

On the Supposed Incoherence of Obligations to Oneself

Janis David Schaab

An influential argument against the possibility of obligations to oneself states that the very notion of such obligations is incoherent: If there were such obligations, we could release ourselves from them; yet, releasing oneself from an obligation is impossible. I challenge this argument by arguing against the premise that it is impossible to release oneself from an obligation. I point out that this premise assumes that if it were possible to release oneself from an obligation, it would be impossible to *violate* that obligation. I point out that there are two interpretations of this assumption, one conceptual and one psychological. I argue that, on both interpretations, the assumption is false—at least according to independently plausible accounts of obligations to oneself and of what it means to waive an obligation. My arguments paint a picture of obligations to oneself that not only challenges the Incoherence Argument, but also illuminates these obligations' relationship to other parts of the moral domain.

Keywords: Obligations to Oneself; Moral Obligation; Ethics; Metaethics; Practical Reason; Second Person

1. Introduction

According to an influential argument, the very notion of obligations to oneself is incoherent. If there were such obligations, the argument goes, we could release ourselves from them; but releasing oneself from an obligation is impossible. We may refer to this as the *Incoherence Argument*.

In this essay, I argue that the Incoherence Argument fails. In particular, I challenge the premise that it is impossible to release oneself from an obligation. I observe that this premise rests on

the assumption that if it were possible to release oneself from an obligation, it would be impossible to *violate* that obligation, because there would be no difference between acting against the obligation and releasing oneself from it. I refer to this assumption as the *Inviolability Thesis*. I contend that, in an argument against the possibility of obligations to oneself, one cannot simply rely on the Inviolability Thesis without justifying it and, what is more, this thesis is harder to justify than it might appear.

I start by providing an analysis of the Incoherence Argument's individual premises (Section 2). I contend that this argument inadmissibly relies on the Inviolability Thesis without justifying it (Section 3). Subsequently, I point out that there are two interpretations of the Inviolability Thesis, one conceptual and one psychological. I note that if the conceptual interpretation of the Inviolability were true, it would indeed support the claim that it is impossible to release oneself from an obligation. However, I argue that this interpretation of the Inviolability Thesis is false—at least according to one independently plausible conception of releasing someone from an obligation (Section 4). I then point out that the psychological interpretation of the Inviolability Thesis, even if it were true, would not obviously support the claim that releasing oneself from an obligation is impossible. In any case, I argue that this interpretation is also false—at least according to one independently plausible conception of obligations to oneself (Section 5). Finally, I point out that my arguments have implications, not only for the coherence of obligations to oneself, but also for their plausibility and their relation to the rest of morality (Section 6).

2. The Incoherence Argument

Early formulations of the Incoherence Argument can be found in Hobbes and Kant [Hobbes 1651: 184; Kant 1797: 6:417]. A more schematic statement was provided by Marcus Singer [1959: 1963]. Singer's argument has three premises [1959: 203]:¹

- (1) If A has a duty to B, then B has a right against A;
- (2) if B has a right against A, he can give it up and release A from the obligation; and
- (3) no one can release himself from an obligation.

From these three premises, Singer concludes that obligations to oneself are impossible because 'the idea involves a contradiction' [1959: 202].

My aim in this paper is to reject the Incoherence Argument by challenging premise (3). To explain why I take this approach, rather than challenge premises (1) or (2), let me briefly outline what these premises entail.

First, note that premises (1) and (2) assume that obligations to oneself are literally *to oneself*. That is, they are 'directed' or 'bipolar' obligations (see May [2015] and Thompson [2004], respectively). A's obligation to ϕ is directed if there is some particular person, B, to whom A owes fulfilment of the obligation so that, by failing to ϕ , A does not merely act wrongly but *wrongs* B. Accordingly, one could challenge the Incoherence Argument by arguing that obligations to oneself are not obligations whose fulfilment we owe to ourselves—instead, they are 'self-regarding' in the sense that they primarily benefit or concern the agent [Singer 1959: 204]. However, Singer plausibly remarks that merely self-regarding obligations are not obligations to oneself in the 'literal sense' [1959: 202]. To illustrate, suppose a firefighter has an obligation to stay fit, which is grounded in the well-being of society. While this obligation

¹ Following Singer, I use 'duty' and 'obligation' interchangeably.

primarily concerns or regards the firefighter, it need not be an obligation to herself, in any interesting sense.

Second, note that premise (1) asserts the correlativity of rights and directed obligations, that is, the thesis that A has an obligation to ϕ to B if and only if B has a *right* to A's ϕ -ing (also see Hohfeld [1917], Kant [1797: 6:221-29]). One could therefore try to undermine the Incoherence Argument by rejecting the correlativity of rights and directed obligations. However, the Incoherence Argument would not be significantly damaged by this. For one could omit premise (1) and restate premise (2) in terms of directed obligations rather than rights, as follows:

(2*) If A has an obligation to B, B can release A from that obligation.

The Incoherence Argument can proceed from premises (2*) and (3) without having to compromise its appeal.

Finally, premise (2)—or (2*)—assumes that all directed obligations are *waivable* by the person to whom they are owed. It follows that one could challenge the Incoherence Argument by arguing that some directed obligations are *unwaivable*, and that some of these are owed to oneself [Cholbi 2018: 111-12; Hills 2003; Timmermann 2006: 516]. However, the only obligations to oneself that could be vindicated in this way are those that are plausibly unwaivable, such as obligations of basic respect for our humanity or obligations to show a minimal level of regard for our well-being [Schofield 2015: 516]. All those obligations to oneself that can plausibly be waived, such as those that might be taken to arise from promises to oneself, would remain vulnerable to the Incoherence Argument [Hills 2003: 132-34].

As we have seen, the appeal of challenging the Incoherence Argument by rejecting premises (1) or (2) is rather limited: in this way, we can at best defend the coherence of unwaivable obligations to oneself or of merely self-regarding obligations. In the remainder of this paper, I therefore grant the truth of these premises and attack premise (3) instead. Consequently, I use

‘obligations to oneself’ exclusively to refer to obligations that are owed to and waivable by oneself.

3. The Inviolability Thesis

Premise (3) states that it is impossible to release oneself from an obligation. The underlying intuition seems to be that, if I can release myself from an obligation, I am not truly obligated. But is that intuition correct? Could we not say that, even if I can release myself from an obligation, I am under that obligation unless and until I *do* release myself [Habib 2009: 549-556; Hill 1991: 147; Muñoz forthcoming; Rosati 2011: 134-35]?

Indeed, isn't the notion of an obligation from which we can release ourselves familiar from our moral phenomenology? Suppose I want to buy a new laptop, so I promise myself to save up some money. A few days later, I am informed that my salary will be raised. I can now afford to buy the new laptop right away. It seems that now, if not before, I can release myself from my promise to save up money. It is not obvious that the mere *possibility* of self-release nullifies my obligation to save up money.

To the contrary, as long as I do not release myself from my promise, I seem to stand in a relation to myself that shares some of the characteristic features of moral obligation. To see this, consider that moral obligations characteristically exhibit what Stephen Darwall calls ‘second-personal’ form: their violation, if unexcused, intrinsically warrants holding the agent accountable by adopting ‘reactive attitudes’ like blame and resentment towards her [Darwall 2006: ch. 5] (also see Strawson [1962]). Such attitudes implicitly address the agent, demanding that she fulfil her obligations, in a distinctive, person-to-person manner. As several authors note, failure to live up to a self-promise from which one has not released oneself typically gives rise to feelings of guilt, remorse and self-blame, and these are naturally interpreted as reactive

attitudes of the sort that Darwall has in mind [Cholbi 2015: 853; 2018: 112; Rosati 2011: 125, 144].²

By blaming myself for failing to save up money for a new laptop, for example, I seem to hold myself accountable, ‘stepping outside of myself’ to address myself second-personally.

However, these phenomenological considerations are insufficient to undermine premise (3) of the Incoherence Argument.³ This premise is motivated by doubts about the *logical possibility* of

² This might seem puzzling. Doesn’t the second-personal structure of obligations imply that they are inherently *inter*-personal? This concern rests on a misunderstanding, however. Obligations are second-personal in the sense that they are based on relations of authority and accountability, which always involve addressing a person *as* a person, i.e., as a ‘you’. Such relations need not be limited to contexts involving several people. In Darwall’s words, “second person” does not entail “second party” [2010a: 217].

³ Of the authors who question the truth of premise (3) of the Incoherence Argument, two (that I know of) provide arguments which go beyond mere phenomenological considerations. However, both authors’ arguments are highly revisionist and therefore do not meet the challenge posed by premise (3) head-on. First, Tim Oakley challenges the claim that one cannot release oneself from an obligation by adopting a revisionist understanding of ‘release’ [2017]. In particular, Oakley supposes that ‘[a] person releases herself from a duty when she does something that either adds a new morally relevant feature to the situation, or changes the weight or stringency of a pre-existing feature, and by doing this re-balances the moral considerations with the result that she no longer has an all-things-considered duty that she previously had’ [2017: 75]. By contrast, I suspect that most authors writing on the topic of directed obligations, Singer included, suppose that B’s releasing A from an obligation owed to B involves the exercise of a distinct *normative power* (also see Section 4), whose function is not to tip the scales so that A no longer has an all-things-considered obligation to ϕ , but to *undercut* or *cancel* whatever obligation A owes to B—be it all-things-considered or merely *pro tanto*. Second, Daniel Muñoz defends the possibility of releasing oneself from an obligation by abandoning the assumption that obligations to ourselves give us normative reasons [forthcoming]. Contrary to Muñoz, I continue to assume that obligations to oneself are normative-reason giving, as do most participants in the debate. However, my arguments in this paper need not be incompatible with Muñoz’s, for they can be interpreted as defending a different subset of our

releasing oneself from an obligation, which cannot be dispelled by the phenomenological plausibility of self-release. This is indicated by Singer's remark that an obligation can be *to* a person even if that person is currently unable to waive it, provided that the *thought* of this person waiving the obligation involves no 'self-contradiction' [1963: 141]. Evidently, Singer believes that the thought of the obligated agent *herself* waiving the obligation does involve such a self-contradiction.

Why does Singer believe this? An important clue is given by what he says about self-promising [1959: 203]:

To promise oneself to do something just *is* to be strongly resolved to do it, and if one were to change his mind and not do what he intended he would not have broken any promise.

This passage suggests that the problem with the notion of an obligation from which I can release myself is that there is no way of *violating* such an obligation. That is, the obligation will never render me *at fault* for failing to ϕ . After all, I will be obligated to ϕ only if I ϕ ; therefore, if I do not ϕ , then I am also not obligated to ϕ . We may refer to the claim that an obligation from which one can release oneself cannot be violated as the *Inviolability Thesis*.

Now, how exactly does the Inviolability Thesis support premise (3) of the Incoherence Argument (the claim that one cannot release oneself from an obligation)? Presumably, Singer relies on what Douglas Lavin calls the *Error Constraint*, according to which 'an agent is subject to a principle only if the agent can go wrong in respect of it' [2004: 425, italics deleted]. The intuitive idea is that a principle cannot be *normative* for an agent unless the agent can violate it. Presumably, part of what makes this idea so intuitive is that, if we conform to a principle no

obligations to ourselves (those arising from self-promises and practical identities, as opposed to those we have by virtue of our basic humanity).

matter what we do, our actions cannot be *guided* by that principle. Since obligations to oneself are naturally interpreted as normative principles of some sort, the conjunction of the Inviolability Thesis and the Error Constraint supports the conclusion that one cannot release oneself from an obligation.

However, Singer does not provide an argument for the Inviolability Thesis. Accordingly, the Incoherence Argument is *incomplete* at best. For even if the implicit appeal to the Error Constraint is unproblematic (which is questionable, as I argue in Section 5), all that Singer's remarks show is that *if* the Inviolability Thesis is true, obligations to oneself are incoherent. Since the falsehood of the Inviolability Thesis is thus a necessary condition for the coherence of obligations to oneself, defenders of such obligations are committed to rejecting the Inviolability Thesis. Therefore, if the Incoherence Argument is to give us any reason to doubt the possibility of obligations to oneself, an independent case for the Inviolability Thesis must be provided.

In the remainder of this essay, I aim to show that motivating the Inviolability Thesis is more difficult than one might expect. I observe that there are two distinct interpretations of the Inviolability Thesis, one conceptual and one psychological. I refer to these as the *Conceptual Inviolability Thesis* and the *Psychological Inviolability Thesis*, respectively. I note that the two theses correspond to two distinct readings of the Error Constraint. I argue that, while the Conceptual Inviolability Thesis can be paired with a compelling, minimal reading of the Error Constraint to support premise (3) of the Incoherence Argument, it turns out to be false—at least according to an independently plausible account of what it means to waive an obligation. Subsequently, I argue that the Psychological Inviolability Thesis not only needs to be paired with a stronger, less compelling reading of the Error Constraint in order to support premise (3) but also turns out to be false—at least according to one independently plausible conception of obligations to oneself.

4. The Conceptual Inviolability Thesis

How we interpret the Inviolability Thesis depends on what kind of *possibility* we take it to invoke. If we take it to be concerned with *conceptual* possibility, we are interpreting the Inviolability Thesis as the claim that it is conceptually impossible to violate an obligation from which one can release oneself—or, equivalently, that acting against such an obligation conceptually entails releasing oneself from it. This interpretation is what I refer to as the *Conceptual Inviolability Thesis*.

Recall that the argument from the Inviolability Thesis to the conclusion that it is impossible to release oneself from an obligation depends on the Error Constraint, which states that ‘an agent is subject to a principle only if the agent can go wrong in respect of it’ [Lavin 2004: 425, italics deleted]. Now, just as the meaning of the Inviolability Thesis depends on which kind of possibility we take it to invoke, the meaning of the Error Constraint depends on what sense of ‘can’ we take it to employ. Fortunately for advocates of the Incoherence Argument, if they employ the Conceptual Inviolability Thesis, they can rely on a minimal interpretation of the Error Constraint which, following Lavin, we may refer to as the *Logical Interpretation*. It states that ‘an agent is subject to a principle only if there is some kind of action such that if the agent did it she would thereby violate the principle’ [Lavin 2004: 426, italics deleted]. Crucially, obligations from which one can release oneself violate the Logical Interpretation if acting contrary to such an obligation conceptually entails releasing oneself from it. This, in turn, is the case if the Conceptual Inviolability Thesis is true.

The Logical Interpretation of the Error Constraint is overwhelmingly compelling. Accordingly, if the Conceptual Inviolability Thesis were true, we would have to conclude that there are no obligations from which we can release ourselves. However, I now want to argue that according to at least one independently plausible and widely accepted conception of what it is to release someone from an obligation, the Conceptual Inviolability Thesis is false. For on this conception,

there is a conceptual gap between releasing oneself from an obligation and merely acting against it.

The conception of releasing oneself from an obligation that I have in mind holds that when we waive an obligation, we exercise a *normative power* or *authority* (see, for instance, Dougherty [2015], Hurd [1996], Owens [2011]). A normative power is an ability to effect a normative change—that is, a change in what duties, rights, or reasons someone has—*merely* by performing a certain act. Accordingly, on this understanding, the act of waiving an obligation, *by itself*, terminates the obligation.

Importantly, waiving is not here understood as terminating obligations by ‘triggering’ or ‘enabling’ some independent normative fact or principle which states that an obligation ceases to exist if and when it is waived. If this were what was meant by ‘normative power’, then by physically harming someone, thus incurring a new obligation (to apologise, compensate, etc.) to them, I would be exercising a normative power. The difference between such an act and a genuine exercise of normative power is that the latter is the ‘source’ of the normative change that results from it; that is, it is ‘that in virtue of which’ the normative situation changes [Chang 2009: 243, italics deleted].⁴

Much of the philosophical discussion surrounding the normative power of waiving focuses on what *kind* of act counts as an exercise of this power: is it exhausted by some mental operation, such as the formation of a propositional attitude, or does it involve an attempt to communicate (see Hurd [1996] and Dougherty [2015], respectively)? The answer to this question is not important for my purposes, however. I want to claim that, no matter what *kind* of act waiving is, it cannot be conceptualised *as* an act of waiving independently from the normative power of

⁴ Note that the term ‘normative power’ is not always used in this specific sense (see, e.g., MacCormick and Raz [1972], Raz [1985; 1999]).

which it is an exercise. In other words, the act-description ‘waiving an obligation’ is conceptually linked to this normative power. If this is true, it immediately follows that there is a conceptual gap between releasing oneself from an obligation and acting contrary to it. After all, to act contrary to an obligation to ϕ one does not have to fulfil an act-description that is conceptually linked to the normative power of waiving; it is simply not to ϕ .

My claim is motivated by the observation that, if waiving is an exercise of a normative power, no one can perform the act of waiving an obligation unless they have this normative power. I may ‘go through the motions’, by engaging in whatever mental or communicative activities are involved in waiving; but unless I have the power to waive, my activities will not amount to an act of *waiving*.⁵ We might say that, independently of the power to waive, nothing *counts* as a waiver. The *concept* of waiving an obligation cannot be applied outside the domain delineated by this power.

This seems independently plausible. If Tim promises Florence that he will walk her dog on Friday, thereby incurring an obligation to Florence, then only Florence can waive this obligation. It is not just that only Florence *may* waive the obligation; rather, only Florence *can* perform the act of waiving with respect to it.

This is so even if both Tim and a third party, Louise, falsely believe that Tim’s obligation is owed to Louise. Suppose Tim writes a letter, saying ‘I, Tim, hereby promise to walk your dog on Friday’, then blindfolds himself and puts the letter into a random mailbox in the hallway of his apartment complex. Tim thereby intends to incur a promissory obligation to the person who

⁵ An understanding of ‘normative power’ that rejects this claim is advanced by Joseph Raz [1999: 100]. We need not worry about Raz’s position here since we are presupposing a different notion of normative powers. But it is worth noting that, as Darwall points out, Raz’s position has the implausible implication that an alarm clock whose ringing we have reason to treat as giving us certain kinds of reasons has normative authority over us [Darwall 2010b: 271-72].

will find the letter in their mailbox, whoever that may be (Tim lives in an all-dog-owner complex). The next day, Tim looks at the mailboxes and discovers that the mailbox he *believes* he put the letter into (he seems to remember it was the one in the top-right corner) belongs to Louise. Unbeknownst to Tim, he actually put the letter into Florence's mailbox, which is right next to Louise's. Tim calls Louise, who has not checked her mailbox yet, and says, 'You are the one to whom I incurred a promissory obligation to walk their dog on Friday'. Louise believes Tim and replies, with the full intention of waiving the obligation that Tim incurred through his letter, 'I hereby release you from that obligation'.

I take it that Louise does not only not waive Tim's obligation (which is to Florence, not to Louise), but she also does not perform an *act* of waiving at all. And the reason is not just that Louise does not *succeed* at terminating any obligation. Arguably, Louise *could* perform a waiver by saying 'I hereby waive *any* promissory obligations that you have towards me' even Tim has no such obligations. By contrast, what Louise actually does—attempting to waive the *specific* obligation that Tim had incurred by his mailbox promise—simply fails to constitute a waiver. The reason is that the power to waive this obligation is vested in Florence, not Louise.

My point here parallels John Rawls's observations about the rules of practices. Rawls notes that the rules of practices are 'logically prior' to the particular cases that fall under them [1955: 25]. The idea is that practices are not merely regulated but defined or constituted by their rules. For example, the game of chess is defined by the set of rules by which it is played. Accordingly, if these rules did not exist, nothing would *count* as playing chess. By the same token, nothing would count as performing any of the actions, or *moves*, that form part of the practice of playing chess. You might move about pieces of wood on a chequered board, but none of your motions would count as castling, pinning or checkmating. In other words, they would not be *describable* as these acts because the relevant *concepts* would not be available. In this way, the rules of practices open up their own, distinct conceptual domains. In Rawls's words, '[w]e may think of the rules of a practice as defining offices, moves, and offenses' [ibid.].

It follows that, unless one follows the rules of a practice to some extent, one is not engaged in the practice at all [1955: 26]. After all, the only way of performing the act of, say, castling is to follow the relevant chess-rule. If one moves around the chess pieces without following this rule, one's motions will not count as castling. Similarly, if one's pattern of actions over time does not follow the rules of chess, one will not count as playing chess.

I do not claim that the act of waiving an obligation depends on the existence of a *social* practice; that is, a practice that is contingently adopted by a concrete historical community. Nonetheless, we might say that waiving is logically posterior to a practice rule in a technical, Rawlsian sense. In particular, no one's act can be conceptualised as a waiver outside the normative context of the power to waive. This power relates to the act of waiving as the rules of chess relate to the moves of chess: it not only regulates but constitutes this act. That is, it determines not only when, how and by whom an obligation *may* be waived but also when, how and by whom it *can* be waived. Accordingly, the act-description 'waiving an obligation' belongs to a distinct conceptual domain. Indeed, we might even say that the power to waive defines an *office* which alone is capable of issuing waivers. It is only *qua* holder of this office, and hence within the conceptual domain created by the normative power to waive, that anyone can perform the act described as a 'waiver'.

Therefore, if waiving is an exercise of a normative power, then releasing oneself from an obligation is fundamentally different, conceptually, from merely failing to comply with it. Thus, the conception of waivers outlined in this section implies that the Conceptual Inviolability Thesis is false: acting against an obligation from which one can release oneself does *not* conceptually entail releasing oneself from it. Accordingly, unless some compelling reason is provided to reject this conception, defenders of obligations to oneself need not worry about the Incoherence Argument in so far as it invokes the Conceptual Inviolability Thesis.

5. The Psychological Inviolability Thesis

Some advocates of the Incoherence Argument might not be satisfied with my argument in the previous section. They might say: 'Maybe it is *conceptually* possible to violate an obligation from which one can release oneself. But no human being is able to violate such an obligation.' This reasoning invokes the second, psychological interpretation of the Inviolability Thesis, which we may call the *Psychological Inviolability Thesis*. It states that it is *psychologically* impossible to violate an obligation from which one can release oneself. I take the underlying idea to be that any motive for not ϕ -ing is equally a motive for releasing oneself from an obligation to ϕ and *vice versa* so that, whenever my motivation is sufficient to make me not ϕ , it will also be sufficient to make me release myself from any obligation to ϕ .

Unlike the Conceptual Inviolability Thesis, the Psychological Inviolability Thesis will not support premise (3) of the Incoherence Argument if it is merely conjoined with the *Logical Interpretation* of the Error Constraint. After all, the psychological impossibility of violating an obligation is compatible with there being an action that would violate it. To support the Incoherence Argument, then, the Psychological Inviolability Thesis must be conjoined with a stronger interpretation of the Error Constraint. In particular, it must be conjoined with what, following Lavin, we might call the *Imperative Interpretation*: 'an agent is subject to a principle only if there is some kind of action such that if the agent did it she would thereby violate the principle *and it is possible for the agent to do it*' [2004: 427, italics altered]. In other words, a principle is not normative for an agent unless she is 'in one way or another imperfectly hooked up with it' [Lavin 2004: 427].

Lavin argues that accepting the Imperative Interpretation comes with significant theoretical costs, including the acceptance of the conception of freedom known as 'the liberty of indifference' and the denial of constitutivism about normativity [2004: 446-57]. I cannot provide a detailed assessment of the Imperative Interpretation here, although I dare say that the

Imperative Interpretation is much less compelling than the Logical Interpretation of the Error Constraint.⁶ In any case, I now want to argue that at least one independently plausible conception of obligations to oneself implies that the Psychological Inviolability Thesis is false.

The conception that I have in mind has been developed, in a series of recent articles, by Paul Schofield [2015, 2019]. Schofield argues that, as time passes, we acquire different ‘perspectives’, where a perspective is understood as ‘a point of view from which one perceives, or feels emotions, or has sensations, or judges a proposition to be true, or wills some particular action, and so on’ [2015: 517]. Schofield’s key claim about obligations to oneself is that, while the *person* who can waive such an obligation is identical to the person who is bound by it, the *perspective* from which she can waive it is not identical to the perspective from which she is bound by it. Crucially, what individuates perspectives and enables them to play this role in Schofield’s conception is *not* the temporal distance between them, as such. Rather, what is important is that different perspectives come with different ‘ends and interests’ [Schofield 2015: 518].

It is worth illustrating this by an example. Suppose I have an obligation to myself to advance my career because of the benefits, like stability and financial security, that it will bring me later in life. On Schofield’s conception, while this obligation is waivable by *me*, I cannot waive it *now*. Plausibly, the perspective from which I can waive the obligation is one that I will come to occupy in the future. The reason is not that the relevant future perspective corresponds to a metaphysically distinct ‘temporal part’ of myself [Schofield 2015: 516-17]. Instead, the reason is that my obligation to advance my career is in an important sense underpinned by ends and interests that I have yet to acquire. It is my future interests in stability and financial security

⁶ The Imperative Interpretation might be more plausible as a constraint on *obligations* than on normative principles *in general*. After all, obligations might be thought to *constrain* or *necessitate* an agent, by commanding actions that she might otherwise not be motivated to perform [Kant 1785: 4:413] (also see Lavin [2004: 437-41]).

that explain why I am obligated to advance my career. And, by hypothesis, I do not have these interests now.

If it is plausible that the relevant perspectives are individuated by their accompanying ends and interests in this way, then, I submit, it is equally plausible that different perspectives are also linked to distinct sets of *motives*. Hence, supposing that I have an obligation to myself to advance my career, and that this obligation can only be waived from my future perspective, it follows that my present motive to, say, visit a music festival is not a motive from which I can release myself from my obligation to advance my career. This motive is simply attached to the wrong perspective to make me release myself from the obligation. Consequently, if I act on this motive and visit a music festival instead of advancing my career, I violate my obligation to myself.

In this way, Schofield's proposal implies that the Psychological Inviolability Thesis is false. However, there is an important limitation. For Schofield, it is crucial that we occupy the different perspectives on our obligations *at different times*, so that we are never bound by an obligation *and* in a position to waive it simultaneously [2015: 521; 2019: 223]. To this extent, Schofield's proposal is true to the spirit of premise (3) of the Incoherence Argument. Fortunately, there is a natural way to modify Schofield's proposal so that it allows for the possibility of being bound by an obligation and in a position to waive it at the same time.

Schofield himself considers the possibility that there need not be any temporal distance between the perspective from which we are obligated and the perspective from which we can release ourselves [2015: 520-521]. After all, as Christine Korsgaard argues, at any given time we can subscribe to several different conceptions of our 'practical identity', which are conceptions of ourselves as an agent that are connected to different sets of values, goals, and norms [1996: 101]. For example, I can simultaneously have the practical identities of *philosopher* and *athlete*. My ends, interests, and motives *qua* athlete are plausibly different from my ends, interests, and

motives *qua* philosopher. By virtue of my practical identity as a philosopher, I might have an interest in, and a motive for, continuing to work on my paper without interruption. By virtue of my practical identity as an athlete, however, I might have an interest in, and a motive for, going to the gym. Therefore, if I have an obligation to myself to go to the gym today, and this obligation can be waived only from my perspective as an athlete, then by continuing to work on my paper, I *violate* the obligation.⁷ After all, my motive to act in this way is not a motive on the basis of which I can release myself from the obligation.

Schofield ultimately rejects the possibility that the tensions between our practical identities suffice to sustain waivable obligations to oneself.⁸ He writes [2015: 521]:

[A]t this point we run up against [the Incoherence Argument]. A woman who has two practical identities simultaneously would be in the position, one might think, to waive whatever apparent duties are generated from the perspective of either of those identities. [...] If this is right, then the putative duty would fail to bind her in the way characteristic of moral duties—that is to say, it would be no duty at all.

We might paraphrase Schofield's complaint as follows: since access to the perspectives corresponding to our practical identities is not restricted to specific points in time, we will release ourselves from any obligation to ϕ that is generated by these perspectives as soon as our motives are strong enough to make us not ϕ —regardless of which perspective these motives are attached to. In particular, Schofield seems to think that I will release myself from my obligation to go to the gym today (from my perspective as the child of my parents), as soon as my motive

⁷ How might one incur such an obligation? My tentative suggestion is that we adopt practical identities, or incur obligations to our existing ones, by making a certain type of *commitment* (see my [2019]).

⁸ Interestingly, Schofield thinks that our practical identities might give rise to *unwaivable* obligations to oneself [2019].

to continue working on my paper is strong enough to make me act against that obligation, although this motive belongs to a different perspective (my perspective as a philosopher).

However, if this is Schofield's complaint, it is inconsistent with his own conception of obligations to oneself. After all, on this conception, obligations to ourselves can only be waived from a certain perspective. I can only waive my obligation to go to the gym from my perspective as an athlete. But from within this perspective, my motive to continue working on my paper instead of going to the gym is not available. It is simply not a motive that I, *qua* athlete, have. Accordingly, I cannot waive my obligation from this motive.

The ultimate aim of Schofield's conception of obligations to oneself is to show that, just as utilitarianism 'fails to take seriously the distinction between persons' by lumping together the interests of distinct individuals in an impersonal utility calculation, we fail to take seriously the distinction between a person's various perspectives if we regard a person's relation to herself as amoral, a matter of mere prudence [Schofield 2015: 523, citing Rawls 1999: 24]. In other words, Schofield wants to show that it is a mistake to assume that *moral* claims are generated only by the 'perspective of a person as a whole', and not by the perspectives corresponding to various sub-sets of a person's ends, interests and motives [ibid.]. Yet, by suggesting that I can waive my obligation to go to the gym from a motive that I only have from within my perspective as a philosopher, Schofield himself makes this very mistake. That is, he assumes that the motives attached to my practical identity as an athlete, as well as those attached to my practical identity as a philosopher, are available to me from an all-encompassing meta-perspective, as if they did not belong to distinct perspectives at all.

Schofield could respond that perspectives tied to specific times in our lives are importantly different from perspectives tied to synchronic practical identities. But he would thereby effectively abandon his conception of obligations to oneself. As we have seen, it is Schofield's notion of *perspectives*, individuated by their accompanying ends and interests (and motives),

that is supposed to help us make sense of obligations to oneself. If all the philosophical work were instead done by the *temporal distance* between these perspectives, Schofield's proposal would seem to collapse into the view that obligations to ourselves are owed to metaphysically distinct temporal *parts* of ourselves after all (*pace* Schofield [2015: 516-17]). Schofield would thereby also undermine his plausible claim that it is persons' distinct perspectives, rather than their metaphysical separateness, that underpins our obligations to *others* as well [2015: 517-19].

Thus, Schofield provides no principled reason for restricting his conception of obligations to oneself to cases of perspectives separated by time. To the contrary, his conception of obligations to oneself, as well as the philosophical outlook that motivates this conception, gives us principled reason to extend it to perspectives attached to practical identities.

Before I conclude this section, two points of clarification are in order. First, my amended version of Schofield's conception does not deny that my motive to keep working on my paper will make me *want to* release myself from this obligation. What this conception denies is that I *can* release myself on the basis of that motive. The reason is that the motive in question is not available from within the perspective from which the obligation can be waived. This indicates that Schofield's conception does not refute the Psychological Inviolability Thesis on *psychological* grounds. Instead, it refutes this thesis by making the *metaphysical* claim that only actions performed from a certain perspective *count* as waiving these obligations. This metaphysical claim contradicts the Psychological Inviolability Thesis because it implies that sometimes, no matter how *strong* our motives to act against an obligation to ourselves are, we may still not release ourselves from the obligation, because whatever we do on the basis of these motives will not *constitute* releasing ourselves.

Second, although my amended version of Schofield's conception denies that my obligation to go to the gym can be waived from my perspective as a philosopher, the ends, interests, and motives associated with that perspective might nonetheless *inform* my decision to waive the

obligation. Although these ends, interests, and motives are not available to me in the first person while I am deliberating from my perspective as an athlete, I might nevertheless *empathise* with them, recognising that I *should* release myself from the obligation ('I should not be so hard on myself!'). This is analogous to the case of obligations to others: If you promise me to meet me for lunch, the resulting obligation cannot be waived on the basis of *your* motive to keep working in order to meet an important deadline (only I can waive the obligation, and I cannot act on *your* motives). However, empathising with you, I can conclude that I should release you from your obligation.

In this section, I argued that Schofield's conception implies that the Psychological Inviolability Thesis is false, even for obligations to ourselves from which we can release ourselves at the time at which they bind us. Consequently, unless some compelling argument against Schofield's conception is provided, defenders of obligations to oneself have nothing to fear from a version of the Incoherence Argument that is based on the Psychological Inviolability Thesis.

6. Implications for Moral Theory

In the previous two sections, I showed that providing support for the Inviolability Thesis, and thus for premise (3) of the Incoherence Argument, is not as easy as one might think. In this section, I want to outline the implications of my arguments for moral theory. In addition to dispelling doubts about the *coherence* of obligations to oneself, my arguments illuminate the nature and plausibility of such obligations as well as their relation to the rest of morality.

To begin with, note that the conception of waivers as exercises of a normative power to which I appealed in Section 4 and Schofield's conception of obligations to oneself together form a coherent and independently plausible picture of obligations to oneself. The former implies that we can waive obligations to ourselves only in so far as we possess the normative power to do so, and thus, we might say, only in so far as we hold the relevant *office*. Schofield's conception adds that we hold the relevant office only from a certain perspective, which is distinct from the

perspective from which we are bound by the obligation. In so far as this perspective is tied to a practical identity, we hold the relevant office in our capacity as someone with this practical identity. For example, if I can waive my obligation to go to the gym today only from my perspective as an athlete, I hold the relevant office only in my capacity as an athlete.

This implication seems plausible. After all, many moral phenomena that are naturally described as obligations to oneself have this structure. Allen Habib appeals to the example of an army captain who also happens to be the army paymaster [2009: note 9]. Habib argues that the captain owes it to himself to pay his salary, in the same way in which he owes it to all the other soldiers to pay their salaries. Importantly, Habib notes that this obligation is owed by the captain, in his capacity as paymaster, to himself, in his capacity as soldier [2009: 555]. Accordingly, if the captain can waive this obligation, he can do so in his capacity as soldier, not in his capacity as paymaster.

The idea that obligations to ourselves depend on our ability to take up different perspectives that come with different normative powers or offices suggests that, when we have such an obligation, our self is in some sense *divided*. Indeed, Korsgaard suggests that when we are under an obligation because of one of our practical identities, our ‘acting self’ is accountable to our ‘thinking self’, which demands that we conform to our practical identity [1996: 104] (also see [2007]). In this way, the overall picture of obligations to oneself that emerges from my arguments in the two previous sections corroborates my earlier observation that obligations to oneself exhibit the same *second-personal* form as obligations to others (see Section 3). That is, this picture suggests that when we waive or hold ourselves to an obligation to ourselves, we address ourselves *as if* we addressed another person.⁹

⁹ Darwall’s second-personal framework also corroborates the conception of waiving as a normative power that I invoked in Section 4. Darwall holds that waiving is an exercise of authority which, by itself, effects normative changes and thus forms part of a distinct conceptual domain (see [2006: 11]).

Crucially, this suggests a significant *symmetry* between obligations to oneself and obligations to others. Just as our obligations to others are characteristically demandable by these persons, putting us in a relation of accountability to them, obligations to ourselves characteristically involve a relation of accountability between two perspectives that we can take up on our own actions.¹⁰ Both types of obligations are thus characteristically accompanied by the disposition to adopt reactive attitudes that address implicit demands in a person-to-person mode. Therefore, if the picture that emerges from my arguments can be sustained, it vindicates obligations to ourselves in the fullest sense; for it implies that these obligations are no less *obligatory*, and no less *moral*, than our moral obligations to others.

By stressing this implication of my arguments, I also want to forestall a potential objection. It might be objected that the most that my arguments can achieve is a pyrrhic victory over the Incoherence Argument (see Munoz [forthcoming]). According to this objection, while my arguments may defend the coherence of obligations to oneself *in some sense*, they abandon the notion that these obligations are truly *to oneself* by suggesting that whoever has such an obligation is in a sense divided.

It is worth noting that this objection threatens to do the Incoherence Argument a disservice. To my knowledge, nearly all defences of obligations to oneself invoke some sort of division of the agent's self (see Habib [2009], Kant [1797: 6:418], Rosati [2011], Schofield [2015, 2019], Timmermann [2006], also see Zylberman [forthcoming]). To say that the best that these defences can achieve is a pyrrhic victory over the Incoherence Argument is to say that they are compatible with the spirit, if not the letter, of the Incoherence Argument. But if that was the case, then the Incoherence Argument would come close to attacking a strawman.

¹⁰ Schofield himself presents his conception of obligations to oneself as second-personal [2015: 516-23].

More importantly, the picture of obligations to oneself suggested by my arguments is *not* compatible with the spirit of the Incoherence Argument. Singer's ultimate goal is to discredit all views that use the expression 'obligations to oneself' as anything other than a colourful way of referring to requirements of *prudence*. Following J. S. Mill, he holds that actions can constitute *moral* failings only in so far as they affect *others* [Singer 1959: 205] (see Mill [1989: 78-9]). Actions that only affect the agent, by contrast, can at most constitute prudential failings. Consequently, for Singer, 'to suppose that one can actually have a moral duty to oneself, in any literal sense, is to confuse [...] an imprudent act with an immoral one' [1959: 203]. What is more, the reason why Singer reserves the label 'moral' for the interpersonal sphere is that he takes his Incoherence Argument to show that we cannot be *accountable* to ourselves in the way in which we are to others [Singer 1959: 205; 1963: 137]. By depicting obligations to oneself as involving a relation of accountability to ourselves, and thus as 'moral' in the same sense as obligations to others, my arguments in this paper decidedly reject Singer's view.

7. Conclusion

I have argued that the Incoherence Argument fails. For one thing, this argument relies on the Inviolability Thesis, which defenders of obligations to oneself are committed to reject. For another, the Inviolability Thesis is harder to support than one might think. On its conceptual interpretation, the Inviolability Thesis is incompatible with the view that to waive an obligation is to exercise a normative power. On its psychological interpretation, it does not straightforwardly support the Incoherence Argument and it is incompatible with the view that obligations to oneself involve two perspectives, one from which the obligation can be waived and one from which it binds. I also pointed out that my arguments, taken together, paint a picture in which obligations to oneself are symmetrical to obligations to others: both types of obligation rest on relations of accountability.

Acknowledgements

For helpful comments and discussions, I am grateful to Joseph Bowen, Stephen Darwall, Ben Sachs, Robert Simpson, Jens Timmermann, two anonymous reviewers, and audiences at Durham, St Andrews, Stirling and UCL.

REFERENCES

- Chang, Ruth 2009. Voluntarist Reasons and the Sources of Normativity, in *Reasons for Action*, ed. David Sobel and Steven Wall, Cambridge, UK: Cambridge University Press: 243-71.
- Cholbi, Michael 2015. On Marcus Singer's "On Duties to Oneself", *Ethics* 125/3: 851-53.
- Cholbi, Michael 2018. Paternalism and Duties to Self, in *The Routledge Handbook of the Philosophy of Paternalism*, ed. Kalle Grill and Jason Hanna, London: Routledge: 108-18.
- Darwall, Stephen 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*, Cambridge, MA: Harvard University Press.
- Darwall, Stephen 2010a. *Precis: The Second-Person Standpoint*, *Philosophy and Phenomenological Research* 81/1: 216-28.
- Darwall, Stephen 2010b. Authority and Reasons: Exclusionary and Second-Personal, *Ethics* 120/2: 257-278.
- Dougherty, Tom 2015. Yes means Yes: Consent as Communication, *Philosophy & Public Affairs* 43/3: 224-53.
- Habib, Allen 2009. Promises to the Self, *Canadian Journal of Philosophy* 39/4: 537-57.
- Hill, Thomas Jr. 1991. Promises to Oneself, in *Autonomy and Self-Respect*, Cambridge, UK: Cambridge University Press: 138-54.
- Hills, Alison 2003. Duties and Duties to the Self, *American Philosophical Quarterly* 40/2: 131-42.
- Hobbes, Thomas 1651 (1991). *Leviathan*, ed. Richard Tuck, Cambridge, UK: Cambridge University Press.

- Hohfeld, Wesley N. 1917. Fundamental Legal Conceptions as Applied in Judicial Reasoning, *The Yale Law Journal* 26/8: 710-70.
- Hurd, Heidi M. 1996. The Moral Magic of Consent, *Legal Theory* 2/2: 121-46.
- Kant, Immanuel 1797 (2017). *The Metaphysics of Morals*, ed. Lara Denis, trans. Mary Gregor, Cambridge, UK: Cambridge University Press.
- Kant, Immanuel 1785 (2011). *Groundwork of the Metaphysics of Morals, A German-English Edition*, ed. and trans. Mary Gregor and Jens Timmermann, Cambridge, UK: Cambridge University Press.
- Korsgaard, Christine M. 1996. *The Sources of Normativity*, Cambridge, UK: Cambridge University Press.
- Korsgaard, Christine M. 2007. Autonomy and the Second Person Within: A Commentary on Stephen Darwall's *The Second-Person Standpoint*, *Ethics* 118/1: 8-23.
- Lavin, Douglas 2004. Practical Reason and the Possibility of Error, *Ethics* 114/3: 424-57.
- MacCormick, Neil and Raz, Joseph 1972. Voluntary Obligations and Normative Powers, *Proceedings of the Aristotelian Society, Supplementary Volumes* 46: 59-102.
- May, Simon C. 2015. Directed Duties, *Philosophy Compass* 10/8: 523-32.
- Mill, John S. 1989. *On Liberty and Other Writings*, ed. Stefan Collini, Cambridge, UK: Cambridge University Press.
- Muñoz, Daniel forthcoming. The Paradox of Duties to Oneself, *Australasian Journal of Philosophy*.
- Oakley, Tim 2017. How to Release Oneself from an Obligation: Good News for Duties to Oneself, *Australasian Journal of Philosophy* 95/1: 70-80.
- Owens, David 2011. The Possibility of Consent, *Ratio* 24/4: 402-21.
- Raz, Joseph 1985. Authority and Justification, *Philosophy & Public Affairs* 14/1: 3-29.
- Raz, Joseph 1999. *Practical Reason and Norms*, Oxford: Oxford University Press.

- Rosati, Connie 2011. The Importance of Self-Promises, in *Promises and Agreements: Philosophical Essays*, ed. Hanoch Sheinman, Oxford: Oxford University Press: 125-55.
- Schaab, Janis 2019. Commitment and the Second-Person Standpoint, *Zeitschrift für philosophische Forschung* 73/4: 511-32.
- Schofield, Paul 2015. On the Existence of Duties to the Self (and Their Significance for Moral Philosophy), *Philosophy and Phenomenological Research* 90/3: 505-28.
- Schofield, Paul 2019. Practical Identity and Duties to the Self, *American Philosophical Quarterly* 56/3: 219-32.
- Singer, Marcus G. 1959. On Duties to Oneself, *Ethics* 69/3: 202-05.
- Singer, Marcus G. 1963. Duties and Duties to Oneself, *Ethics* 73/2: 133-42.
- Strawson, Peter F. 1962. Freedom and Resentment, *Proceedings of the British Academy* 48: 1-25.
- Thompson, Michael 2004. What is it to Wrong Someone? A Puzzle About Justice, in *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, ed. R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith. Oxford: Clarendon Press: 333-84.
- Timmermann, Jens 2006. Kantian Duties to the Self, Explained and Defended, *Philosophy* 81/317: 505-30.
- Zylberman, Ariel forthcoming. Relational Primitivism, *Philosophy & Phenomenological Research*.