent unknowable), the fact that she was told that she has O^* (thus she now knows that she has O^*) does not entail that she now actually knows that she has O^* (since she argues that it is impossible for her to understand what is O^* - in order to understand O^* , she argues, she must understand all instances of (S) (she concludes from (15-25) that this is not possible. Hence, by formulating (15-25), a shirker argues that not only her *possibility to know* an obligation is not sufficient to make this obligation knowable; she now argues that even her *having actual knowledge* that she has an obligation is not sufficient to assert that she knows that she has this obligation because, as she argues, this knowledge is inconsistent.

My own answer is as follows. We must agree that (15-25) poses a real problem for Access. In the discussion of (15-25), I argue that this argument is not sound, but my diagnosis relies on the idea that obligations are ethical, but not logical entities, - that is, we must not assert that we know O only if O is appropriately logically explained (see my argument for explanatory regress).

Regarding (WA), I raised several objections to this disjunctive proof. My first worry concerns the consistency of (WA) - given (C1), (that is, an assumption that our primary obligation A (for example, not to burn the will) is unknowable), (WA) entails that (S) - an obligation to refrain from making A unknowable - is both knowable and unknowable. My second worry is that (WA) entails a new infinite regress. (WA) asserts that if O_1 is not knowable, then O_2 must be knowable. O_2 is defined by (WA) as ($O_2 \vee O_1$).

But if O_2 is $(O_2 \vee O_1)$ then, by replacing O_2 with $(O_2 \vee O_1)$, we have that O_2 is $(O_2 \vee O_1) \vee O_1 \dots$). Our shirker might then ask: even if, as you say, O_2 – that is, $O_2 \vee O_1$ is knowable, what is a reason to think that $(O_2 \vee O_1) \vee O_1$ is knowable? And what is a reason to think that $(O_2 \vee O_1) \vee O_1 \vee O_1$ is knowable? A shirker can thus object that since O_2 entails a new regress, O_2 is not properly defined by (WA).

My third worry is this. By (WA), if O_1 is not knowable, O_2 is knowable. Suppose that O_2 is now knowable (thus I have O_2), and O_2 says that I have a disjunctive obligation $(O_2 \vee O_1)$. That is, (WA) accepts that

(D) The fact of my having O_2 grounds that $(O_2 \vee O_1)$

(D) is an instance of the principle, known in the literature about grounding as *Disjunctive Grounding*. According to this principle, if there is a fact P , then P fully¹⁷ grounds that either P or Q . Or, technically:

(DG) $P \rightarrow P < (P \vee Q)$

But (DG) is an invalid principle (see Fine ,2010, p.117 and Litland, 2015). In order to show that (DG) (and thus (D)) is not valid, consider the sentence (O_2) such that (O_2) is

(O_2) Either the fact of my having O_2 grounds that $(O_2 \vee O_1)$, or the sentence with label (O_2) is true¹⁸
 $O_2 < (O_2 \vee O_1) \vee \top \ulcorner O_2 \urcorner$

Note that (O_2) is an instance of (DG) because, by (WA), O_2 says of itself that it grounds a disjunction $(O_2 \vee O_1)$. The argument will be as follows. Suppose that it is the case that O_2 (i.e., $O_2 < (O_2 \vee O_1)$). Assume (DG). Assume the following plausible principles (more about these principles in Fine, 2010)

(Factivity)	$P < / < Q \rightarrow P, Q$
	(If P grounds Q (fully or partially), then both P and Q are the case)
(Truth-grounding)	$P \rightarrow P < \top \ulcorner P \urcorner$
	(If P , then the fact that P partially grounds the fact that P is true)
(Irreflexivity)	$(P < Q) \rightarrow \sim (Q < P)$
	(If P grounds Q , then it is not the case that Q grounds P)

We have then the following argument

(I) $O_2 < (O_2 \vee O_1)$	(Assumption)
(II) $(O_2 < (O_2 \vee O_1)) < (O_2 < (O_2 \vee O_1)) \vee \top \ulcorner O_2 \urcorner$	(I), (O_2)
(III) $(O_2 < (O_2 \vee O_1) \vee \top \ulcorner O_2 \urcorner) < \top \ulcorner O_2 \urcorner$	(II), (O_2), (Truth-grounding) ¹⁹
(IV) $\top \ulcorner O_2 \urcorner$	(III), (Factivity)
(V) $\top \ulcorner O_2 \urcorner < (O_2 < (O_2 \vee O_1)) \vee \top \ulcorner O_2 \urcorner$	(IV), (DG)

¹⁷ P fully grounds Q if nothing needs to be added to P in order to get a full ground for Q . We say that P partially grounds Q there is a plurality of facts Γ such that Γ , together with P , fully grounds Q .

¹⁸ Note that the sentence (O_2) is just the same assertion as the assertion of the obligation O_2 – that O_2 grounds that $O_2 \vee O_1$.

¹⁹ Given that, by assumption, our obligation O_2 is equivalent to $(O_2$ grounds that $O_2 \vee O_1)$, if we have that $(O_2$ grounds $(O_2 \vee O_1) \vee \top \ulcorner O_2 \urcorner)$ is true (from (II)), then it is the case that either $(O_2$ grounds $(O_2 \vee O_1))$ or $(\top \ulcorner O_2 \urcorner)$ is true. So, if O_2 is $(O_2$ grounds that $O_2 \vee O_1)$, in any case we have that $\top \ulcorner O_2 \urcorner$

(VI) $\top \supset O2 \supset \neg (O2 < (O2 \vee O1)) \vee \top \supset O2 \supset \neg$	(V), (Subsumption) ²⁰
(VII) $(O2 < (O2 \vee O1)) \vee \top \supset O2 \supset \neg < \top \supset O2 \supset \neg \& \top \supset O2 \supset \neg < (O2 < (O2 \vee O1)) \vee \top \supset O2 \supset \neg$	(III), (VI), &-intro
(VIII) \perp	(VII), (Irreflexivity)

Thus, given this argument, I conclude that (WA), as long as it relies on implausible principle – (DG) – entails a violation of Irreflexivity, and thus cannot be true. Also, from Fine's (2010) logic of ground we know the source of the problem with (DG). Since (WA) presupposes that the fact of my having O2 grounds the disjunctive fact that either O2 or O1, then O2 alone must be a ground for the whole disjunction. But this is not true, because O2 itself is partially grounded in this disjunction, (O2 \vee O1). Following this argument, we conclude that (WA) cannot be true.

Let us now proceed to the next important issue. In the third part of the article I have mentioned that (S), *from a certain perspective*, is an unprovable obligation, and if unprovability entails unknowability, then (S) is an unknowable obligation. An anonymous referee of the journal raised an objection that I gave no argument to think that unknowability entails unprovability, so it is not clear if (S) is unknowable.

Let me clarify my position. First, I do not want to assert that, for any P, unprovability of P entails unknowability of P *necessarily*. Suppose we have proof that P is unprovable. So we have proof that it is false that P is provable. If so, then from the principle that *what is provably false is provably impossible* we should conclude that it is impossible to prove P. Thus, if P is knowable only if provable, P is necessarily unknowable. Surely, this is not my position. I do not claim that (S) is necessarily unknowable. I claim that (S) is unknowable only *from a certain perspective* – if (S) is taken as a logical truth, and I have explained that there are some cases in which we do not need to prove (S) in order to have (S).

But what is a reason to think unprovability could entail unknowability? Consider the following argument

- (A) Suppose that Q is our best arithmetic system
- (B) Q cannot prove its own consistency (from Gödel's incompleteness theorem)
- (C) Thus, we do not know if Q is consistent.

My reason to think that (S) is unknowable is similar to the reason to think that Q's consistency is unknowable. (S) is universally quantified obligation – an obligation about any obligation, including (S). If so, there must be a sufficiently strong background theory of obligation in which (S) could prove its own consistency. But from Gödel's lesson, we know that (S) is not able to do that. Thus, *from that perspective*, we conclude that (S)'s consistency is unprovable. We do not know if (S) is consistent, thus we do not know if there is such consistent obligation as (S).

Let us assume however that (S) is not logical, but an ethical statement. First, such ethical obligations as [You must respect your parents] need not to logically proved. If I have an obligation to respect my parents, and I know that this is my obligation, then I know that I must respect my parents. Consider again a shirker's case with the obligation not to burn the will. A shirker asserts that she does not know if she has such obligation as to donate to charity, because her possibility to get the knowledge that she must donate to charity (assume that the will asserts that a shirker must donate to charity) depends on contingent fact – her reading the will. But even if a shirker disagrees that she has an obligation to donate to charity, she must agree that she has (ethical) obligations. Suppose, that a shirker claims

(S1) I do not know that I have obligations

(S2) Thus, by (Access), I do not have obligations

²⁰ Subsumption is here a principle according to which every instance of full ground is an instance of partial ground.

Despite the fact that a shirker, of course, could claim that she does not know if she has an obligation to donate to charity, because her having this obligation is dependent on the contingent fact of her reading the will (it is possible for her not to read the will, and thus to make her obligation to donate to charity unknowable), a shirker cannot assert that she doesn't have obligations because the fact that she has obligations is unknowable to her. Unlike the case with her obligation to donate to charity, the fact that our shirker has an obligation does not depend on any contingent fact. Thus, (S1-S2) is a wrong argument.

But consider the situation with (S). On what contingent fact the obligation that I must refrain from making my obligation unknowable is dependent? Well, a shirker might respond that my having (S) is dependent on the fact of my reading of Sider's article on this problem (or any other paper in which this problem is explained). Thus, only a few people know that there is such obligation as (S). Sorensen (1995), however, disagrees with this answer of the shirker. By Sorensen, we do not need to have a thorough knowledge of ethics and philosophy in order to understand that we have some obligations.

My proposal is this. For any particular obligation *O*, a shirker, of course, can object that the fact that she can know *O* is unknowable to her. However, she cannot assert that she doesn't have obligations at all. So, she agrees that she has obligations. But if she understands that she has obligations, then she must understand, of course, that she must not make her obligations unknowable, for any particular obligation – due to the fact that, if I have an obligation (say *O*), I must do *O*, and then I must refrain from making *O* unknowable. The idea behind this argument is that (S) is a necessary part of any obligation, since it belongs to the essence of obligations that I must fulfill them, and thus I must refrain from making them unknowable. Of course, such particular obligations as to donate for charity, not to burn the will, to pay rent etc. could be unknowable. But (S) – our universal obligation – cannot be unknowable. Since (S), as I argue, is a necessary part of the very notion of obligation, then if someone claims that she doesn't know if she has (S), it follows from this claim that she doesn't know if she has obligations. Given that the claim that I don't know if I have obligations *at all* is false, I conclude, by contraposition, that a claim that someone could not know that she has (S) is also false.

So, according to my proposal, if we have obligations, then we have (S), because our having (S) follows from the nature of obligations. But how we can understand what is (S)? Suppose that (S) is a logical statement. Then (S) must be provable. But it seems that (S) is not logically provable. Also, a shirker can ask: if (Access) is true, then we have (S) only if we have (SS) (an obligation to refrain from making (S) unknowable), we have (SS) only if we have (SSS), and so on. We must explain them all to prove that we have (S) (see my argument for explanatory regress from the 2-nd part of the paper). Thus, it generates an infinite regress of explanation. As long as I agree that this shirker's counter-argument is at least partially right, we must not think that (S) is a logical statement.

Suppose now that (S) is an ethical statement. If so, we do not need then to explain (SS), (SSS) etc. in order to have (S). A shirker argues that it is not knowable to her that she has (S). But, according to my argument, she cannot claim that (S) is unknowable to her. She cannot claim that (S) is unknowable to her for the same reason why she cannot claim that the fact that she has obligations is unknowable to her. If she has obligations, then she has (S). *In this respect*, we can assert that (S) is knowable for us. If (S) is ethical, we don't have (SS), (SSS), and so on. Thus, we have no infinite regress, so we do not have our main problem with the explanation of (S), described above.

References

- ARMSTRONG, D.M. 1974. Infinite Regress Arguments and the Problem of Universals. *Australasian Journal of Philosophy*, **52**(3): 191–201.
- BERKER, S. 2019. The Explanatory Ambitions of Moral Principles. *NOUS*, **43**(4): 904-936.
- FINE, K. 2010. Some Puzzles of Ground. *Notre Dame Journal of Formal Logic*, **51**(1): 97–118.

- LITLAND, J. E. 2015. Grounding, explanation, and the limit of internality. *Philosophical Review* **124**: 481–532.
- ROSEN, G. 2017. What is a Moral Law? *Oxford Studies in Metaethics*, **12**: 135-159.
- SHER, G. 2009. *Who knew? Responsibility without awareness*. Oxford, Oxford University Press.
- SIDER, T. 1995. Sorensen on unknowable obligations. *Utilitas*, **7**: 273–279.
- SORENSEN, R.A. 1995. Unknowable obligations. *Utilitas*, **7**: 247–271.
- WIELAND, J.W. 2014a. Access and the shirker problem. *American Philosophical Quarterly*, **52**(3): 289-300.
- WIELAND, J.W. 2014b. *Infinite regress arguments*. Amsterdam, Springer.
- ZIMMERMAN, M.J. 1996. *The concept of moral obligation*. New York, Cambridge University Press.

Submitted on December 8, 2021.

Accepted on March 16, 2022.