



SUSANNE BURRI 

IF YOU CARE ABOUT A RULE, WHY WEAKEN ITS ENFORCEMENT DIMENSION? ON A TENSION IN THE WAR CONVENTION

(Accepted 25 February 2022)

ABSTRACT. In *War by Agreement* (Oxford and New York: Oxford University Press, 2019), Yitzhak Benbaji and Daniel Statman argue that the ‘war convention’ – i.e. the international laws and conventions that are widely accepted to govern the use of force between sovereign states – represents a morally binding contract. On their understanding, the war convention replaces a pre-contractual morality governed by principles that so-called reductive individualists have identified and argued for over the past twenty years. This paper argues that if we take Benbaji and Statman’s contractarian interpretation of the war convention seriously, we have to conclude that its *in bello* rules have moral force only in contexts where its *ad bellum* rules were breached non-culpably. While Benbaji and Statman do not attend to this limitation of their argument, it has important ramifications in practice, as it is frequently unclear in the context of war whether one’s adversary is acting in good faith.

I. INTRODUCTION

Consider the following case:

Neighbours. You live on the edge of town. There is only one house close to yours, which is inhabited by Neighbour. Neighbour and you both like to throw parties that end around midnight. If you are not hosting a party or attending one, you both prefer peace and quiet over neighbourly noise. According to your local law and customs, any significant noise after 10 p.m. counts as a nuisance that must where possible be avoided. Over a cup of tea, you and Neighbour reach the following agreement: first, you each grant the other a permission to keep noise levels high until midnight. Second, you each waive your right to call the police about noise-related matters even if noise levels remain high after midnight.

There is something puzzling about the agreement you reach in Neighbours. While it prohibits partying after midnight, it also weakens the pre-contractual normative standing of someone affected by such partying by removing the right to call the police. If you and

Neighbour are genuinely interested in granting each other the *limited* permission to host parties that end *by midnight*, it seems that you should have agreed not to call the police before midnight, but reserved the right to escalate matters if noise levels remain high thereafter. Under what circumstances, if any, might your agreement nevertheless make sense? This is, in essence, the question that I aim to resolve in this paper.

Whether we can make sense of your agreement in Neighbours matters for the ethics of war. In *War by Agreement: A Contractarian Ethics of War* (Oxford and New York: Oxford University Press, 2019), Yitzhak Benbaji and Daniel Statman argue for a contractarian understanding of the international laws and conventions that are widely accepted to govern the use of force between sovereign states (Benbaji and Statman refer to these as the ‘war convention’; I follow their terminology). Benbaji and Statman grant that reductive individualists have over the past twenty years painted an accurate picture of the ethics of war as it obtains in the absence of a relevant agreement. Reductive individualists have not, however, paid attention to the fact that contractual agreements can sometimes replace pre-contractual morality. The war convention is best seen as a morally binding agreement that *replaces* the pre-contractual moral rules that would be binding in its absence – or so Benbaji and Statman claim.

On Benbaji and Statman’s view, the parties to the war convention aim to secure the legitimate goal of international peace. To this end, they agree that all unilateral first use of force should be prohibited, even if some such use is pre-contractually permissible. They also agree on a contractual right to resort to defensive force, largely because this serves as the main enforcement mechanism of the prohibition on first use of force. The parties to the contract moreover agree that, once a war has broken out, symmetric rules should apply to everyone involved. Roughly speaking, all combatants are required to avoid inflicting on each other harm that does not serve to weaken the other side militarily. All combatants are, moreover, required never to target non-combatants, and may risk incidental harm to non-combatants only under restrictive circumstances.

In this paper, I investigate whether we can make sense of the war convention as a morally valid contract that well-intentioned parties

might have agreed to. Crucially for my purposes, the pre-contractual morality of war that reductive individualists have sketched over the past two decades implies that the just side in a conflict enjoys important normative advantages over the unjust side. Combatants on the just side are, for example, generally permitted to target enemy combatants, whereas combatants on the unjust side lack such a permission.¹ The war convention, by contrast, recognises no such asymmetric permissions for the side it deems just, i.e. the one engaged in fighting a defensive war. In this way, the convention's *in bello* rules weaken the pre-contractual moral standing of those who try to enforce its *ad bellum* stipulations. There is thus a tension in the war convention that resembles the tension inherent in the agreement you reach in Neighbours. Both agreements lay out what may be done in response to a breach of their stipulated first order rules, and by doing this, both agreements seem to cater primarily to the interests of the party *in breach of the stipulated first order rules*.

Using the Neighbours case as an analogy, I investigate three attempts at making sense of the war convention's enforcement rules. The first argues that the convention's stipulated enforcement rules are justified because they help prevent the escalation of violence. The second maintains that using restraint in the enforcement of rules helps maintain friendly relations, which facilitates the resumption of a peaceful coexistence once a conflict comes to an end. The third attempt argues that contractually constrained enforcement rights are appropriate in situations where everyone is acting in good faith.

I show that the first two argumentative strategies fail. The third is more promising. It problematically implies, however, that there are significant limits to the circumstances under which the convention's *in bello* rules are morally binding. More precisely, it implies that the rules lack moral force in cases where a counterparty acted in bad faith, which it will seem reasonable to assume in many of the situations that the convention – at least on the face of it – was meant to regulate.

The argument that Benbaji and Statman provide in favour of the war convention's puzzling enforcement rules follows the third argumentative strategy laid out in this paper. It is, for this reason, not without merit. At the same time, Benbaji and Statman fail to

¹ See e.g. Jeff McMahan, *Killing in War* (Oxford and New York: Oxford University Press, 2009).

recognise that their argument implies an important limitation to the moral reach of the convention's *in bello* rules.

The remainder of this paper proceeds as follows. In section II, I introduce those aspects of the war convention, reductive individualism, and of Benbaji and Statman's contractarianism that are needed to bring out the tension inherent in the war convention that is the topic of this paper. In sections III to V, I discuss what seem to me the three main ways in which one might try to resolve this tension. Section VI concludes.

II. REDUCTIVE INDIVIDUALISM AND A CONTRACTARIAN INTERPRETATION OF THE WAR CONVENTION

The war convention's key stipulations are as follows. First, with respect to the justice of resort to war, i.e. the *jus ad bellum*, the convention prohibits all unilateral first use of force, thus imposing a blanket ban on what I will refer to as 'aggressive wars'. The only just cause for unilateral resort to force is to defend oneself – or an affected victim – against an aggressor. On Benbaji and Statman's interpretation, the convention permits resort to defensive force because this is the main mechanism through which the prohibition on aggressive wars can be enforced.² Next, in terms of the justice of fighting wars, i.e. the *jus in bello*, the war convention treats contractually just and unjust combatants alike. All combatants are at all times required to avoid inflicting on each other harm that does not serve to weaken the other side militarily. All combatants are, moreover, required never to target non-combatants, and must accept certain risks to their safety to minimise loss of non-combatant life where such loss is unavoidable and proportionate.³

² Benbaji and Statman, *War by Agreement*, p. 74.

³ Throughout this paper, I leave open whether the war convention equips combatants with symmetric *permissions* to inflict harm. Some scholars assume that it does, e.g. Michael Walzer, *Just and Unjust Wars* (New York: Basic Books, 2015, fifth edition), or Benbaji and Statman, *War by Agreement*. Others argue that legal *in bello* rules simply impose symmetric prohibitions on inflicting harm. On this second reading, the convention's *in bello* rules 'permit' the infliction of harm only in the weak sense that combatants are exempt from criminal prosecution if they inflict harms that are not explicitly prohibited (see e.g. Janina Dill and Henry Shue, 'Limiting the Killing in War: Military Necessity and the St. Petersburg Assumption', *Ethics and International Affairs* 26: 36–63 and Adil A. Haque, *Law and Morality at War* (Oxford and New York: Oxford University Press, 2017). What matters for my purposes is that on both interpretations, the war convention lowers the normative standing of just combatants by depriving them of normative advantages they would enjoy over unjust combatants in its absence.

Michael Walzer argues that the war convention is not merely a legal and conventional construct, but captures the essence of the morality of war as well.⁴ Even though the combatants who are fighting an aggressive war lack a just cause, the injustice of their war is not *their* injustice, and, for this reason, it is appropriate that unjust combatants should enjoy the same moral standing as their adversaries. Over the past twenty years, so-called reductive individualists have pushed back against Walzer's view. They argue that the morality of war reduces to our ordinary morality; war is not an activity that is normatively exceptional.⁵ They also contend that the rights and duties of collectives – such as those of a citizenry or an army – supervene on the rights and duties held by the individuals who make up these collectives.⁶ Based on these assumptions, they argue that just and unjust combatants cannot be regarded as moral equals. On their view, unjust combatants are generally liable to be killed, because they share in the moral responsibility for the unjust threats that their armies pose. Even if they have an excuse of duress or of ignorance, their actions are not morally on a par to the actions of just combatants. Just combatants are not generally liable to be killed, as they are not usually responsible for any unjust threats. Finally, some unjust civilians are liable to defensive harm in virtue of their moral responsibility for the unjust threats their combatants pose.⁷

At least on the face of it, it seems unlikely that Walzer's view and the claims of reductive individualists can be reconciled. Yet this is precisely what Benbaji and Statman set out to do. They take, as a starting point, the insight that we can sometimes alter pre-contractual moral rules by entering into morally binding agreements. In *Neighbours*, it seems appropriate to say that before entering into an agreement, you and Neighbour both enjoy a right to peace and quiet after 10 p.m., which is matched by a duty to keep noise levels down

⁴ Walzer, *Just and Unjust Wars*.

⁵ See e.g. McMahan, *Killing in War*.

⁶ *Ibid.*

⁷ See e.g. Helen Frowe, *Defensive Killing* (Oxford and New York: Oxford University Press, 2014), pp. 162–87. The claim that unjust non-combatants may be liable to defensive harm is controversial among reductive individualists. For an argument that they are not, see David Rodin, 'The Moral Inequality of Soldiers: Why the *Jus in Bello* Asymmetry is Half Right', in *Just and Unjust Warriors: The Moral and Legal Status of Soldiers*, edited by David Rodin and Henry Shue (Oxford and New York: Oxford University Press, 2008). For an argument that they rarely are, see Cécile Fabre, 'Guns, Food, and Liability to Attack in War', *Ethics* 120 (2009): 36–63; McMahan, *Killing in War*, pp. 213–231.

after 10 p.m.⁸ This changes once you grant each other a permission to throw parties that last until midnight.

Benbaji and Statman develop a careful argument in favour of the idea that the war convention is best understood as a morally valid contract that has been voluntarily entered into by all relevant parties due to its mutually beneficial nature. They grant that reductive individualists paint an accurate picture of the ethics of war as it obtains in the absence of a relevant agreement. The war convention, however, replaces the pre-contractual morality of war with a different set of morally binding rules. Here, I grant this general outlook for the sake of argument, and proceed to investigate why the parties to the contract might have agreed to a rather peculiar combination of contractual *ad bellum* and *in bello* rules.

What renders the war convention's combination of rules peculiar is its insistence that even though aggressive wars are contractually unjust, contractually just combatants should nevertheless not enjoy any normative advantages over unjust combatants once a war has broken out. This stands in contrast to the pre-contractual rules, which impose on unjust combatants prohibitions that do not apply to just combatants. At least on the face of it, it seems that if the parties to the contract had been serious about banning all aggressive wars, they should have kept the pre-contractual asymmetry between just and unjust combatants in place, possibly even rendering it more pronounced, thus enabling more effective defensive wars. If, by contrast, the parties did not actually care about prohibiting aggressive wars, and wanted primarily to ensure that whenever wars are fought, they are fought according to some basic and symmetric rules, then they should simply have agreed to that.

In what follows, I investigate three ways in which one might try to explain why the parties to the war convention settled for weak enforcement rights even if they cared about the convention's *ad*

⁸ I assume here that your local law and customs create a morally valid claim-right to peace and quiet after 10 p.m. More precisely, I assume that your local law and customs provide a reasonable sharpening of two vague moral rights that we generally possess, i.e. the right to peace and quiet on the one hand, and, on the other hand, the right to impose on others certain burdens that are the by-product of activities we have reason to value. Where local law and customs provide a reasonable specification of otherwise vague moral rights, this specification is morally binding, even if other specifications would be reasonable as well. Reasonable law and customs give rise to legitimate expectations, and they enable us to plan ahead and coordinate our behaviour, thus providing us with *pro tanto* moral reasons in favour of respecting them. For a pertinent discussion, see Tony Honoré, 'The Dependence of Morality on Law', *Oxford Journal of Legal Studies* 13 (1993): 1–17.

bellum prohibition on aggressive wars. I conclude that the first two explanation attempts fail. The third is more promising, but it significantly limits the moral reach of the convention's *in bello* rules.

III. EXPLANATION I: PREVENTING AN ESCALATION OF VIOLENCE

Consider again the situation in Neighbours. Suppose that while discussing the idea of extended partying permissions, your neighbour proposes to combine such extended permissions with a prohibition on calling the police even if the neighbourly contract is violated. You are taken aback, and ask your neighbour to clarify his thinking. He responds as follows: 'Well, it's very embarrassing for the host of a party if the police show up to shut the party down. Naturally, it creates resentment in the host towards whoever called the police. I, for one, wouldn't be surprised if the host decided to take revenge. By poisoning the caller's pet, say. Or by collecting cat droppings, and scattering them in the caller's backyard. It's best to just talk to each other if we have a problem, isn't it?' You do not know Neighbour well, and you are unsure whether his words are intended as a threat. You decide, at any rate, never to call the police on Neighbour if it can be avoided, as you care about your dog Jax. For Jax's sake, you agree not to call the police, figuring that if the prohibition is contractual, then at least it goes both ways.

In the context of war, analogous reasoning is sometimes adduced to justify the war convention's symmetric *in bello* rules. The key idea is that symmetric constraints on the use of force can help prevent an escalation of violence. Whatever weapons or tactics one side makes use of, the other is likely to match or outdo if – as seems likely – it deems it conducive to its ends.⁹ Even if one side is fighting a contractually just war, it is thus advisable for that side to use restraint and to desist, for example, from targeting enemy civilians, no matter how compelling the evidence that they are liable to be killed. After all, the opposing side has already shown that it will not respect the territorial integrity and the right to life of the combatants of the state that it has attacked, and if mutual restraint is necessary to keep non-combatants safe, then this will usually be the best policy for the

⁹ See e.g. Walzer's discussion of Carl von Clausewitz' ideas in *Just and Unjust Wars*, pp. xxx, 13–15, and 23–24, as well as Christopher Kutz, 'Fearful Symmetry', in *Just and Unjust Warriors. The Moral and Legal Status of Soldiers*, edited by David Rodin and Henry Shue (Oxford and New York: Oxford University Press, 2008), esp. p. 75.

victim state from both a prudential and a moral perspective. More precisely, both the prudential and the moral value of what the just combatants can expect to achieve by targeting liable enemy civilians will usually be outweighed by the expected disvalue of the havoc the enemy would wreak by targeting civilians. Consider again the structurally similar situation in Neighbours: there, both a prudential and a moral calculation make it advisable not to call the police to protect Jax.

This first explanation goes some ways towards justifying symmetric *in bello* rules, just as it goes some way towards justifying an agreement not to call the police if you are dealing with a dodgy neighbour. To the extent that it does, however, it simultaneously weakens Benbaji and Statman's core claim that the war convention's *in bello* rules are morally binding because they are part of a valid contract. To the extent that the *no escalation of violence* explanation is to the point, it is the looming threat of unjust violence that renders it prudentially and morally advisable to agree to weak enforcement rules. Such a threat cannot, however, underlie a relevantly voluntary and, for this reason, morally valid agreement.¹⁰ If we do not have effective means at our disposal to ensure that an unjust threat will not be acted upon, we can be both prudentially and morally required to take it into account. Any such requirement, however, exists independently of what we agree to, and whatever rules that we agree to due to a threat of wrongful harm are not thereby imbued with any additional moral force.

Suppose that, in Neighbours, Jax dies of natural causes. A few days after Jax's death, your neighbour hosts a party that shows no signs of dying down well beyond midnight. You knock on Neighbour's door, and he promises to wrap things up, but an hour later, the party is still going strong. If you then decide to call the police, figuring that you can deal with a barrage of cat droppings on your lawn the next day, it seems to me you are acting morally entirely permissibly. Neighbour is culpably in breach of his contractual obligations, and by calling the police, you are making use of an enforcement right that you agreed to waive only because Neighbour threatened unjust harm. By analogy, if a victim state is confident that it can better protect itself against an aggressor's unjust violence if it

¹⁰ In contract law, such an agreement would be unenforceable due to duress. See e.g. Randy E. Barnett, 'A Consent Theory of Contract', *Columbia Law Review* 86 (2) (1986): 269–321, at p. 318.

makes use of its entire range of pre-contractually available enforcement rights, then the fact that it has agreed to use restraint to keep an unjust enemy in check does not morally speak in favour of using such restraint.

IV. EXPLANATION II: MAINTAINING FRIENDLY RELATIONS

A second explanation of weak enforcement rules seems appropriate in contexts where relationships are rather more amicable. Suppose that, in *Neighbours*, Neighbour is not a stranger (and possible lunatic), but a close friend whom you have known since childhood, and who has just bought the house next to yours – with your enthusiastic support. While discussing extended partying opportunities, you agree that you will attend most of each other's parties anyway, and that strict rules are therefore largely unnecessary. Still, you want to make sure that there is no bad blood due to parties not attended by both. You thus agree on a mutual permission to host parties until midnight, and you decide that, as friends, you will resolve any differences amicably, without the involvement of law enforcement.

Against this backdrop, the agreement that you reach in *Neighbours* seems reasonable. Both you and Neighbour are interested in remaining friends, and you both agree that it might sour relations between you if you involve the police should a dispute arise. An agreement not to call the police thus seems mutually beneficial, and at least potentially morally binding.

In the international realm, one might tell the following analogous story. In the absence of a world government that is capable of adjudicating international disputes, it is crucial that states develop and maintain friendly relations. To flourish, each state depends on the respect, the cooperation, and the good will of others. For this reason, it is helpful if states develop strong economic ties, thus aligning their interests. It is helpful also if they agree to constrain their use of force should a conflict turn violent.¹¹ Symmetric *in bello* rules that provide strong protections to civilians are a relatively amicable way of resolving disputes because they provide all parties to the dispute with a way of dividing its population into a part that it

¹¹ As Henry Sidgwick puts it in *The Elements of Politics*, this helps avoid 'causing bitterness that will long outlast' the fighting (quoted in Walzer, *Just and Unjust Wars*, p. 132).

agrees may be targeted, and another part that it stipulates may not. This provides each party to the conflict with valuable discretion.¹² Respecting the right to life of citizens that an opponent has deemed 'off limits' helps resume friendly relations once a war comes to an end.

As in Neighbours, this narrative appears promising. It renders the war convention's *in bello* rules at least potentially mutually beneficial, and, for this reason, at least potentially morally binding, even if the parties to the contract care greatly about the convention's prohibition on fighting aggressive wars. There is, however, a difficulty.

Consider again the case of Neighbours under the assumption that Neighbour is your friend. It seems to me that if the two of you were to properly think things through, you might want to agree to call the police only as a *last resort*, but you would not want to completely waive your right to do so. Once there is a need for enforcement, this means that some rule has been broken. If the infraction is minor, it is sensible that two friends would aim to resolve things amicably. But if the infraction is serious, and if amicable ways of trying to resolve things are not productive, this inevitably means that previously friendly relations are no longer as chummy. There may then come a point at which the victim of a contractual breach wishes primarily to enforce the contract and is no longer interested in treating the counterparty amicably. The victim may judge, in fact, that taking more drastic steps not only increases their chance of getting their due, but is also their most promising means of salvaging what is left of the relationship. In light of this, you and your friend do best if you either leave the choice of appropriate enforcement means to whoever happens to be a victim of the other's excessive partying, or if you contractually stipulate what common sense dictates anyway: that you will call the police only as a last resort.

The situation is similar in the international realm. In a situation where one state launches an aggressive war, thus breaching its contractual obligations, the relationship between the aggressor and its victim has already soured. It is thus predictable that a victim of aggression may wish to make use of its entire range of pre-contractually admissible enforcement rights, and would regret having granted to enemy combatants some privileges they would not

¹² Frowe, *Defensive Killing*, p. 169; George I. Mavrodes, 'Conventions and the Morality of War', *Philosophy and Public Affairs* 4 (1975): 117–131. See also Walzer, *Just and Unjust Wars*, p. 116.

otherwise have had. If some basic trust remains, the victim may choose to use restraint in the hope that this will help rebuild relations once the violence ends. At the same time, if it judges that the costs of using restraint outweigh the benefits, it may want to act on this judgement. As all of this is entirely predictable, it seems that parties interested in maintaining friendly relations should not waive any of their pre-contractual enforcement rights, but should at most agree to what morality and common sense require anyway, namely that the harsher pre-contractually available enforcement mechanisms be used only as a last resort.

Admittedly, it can be difficult to stick to a resolution to resolve conflicts amicably once emotions run high. It might thus be suggested that waiving particular enforcement rights helps us stick with a resolution not to use them, thus keeping us from doing what we would later regret. In *Neighbours*, you may very much want law enforcement to put an end to your friend's excessive partying after a bad day at work that has left you with zero energy to confront your overly merry friend directly. At the same time, you might judge this an unduly hostile reaction that you really ought to avoid. Having waived your right to call the police might help you act in accordance with your considered judgement.

This seemingly promising suggestion cannot salvage the *maintaining friendly relations* explanation, however. To see why, suppose that a victim respects contractual stipulations, and does not call the police if he or she has contractually waived the relevant right. Such a victim will take seriously any agreed-upon last resort clause as well, and will resort to harsher enforcement mechanisms only once more amicable means have been exhausted. For the reasons just laid out, this seems preferable to a situation where the harsher enforcement mechanisms are no longer available at all. If, by contrast, a victim ignores contractual stipulations once emotions run high, it does not matter whether the victim has contractually agreed to call the police only as a last resort or not at all – the victim will immediately leap to drastic enforcement mechanisms either way.

V. EXPLANATION III: DEALING WITH GOOD FAITH MISTAKES

A third explanation in favour of weak enforcement rights even if the parties to the contract care about the stipulated first order rules is as

follows. Some first order rules are such that it can at times be difficult to know whether one is violating them. One might thus violate them *in good faith*, i.e. without being aware that one is violating them, and in the presence of evidence suggesting that one is not, in fact, violating them. Where such good faith mistakes are likely, it is reasonable to impose contractual limits on what should count as admissible enforcement. Ruling out some of the harsher enforcement mechanisms is sensible because of the risk that one might become a good faith violator oneself.

To consider this third explanation in more detail, imagine that, in Neighbours, neither you nor Neighbour have a precise way of telling the time. The police are able to tell the time with accuracy, but it is expensive for them to do this, and they charge a hefty fee for the service (to be paid by the person partying past midnight, or to the person calling before midnight, as the case may be). With these details in place, you and Neighbour might predictably find yourselves in a situation where the host to a lively party is confident that it is not yet past midnight, and where the sleep-deprived neighbour is equally confident that midnight must have come and passed. If you and Neighbour trust each other to act in good faith, it may be in your mutual benefit contractually to rule out calling the police. Doing so commits you to resolve neighbourly disputes amicably, and might save you lots of money.

Compared to the first two explanations, this third explanation has a key advantage. Both the first and the second explanation implicitly grant that whoever breaches the first order rules may well do so culpably. But once a party is in culpable breach of the central terms of a contract and shows no willingness to make amends, this morally releases its counterparties from whatever contractual obligations they have towards the culpable party. Suppose that I have a pear tree, and you have an apple tree. I promise you a weekly delivery of 20 pears throughout September and October. You promise, in return, a weekly delivery of 20 apples during the same time period. Once September rolls around, I send you two bags of pears over two weeks, receiving nothing in return. I give you a call, and you tell me that you have decided to keep your apples to yourself. This morally releases me from my contractual obligations towards you: the fact that you are culpably ignoring yours cancels mine.

Of course, the war convention and the contract in *Neighbours* are importantly different from the produce example just introduced. Both the war convention and the contract in *Neighbours* stipulate rules that are meant to apply once the contract's first order rules have been breached. That the first order rules might be breached, and possibly culpably so, is thus anticipated by the contracting parties, who have contractually laid out what rules should come into force if the contract's first order rules are breached culpably. If all the parties to the contract have voluntarily agreed to it, its stipulations should thus remain binding even if a culpable breach occurs.¹³

But this seems to me too quick. It matters for the moral validity of a contract that its stipulations were agreed to by parties negotiating in good faith. Parties negotiating in good faith consider what rules are in their expected interest *on the assumption that they themselves – along with all the other parties to the contract – will make a reasonable effort to abide by the agreed-upon rules*. If part of a contract caters to the interests only of parties who expect to breach it in a culpable manner, then that part of the contract is morally void. If there are no legitimate interests to ground part of a contract in, then that part of the contract lacks a moral justification that could render it valid.¹⁴ In *Neighbours*, we deem the agreement that you enter into mutually beneficial if both you and Neighbour are better off if you both comply with the agreement as opposed to the pre-contractual moral rules. For the contract's first order rules, this is plausibly the case if you both like to host parties that last until midnight, and if you both do not greatly mind the noise that you are exposed to if the other hosts a party. Such preferences are morally legitimate, and can thus ground a morally valid agreement to depart from pre-contractual morality. We cannot, however, apply any structurally similar thinking to your agreement not to call the police even if the other party is in culpable breach of the contract. On the assumption that both you and Neighbour make a good faith effort to comply with the

¹³ Benbaji and Statman make this point in *War by Agreement*, p. 164.

¹⁴ In chapter 2 of *War by Agreement*, Benbaji and Statman argue that a contract is morally valid if it satisfies the conditions of 'mutual benefit', 'fairness', and 'actuality'. 'Mutual benefit' is satisfied if all the parties to the contract expect to be better off if the pre-contractual rules are replaced by contractual ones (pp. 43–45). Crucially for my purposes, Benbaji and Statman do *not* discuss how we should think about contractual rules that cater to the interests only of parties who do not intend to abide by the contract's first order rules. While they argue that 'chauvinistic or racist preferences' should be excluded as indecent (p. 45), they leave unaddressed how we should think about rules that aim to change the normative landscape in a way that seems desirable only from the perspective of culpable non-compliers.

rules of the agreement, this stipulation is of no benefit to anyone, as it never comes into play. It comes into play only if there has been a culpable breach, in which case it protects the party who has culpably breached the contract. But only a bad faith negotiator would aim to protect the interests of someone who is in culpable breach of the contract. The relevant stipulation thus lacks legitimate interests that might ground it and render it morally valid.

The situation looks different when we are dealing with *good faith* breaches of contractual first order rules. To breach a contract in good faith is to breach it while believing, based on reasonable evidence, that one is abiding by its stipulated terms. Avoiding good faith breaches is something that is largely outside of our control. More precisely, such breaches can happen even if we take our contractual obligations seriously, and make a reasonable effort to live up to them. Consider again the version of Neighbours where the police charges a hefty fee for telling you the time. If you and Neighbour agree not to call the police as long as both parties to the contract are acting in good faith, this waiving of enforcement rights is in your interest even if you both intend to abide by the contract's first order rules. In disputes where it is reasonable to assume that everyone is acting in good faith, you and Neighbour will both have to admit that it may be before or past midnight, even if you disagree about the relative likelihood of the two alternatives. As you both have to admit that your neighbour may be within their contractual rights, you know that calling the police might set you back financially. If you and Neighbour can work things out between yourselves, this is thus preferable to everyone involved.

With these ideas in hand, let us consider the war convention. At first sight, one might be tempted to argue that the *good faith explanation* in favour of weakened enforcement rights may be sound in principle, but lacks practical application when it comes to the war convention. More precisely, it might seem strange to think that there are 'honest mistakes' when it comes to launching an aggressive war. On closer inspection, however, such mistakes are not unusual. As the tensions in a geopolitical region rise, the parties to the conflict may want to secure their borders and strengthen strategically significant positions. They may also want to show off their military prowess to deter an enemy who is perceived as potentially aggres-

sive. Some such efforts can easily be misinterpreted as acts of aggression by an alert enemy, who in turn decides to strike back. The enemy's 'defensive' strike, in turn, will seem aggressive to the side at which it is aimed, and in this way the escalation of violence might continue until the parties to the conflict are fighting a full-blown war everyone hoped to avoid.

If there is such a dynamic, it is beneficial for everyone involved if the *in bello* rules apply symmetrically and put clear constraints on the violence the parties may permissibly resort to. If tensions in some region run high, and a particular government reasonably reads an enemy's actions as aggressive while also being aware that the enemy might understand the situation differently, striking back in a restrained manner, as well as granting the enemy similar rights to use force, may be the best available option. Striking back in a restrained manner signals to the opponent that it faces resistance if it tries to advance, but it also shows basic respect for the opponent. Finally, if the government is committing a good faith error and the military actions that it decides upon are aggressive as opposed to defensive, the wrong that it and its combatants commit is, at the very least, more limited than it would be if it used less restraint.¹⁵ Agreeing to the war convention's *in bello* rules in circumstances where a good faith assumption is reasonable can thus be based on the legitimate preferences of parties who intend to comply with their contractual obligations, and such rules can for this reason be validly agreed to.¹⁶

Importantly, however, not all breaches of the war convention's first order rules happen in good faith. Where an aggressor acts culpably, there can be no valid contractual weakening of the victim's pre-contractual moral standing. Instead, the asymmetric *in bello* rules sketched by reductive individualists stay firmly in place. This is not to say, of course, that a victim state ought to make use of its full range of pre-contractual enforcement rights. Both the *no escalation of*

¹⁵ If the war convention's *in bello* rules assign symmetric permissions, the combatants who are committing a good faith error do no wrong as long as they stick to the rules. If the war convention merely carves out actions that are not subject to criminal prosecution without fundamentally altering what is pre-contractually morally permissible, then the combatants on the side that is committing a good faith error do less wrong if they abide by the war convention's *in bello* rules as opposed to fighting in accordance with their perceived pre-contractual permissions. See fn. 3.

¹⁶ Of course, limiting the wrong that they commit may be in the interest also of parties who culpably breach the convention's *ad bellum* rules. Culpable parties may for this reason choose to fight with restraint. They cannot, however, limit the wrong that they do through relevant contractual agreements – or so I argue in this section.

violence and the *maintaining friendly relations* explanations highlight that the victim may reasonably decide not to make use of certain enforcement rights, and may possibly even be morally required to abstain from using them. What I hope my discussion has clarified, however, is that no such restraint can be required *because the victim state has entered into a morally binding agreement that asks for such restraint*.

If what I have argued for in this section is correct, then the war convention's *in bello* rules are morally binding as part of a contractual agreement only in contexts where its first order rules were breached in good faith. The tricky thing is that, in practice, it may frequently be difficult to know whether one's opponent is breaching the war contract in good faith. Under such tangled circumstances, one should maybe start out assuming that any perceived breach happened in good faith, and change one's assessment of the situation only in the presence of clear evidence to the contrary, thus giving the enemy the benefit of the doubt. This should help ensure that one takes seriously the possibility that the enemy is acting non-culpably even in situations where it is tempting to assume otherwise. Starting out by giving the enemy the benefit of the doubt is by no means guaranteed, however, to lead to an accurate assessment of the situation. It is no wonder, then, that war is a normatively difficult business.

Admittedly, there may be a compelling explanation in favour of contractually weakened enforcement rights that my discussion so far has failed to consider. On the argument that Benbaji and Statman put forward, symmetric *in bello* rules are essential for ensuring that soldiers are willing to fight the wars that they are ordered to fight.¹⁷ If the conventional rules of war prohibited fighting for the unjust side and threatened criminal prosecution for doing so, this would render it difficult to fight defensive wars, as soldiers would refuse to fight if they had any doubts about the justice of their war.¹⁸ Suppose that Benbaji and Statman are correct about their empirical claims. If so, this suggests a fourth reason why one might want to agree to a contractual weakening of enforcement rights. Roughly speaking, weakening the rights might strengthen the extent to which they can

¹⁷ Benbaji and Statman, *War by Agreement*, pp. 117–118.

¹⁸ *Ibid.*

be acted upon, thus increasing what we might call ‘overall enforceability’.

Contrary to initial appearances, it seems to me that Benbaji and Statman’s idea does not introduce an independent fourth explanation in favour of contractually weakened enforcement rights. Their ideas are instead covered by the *dealing with good faith mistakes* explanation. According to Benbaji and Statman, what makes just combatants hesitant to obey orders is their *awareness that they might be committing a good faith mistake*. As I have argued in this section, it is quite reasonable for the parties to a contract to want to put limits on the costs that they will incur if they commit a good faith mistake. What Benbaji and Statman fail to notice is that their argument implies an important limitation of the moral reach of the war convention’s *in bello* rules.¹⁹

There might, of course, be further explanations that are not reducible to the three explanations that I have discussed in this paper. If so, they might vindicate the war convention’s contractual weakening of pre-contractual enforcement rights even in cases where the agreement is breached in a culpable manner. I do not currently see, however, what such further explanations might look like.²⁰

¹⁹ Benbaji and Statman do claim that the war convention is an agreement between ‘decent states’, and that decent states ‘tend to respect the contractual duties they undertake’ (War by Agreement, p. 72). This sounds as if Benbaji and Statman accept the limitation to their argument that I lay out in this paper. At the same time, their discussion clarifies that they think of decent states primarily as states that are committed to satisfying the war convention’s *in bello* rules (see War by Agreement, ch. 7). On their view, a state that knowingly breaks the contractual agreement not to launch any aggressive wars continues to count as decent as long as it reasonably believes that its aggressive war would be pre-contractually justified. As I see it, this way of looking at things misses the key consideration that states waive their right to whatever unilateral first use of force would be pre-contractually permissible by agreeing to the war convention’s *ad bellum* rules.

²⁰ An anonymous reviewer suggests that a promising explanation proceeds along the lines of Walzer’s argument in *Just and Unjust Wars*. The key idea is that the parties to the war convention are political leaders, and it is they who decide whether wars should be started, and, in this way, whether contractual *ad bellum* rules are respected. The people actually caught up in the fighting are very different from this political elite, and the contractually stipulated *in bello* rules aim to benefit primarily those caught up in the fighting. In *Neighbours*, this explanation is most akin to arguing that a rule not to call the police aims to protect primarily the guests at a party, who are there to have a good time. I believe that this explanation rightly draws our attention to the fact that those who are caught up in a contractually unjust war may frequently have an excuse of ignorance or duress. This is something that contractually just combatants are required to take into account *in accordance with pre-contractual moral rules*. It is not possible, however, to introduce by contractual fiat rules that aim to benefit a group of people one has a special relationship to for the specific circumstances where one has culpably breached the contract’s first order rules. Culpably breaching a contract is something that all contracting parties can and ought to avoid, and it is, for this reason, not something they can morally validly make self-serving provisions for.

VI. CONCLUSION

According to Benbaji and Statman, the war convention is a morally binding agreement that replaces the pre-contractual morality of war. While the pre-contractual morality endows just combatants with permissions that unjust combatants lack – including, potentially, permissions to target unjust non-combatants who are responsible for the unjust war that their combatants are fighting – the war convention treats contractually just and unjust combatants alike.

I have argued that if we take Benbaji and Statman's basic picture at face value, then there is something puzzling about the war convention, in that it regulates the enforcement of its first order rules in a way that seems to cater primarily to the interests of the party in breach of these rules.

I have investigated three ways in which one might seek to make sense of this peculiarity. According to the *no escalation of violence* explanation, the victim of a breach of first order rules may be well advised to use restraint in their response to keep a dodgy enemy in check. I have argued that this explanation can explain why using restraint may be both rational and morally required, but that restraint cannot, in such a case, be morally required because it has been contractually agreed upon.

According to the *maintaining friendly relations* explanation, the victim of a breach of first order rules may wish to use restraint in their response to ensure that friendly relations can be resumed once the conflict is resolved. I have argued that this explanation can render it appropriate to use harsher enforcement mechanisms only as a last resort, but that it cannot render it in the parties' interest to completely waive pre-contractually available enforcement rights.

Finally, according to the *dealing with good faith mistakes* explanation, the parties to a contract may wish to limit enforcement rights in cases where the first order rules of a contract have been breached in good faith. I have argued that this explanation is successful – it can render the war convention's *in bello* rules morally binding in the context of good faith breaches of its *ad bellum* rules. Not all such breaches happen in good faith, however. Where a breach is committed in bad faith, any contractual agreement to limit pre-contractually admissible enforcement rights must be rejected as morally invalid. In practice, whether a party to the war convention is bound

by its *in bello* rules can thus depend on facts about the moral character of its adversary's actions that are difficult to establish.

The argument that Benbaji and Statman offer in favour of the war convention's *in bello* rules falls under the *good faith* explanation. It thus applies only in cases where the convention's *ad bellum* rules were breached non-culpably. This is a point that Benbaji and Statman fail to make clear, and one that complicates issues tremendously, at least in practice. If the war convention's *in bello* rules were morally binding independently of any *ad bellum* considerations, it would be relatively straightforward for combatants to conform their behaviour to the moral rules that applied to them. The same is not true if reductive individualism provides the right way of thinking about the moral rules that apply to combatants. Part of the appeal of Benbaji and Statman's contractarian project is thus that it promises to greatly reduce the moral complexity that combatants face, while at the same time granting that the 'deep' morality of war is reductive individualist in nature. In this paper, I have argued that even though Benbaji and Statman's contractarian project is not without merit, it fails to establish that the war convention's *in bello* rules are morally binding independently of *ad bellum* considerations. Once we take reductive individualism seriously, it seems that we are stuck with the conclusion that combatants will frequently have trouble establishing what moral rules apply to them.

ACKNOWLEDGEMENTS

For very helpful written comments, I thank Janina Dill, Cécile Fabre, Linda Eggert, Hillary Nye, Anat Shapira, and an anonymous reviewer of this journal. I have greatly benefited, as well, from insightful discussions at a 'War by Agreement' roundtable organised by Janina Dill, Cécile Fabre, and Cécile Laborde. The roundtable took place in December 2020, and its participants included Yitzhak Benbaji and Danny Statman.

FUNDING

Open Access funding enabled and organized by Projekt DEAL.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

*Department of Philosophy
University of Konstanz, Fach 18, 78457, Constance, Germany
E-mail: susanne.burri@uni-konstanz.de*

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.