



In Bello Proportionality: Philosophical Reflections on a Disturbing Empirical Study

Stephen de Wijze, Daniel Statman & Raanan Sulitzeanu-Kenan

To cite this article: Stephen de Wijze, Daniel Statman & Raanan Sulitzeanu-Kenan (2022): In Bello Proportionality: Philosophical Reflections on a Disturbing Empirical Study, Journal of Military Ethics, DOI: [10.1080/15027570.2022.2131104](https://doi.org/10.1080/15027570.2022.2131104)

To link to this article: <https://doi.org/10.1080/15027570.2022.2131104>



Published online: 21 Oct 2022.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



In Bello Proportionality: Philosophical Reflections on a Disturbing Empirical Study

Stephen de Wijze^a, Daniel Statman^b and Raanan Sulitzeanu-Kenan ^c

^aManchester Centre for Political Theory (MANCEPT), University of Manchester, Manchester, United Kingdom; ^bDepartment of Philosophy, University of Haifa, Haifa, Israel; ^cFedermann School of Public Policy, The Hebrew University, Jerusalem, Israel

ABSTRACT

A recent empirical study has argued that experts in the ethics or the law of war cannot reach reasonable convergence on dilemmas regarding the number of civilian casualties who may be killed as a side effect of attacks on legitimate military targets. This article explores the philosophical implications of that study. We argue that the wide disagreement between experts on what *in bello* proportionality means in practice casts serious doubt on their ability to provide practical real-life guidance. We then suggest viewing *in bello* proportionality through the prism of virtue ethics.

KEYWORDS

Proportionality; experts; just war theory; virtue ethics

1. Introduction

In bello proportionality requires that soldiers make sure that the military force they use does not cause disproportionate harm to civilians who are located in proximity to military targets.¹ In the well-known words of the Geneva Convention, the warring sides are required to “refrain from deciding to launch any attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.”²

While most writers agree that the implementation of the proportionality principle is hard and often involves grey zones,³ lawyers, politicians and journalists generally assume that it provides a reliable criterion to determine – both *ex ante* and *post factum* – the moral and legal permissibility of military attacks that involve harm to civilians. That is why they often feel confident in condemning allegedly disproportionate attacks. A recent empirical study that we conducted (Statman et al. 2020, hereafter: “STUDY”), however, raises serious doubts about this assumption. The current article presents an outline of this study and reflects on its philosophical and normative implications. Our contention is that if the findings of STUDY are valid, it calls for a radical re-evaluation of *in bello* proportionality.⁴

The skepticism that STUDY encourages with regard to *in bello* proportionality are closely related to skepticism regarding proportionality in general, especially in the

context of constitutional law.⁵ Although STUDY does not by itself show “the limited usefulness of the proportionality principle” in this wider context, it does make one concerned that this is the case. Indeed, Anne Peters has recently noted (2021) that STUDY confirms Chang and Dai’s diagnosis (2021) that “proportionality does not offer a rational decision-making tool.”

2. A brief summary of STUDY

The purpose of STUDY was to examine whether the proportionality condition provides reliable guidance for solving real-world dilemmas concerning the legitimacy of collateral harm that results from military attacks. To do so, STUDY sought to test the proportionality judgments of experts, about whom it is reasonable to assume that they have the best understanding of *in bello* proportionality.

In order to test the reliability of experts in other fields, the best practice is to compare their judgments or predictions against reality. For instance, to determine whether the tools used by meteorologists are reliable, their predictions are compared to actual weather conditions that obtain. Similarly, the efficacy of concepts and theories used by political scientists are tested by comparing their predictions to actual political developments. In both examples, predictions are compared in time with reality and this serves as an objective criterion to test the validity and reliability of the respective experts. However, with normative judgments, there are no facts of the matter in objective reality that can serve to either confirm or falsify them. Normative judgments are about what ought to be, not about what there is, hence looking at what there is – at reality – cannot help in establishing their validity or reliability. In this sense (though not in others), normative judgments are quite clearly not objective; even radical moral realists do not think that we can settle fundamental moral disputes by simply looking at natural facts.

To (partially) overcome this challenge, STUDY proposed to test the reliability of expert judgments about war indirectly, by testing the level of convergence between experts on questions regarding *in bello* proportionality. The idea underlying this method is that reasonable convergence among experts is a necessary condition for the reliability of their judgments. Needless to say, a reasonable and even perfect convergence is not *sufficient* to establish the correctness of the particular converged upon view. Consensus about falsehoods is possible. However, a *lack* of (reasonable) convergence does seem sufficient to undermine claims for knowledge and for expertise.

Who, then, has the expertise with regard to proportionality judgments concerning war? STUDY assumed that if there are experts in this domain, they are researchers working on the ethics and/or the law of war.⁶ STUDY obtained close to three hundred (N = 289) responses from such academic experts in mainly 10 countries.⁷ In addition, STUDY included a sample of military officers (N = 234) from two countries (US and Israel) and a representative sample of the US population (N = 960). Our focus in this article will be on the sample of academic experts.⁸

Participants in STUDY were presented with descriptions of military operations in the context of a conventional war between two countries, where they had to determine the maximum permissible number of collateral civilian casualties in two hypothetical examples. The first vignette concerned a strategic target: the enemy headquarters, and

the second, a tactical one: a relatively small military airbase. To avoid political bias, the countries in the vignettes were not real ones but labeled “country A” and “country B”. For simplicity, in this article we focus on the strategic target. The vignette describing it was the following:

Headquarters: Assume that country B unjustly invaded country A with the purpose of annexing a large area of territory over which B falsely claims sovereignty. The war has been going on for two weeks now with substantial casualties on both sides, soldiers and civilians ... Thanks to its success in partially neutralizing the air defense system of Army B, army A is now in a position to attack the main headquarters of army B ... Unfortunately, the headquarters are located close to a densely populated area and cannot be attacked without causing collateral damage to civilians residing nearby ... Army A has to make a decision on this attack today, on the basis of the information mentioned above.

After reading the vignette, respondents were asked whether the target in question was a legitimate military target for attack (regardless of collateral damage). Respondents who considered the target as legitimate – an overwhelming majority of them – were then asked whether they thought it was permissible to attack the target at almost any risk to enemy civilians, at no risk to enemy civilians, or provided that the estimated number of civilian casualties would not exceed a certain number which the subjects were asked to provide. Alternatively, they could choose an option stating that “I cannot offer a reasoned answer to this question.”

The results showed that experts were less likely to choose the extreme responses (either “zero casualties” or “almost any number”) compared to lay respondents. However, among experts who opted for a non-extreme response the distribution of the numerical answers were highly dispersed. The median (max) number of casualties that respondents thought could be risked in the case of the attack on the enemy’s headquarters was 125. Comparing the responses of the 25th and 75th percentiles (interquartile-range), the difference between their responses was 575, *4.6 times the median response*. To conclude, then, STUDY demonstrated the absence of reasonable convergence among experts regarding a paradigmatic case of collateral harm in war.

3. STUDY and moral vagueness

3.1 What should we make of these results?

An initial thought is that moral questions where no reasonable consensus can be achieved even among experts reflect moral vagueness. How exactly to interpret such vagueness – as ontic (Schoenfield 2016), semantic (Sud 2019), or in some other way – is a question to which we cannot do justice here. However, it is important to underline how the claim of vagueness seems even more warranted in the case of proportionality than in other cases used in the literature. While with questions such as at which stage a fetus becomes a human being (Schoenfield 2016, 262–263), one can imagine scientific and philosophical developments that would reduce the grey zone concerning the notion of a fetus. This is not the case with the notion of proportionality. It is very hard to think of any scientific or philosophical progress that could reduce the size of the grey zone regarding the question of how many civilian casualties are acceptable when attacking an enemy’s headquarters.

We leave aside the question of which theory of moral vagueness is best, as well as the question of how this theory bears on the notion of *in bello* proportionality. We stipulate, for the sake of discussion, that there *is* a true answer regarding the exact harm that may be brought about collaterally in cases like HEADQUARTERS, namely, that an omniscient subject who spoke a perfect language *would* be able to provide clear and exact answers to such cases. But since no one possesses such qualities, it seems that no one is in a position to claim that he or she is justified in believing that N , rather than $N+35$ or $N-50$ is the correct number of permissible collateral civilian casualties in a given military operation. While there is a wide consensus concerning the permissibility of attacks involving very few civilian casualties and about the impermissibility of attacks involving extremely large numbers of enemy casualties, there is no such consensus with cases between these extremes. The numbers taken from STUDY look more like *guesses* than like the application of some well-defined criterion.

Since reasonable convergence of opinion is a necessary condition for defining expertise in a particular field, it seems that in the field of *in bello* proportionality there are no experts. Given the endemic disagreement between experts about the right balance between the value of attacking a military target and the harm that will be caused to enemy civilians, claims to expertise in this field are merely a pretense. There is no basis for believing that people who teach and write on the ethics and law of war are more qualified than non-experts to provide reliable answers to questions regarding the legitimacy of attacks involving collateral harm to civilians. More accurately, such people are no doubt “academic” experts – they are much more acquainted with the relevant moral and legal texts than non-experts and much better in understanding the history and theory of the *in bello* rules – yet this sort of expertise does not make them experts in making actual judgments about real-life *in bello* dilemmas.⁹ Regarding such dilemmas, it would seem that there is no reason to defer to them rather than anyone else.¹⁰

There is a caveat to this conclusion. STUDY demonstrated that (academic) expertise makes *some* difference because it reduces the likelihood of choosing extreme options – “no casualties” or “any number.” But we expect experts (in actual *in bello* judgments) to provide better advice than a nebulous general claim that “it is okay to bring about the death of civilians as a side effect, though not *any* number of civilians.”

Finally, note that STUDY’s method for undermining the reliability of experts in proportionality judgments is different from that used in the last decade or so in order to undermine the moral expertise of philosophers, a method based on showing that philosophers’ intuitions are as susceptible as non-philosophers to cognitive and other biases.¹¹ Although STUDY also finds evidence for such biases, it remains neutral concerning the general conclusion that might follow regarding the reliability of philosophical intuitions.¹² The challenge STUDY raises about the expertise of moral philosophers is not based on their (equal) susceptibility to various biases, but on the fact that they cannot reach a reasonable consensus in questions regarding proportionality.

4. Peer disagreement

The philosophical debate most relevant to the analysis of STUDY concerns peer disagreement. Most examples of such disagreement involve a small number of people who

disagree on some issue, which leads them to question the reliability of their judgments. Here we find the “Equal Weight View,” which holds that such questioning is indeed warranted.¹³ If A has no reason to believe that she is epistemically superior to her interlocutors with regard to some question, then, given that both have the same training, are intelligent to the same degree, and have the same (relevant) knowledge, it would be irrational for A to stick to her judgment with the same confidence that she had prior to realizing that her epistemic peers thought differently. According to Sara McGrath, given this state of affairs, “the reasonable course is to suspend judgment about whether *p* until further evidence becomes available” (McGrath 2008, 93). And if there can be no claim to knowledge, then neither can there be a claim to expertise.¹⁴

The debate about the Equal Weight View is still ongoing with no clear-cut determination. Our purpose here is to show that STUDY provides a particularly powerful case of peer disagreement against which the standard objections to the Equal Weight View fail. If we are right, then STUDY can be seen as the paradigm case of peer disagreement, one for which the Equal Weight View is the most convincing.

a. Denying peerhood

The most natural response to the Equal Weight View is to concede that in cases of true epistemic peerhood we should either abstain from judgment or at least lower our confidence in our abilities to pass judgment. However, we then have to note that, in real life, there are hardly any cases that fit this description (Elga 2007). In real life, after all, we can almost always rely on some non-arbitrary consideration to justify the thought that regarding the question in dispute we are better qualified to judge than our interlocutors; perhaps because we have read a bit more on the subject, or because we are better at avoiding bias, or because we are not personally invested in our judgments, or because our interlocutors are making their judgments when tired or distracted, and so on. And in the absence of positive evidence showing that we are epistemically equal to our interlocutors (regarding the question in dispute), we would be justified in believing that we have the better judgment.

However, this response is unavailable in cases like STUDY. The group of interlocutors is very large and we know nothing about the specifics of their education, intelligence, or political views. All we know about them is that they are experts in the sense described above; they publish academic research on the law and the ethics of war. We might think their judgments incorrect as we are more intelligent, but we have no good grounds for believing this to be true of *all* the experts. And although we might be better than many other respondents in some of the epistemic qualities that are pertinent to the issue of proportionality, it is no less likely that we are *worse* than many of them in these qualities. These possibilities, as it were, cancel out each other, leaving us with no basis for believing that we are epistemically superior to the many respondents who hold different judgments about a paradigmatic dilemma such as how many civilians it would be morally permissible to kill in an attack on the enemy headquarters.

b. Relying on a private intuition

Peter van Inwagen (1996) has argued that, in cases of peer disagreement, it is sometimes rational to rely on one’s own private insight that *p*, even if one’s equally intelligent and

knowledgeable peers believe not- p . The basic idea is that if we have access to a special body of evidence to which our peers lack access then this disagreement needs not give us a reason to think that we are in error, or to lose confidence in our judgment.

The standard response to this argument points out that this reasoning goes both ways. If we have reason to trust our epistemic superiority on the basis of some kind of private evidence or incommunicable insight, then our interlocutors may also have it, which means that we lack justification for believing that our insight that p is more reliable than our peers' insight that not- p (Feldman 2006).

This response is well-taken, yet it is more convincing in some contexts than in others. In disagreements about the existence of God, for instance, the believer might have a stronger basis for trusting her subjective experience because usually non-believers do not have an analogous atheist subjective experience or insight. So in the religious context, the van Inwagen approach does seem to provide subjects with some justification to trust their judgments. However, this argument does not apply to STUDY. Those saying that the attack on the headquarters justifies only, say, ten civilian casualties, cannot plausibly claim that their judgment is based on some kind of "private experience," which their peers, who think that the attack justifies fifty or one hundred casualties, assumingly lack. In cases like STUDY, there is nothing analogous to a religious or mystical experience that could ground such a move.

Perhaps some respondents have had battlefield experience, which might have given them an insight of the sort van Inwagen has in mind when thinking about what proportionality in warfare means in practice. But it is safe to assume that the overwhelming majority of the academic experts who participated in STUDY had no such experience. Their judgments about proportionality came from the same philosophical and legal texts studies (on average) by all of them.

c. The right to trust our intuitions

The last objection to the Equal Weight View focuses on the right to self-trust (see, for instance, Wedgwood 2010 and Schafer 2015). According to Wedgwood, "it is rational to have a special sort of 'fundamental trust' in one's own intuitions, but it is not even possible to have the same sort of 'trust' in the intuitions of others" (Wedgwood 2010, 20). For the sake of argument, we concede that sometimes it is rational to maintain our confidence in our judgments even when we realize that people who look like our epistemic peers think otherwise (Rattan 2014). But sometimes it is clearly irrational to do this – and cases like STUDY provide a paradigmatic example. If many people, whose intelligence, knowledge, and moral sensitivity we cannot doubt, disagree with me, and if we cannot claim that our judgment is based on either a sophisticated argument that only we comprehend, or on some private insight to which only we have access, then there is no basis for privileging our own judgment over that of others.

There is another argument against the self-trust response, which again seems especially pertinent to cases like STUDY. Numerous studies have demonstrated the unrealistic confidence of people in their achievements, qualities, chances of success and so on (for a brief summary, see Thaler and Sunstein 2008, 31–33). This confidence, as Thaler and Sunstein note, "is a pervasive feature of human life; it characterizes most people in most social categories" (33). Thus, when people are confident about themselves,

a likely supposition is that they are *over*-confident.¹⁵ Their subjective confidence should not be taken as indicating some deep insight to which only they have access, but as indicating unrealistic optimism about their epistemic capacities.

All the more so when one realizes that a great number of evaluators who are most probably (on average) similarly intelligent and morally sensitive as oneself give significantly different answers to a particular question, as is shown in STUDY. In such circumstances, only extreme vanity or delusion would enable the maintenance of “fundamental trust” (in Wedgwood’s terms) in one’s own intuitions. STUDY provides an excellent illustration of Parfit’s point regarding the significance of disagreement between many people in ideal conditions:

If we had strong reasons to believe that, even in ideal conditions, we and others would have deeply conflicting normative beliefs, it would be hard to defend the view that we have the intuitive ability to recognize some normative truths. We would have to believe that, when we disagree with others, it is only we who can recognize such truths. But if many other people, even in ideal conditions, could not recognize such truths, we could not rationally believe that we have this ability. How could we be so special? And if none of us could recognize such normative truths, we could not rationally believe that there are any such truths. (Parfit 2011, 546)

In response to our support for the Equal Weight View in cases of peer disagreement between many subjects, like in STUDY, one might argue that we should assign more weight to the quality of arguments underlying the different views. If experts reach starkly opposite conclusions within a difficult field such as *in bello* proportionality, we should look at their arguments and compare them to each other, which might make us see that some views are more persuasive than others; that they are not after all equal. However, when it comes to proportionality judgments, once there is agreement on the military value of an attack, one rarely finds *arguments* purporting to show that this attack justifies a collateral harm of twenty, fifty or a hundred civilian casualties. The experts can agree on all relevant premises – that the military target is very valuable, that it nonetheless must be balanced against the harm to civilians, and that if the harm is excessive the attack should be called off – and still radically disagree about how many civilian deaths would be excessive relative to the value of attacking the enemy’s headquarters. Debates on such questions are not typically about arguments, exploring the truth of premises or the logical connection between premises and conclusions, but about *intuitions*. Some intuit that given the military importance of attacking the headquarters, 100 casualties is an acceptable cost. Others intuit that any number above twenty would be excessive. None of them offer *arguments* to support these verdicts. None *can* offer such arguments.

To sum up, STUDY provides a strong example of peer disagreement for which the Equal Weight View seems the most fitting approach. What follows is that experts working on the law and ethics of war – and *a fortiori* non-experts – must acknowledge that in questions regarding *in bello* proportionality there is an extremely wide spectrum of reasonable answers, and that the confidence they have in their particular judgment *within* that spectrum should be seriously reconsidered.

5. Normative implications

The insights from the previous section is that while in extreme cases of either a very few civilian casualties or very many such casualties, experts can offer reliable guidance, while

in less extreme cases – *which are typical of most dilemmas in this area* – they are in no position to do so.

One reaction to STUDY would be to see it as confirming the realist position about morality and war. For realists, the unreliability of the proportionality constraint would be taken as yet another indication of just war theory's failure in regulating war, and a confirmation of the fundamental realist view that *inter arma silent leges*. But STUDY does not entail this implication because exposing the unreliability of the proportionality test is insufficient to demonstrate the unreliability of other *in bello* principles, mainly that of non-combatant immunity. Also, as pointed out earlier, STUDY does not reject the notion of proportionality when facing extreme cases. STUDY, then, does not offer support for realism.

An opposite reaction would be to see STUDY as confirming pacifism. Given the strong moral presumption against killing people, especially against killing the innocent, there is a heavy onus of justification on whoever wishes to knowingly bring about the death of civilians in war (more accurately, to knowingly put innocent lives at serious risk). Since STUDY shows that this onus cannot be met, the result is that attacks predicted to kill civilians are allowed only in the extreme case of very few casualties, otherwise they are morally ruled out. If you are not reasonably sure that (collaterally) killing civilians is permissible, then, morally speaking, the only option open to you is to refrain from the killing. In contemporary conflicts, in which the boundaries between the battlefield and civilian life are very much blurred, this result would in fact entail pacifism, or a position close to it.

We obviously cannot enter here into a general discussion of pacifism. For the sake of the present article let us just say that since most philosophers and, surely, most lay people and politicians are *not* pacifists, this reaction to the results of STUDY is not a real option for them.

Is there, then, a non-realist, non-pacifist way of justifying military attacks that are predicted to bring about the death of civilians without sufficient certainty that the killing is proportionate? Benbaji and Statman (2019) have recently offered a contractarian answer (142–152). According to contractarianism, since allowing collateral harm only in extreme cases would significantly impede the ability of states to fight and win just wars, the parties on the international plain undertake a rule that permits attacks even in non-extreme cases. The parties are allowed to act on the basis of their best judgment about the proportion between the military value of the attack and the expected harm to civilians in spite of STUDY demonstrating that such judgments are unreliable. Such a rule is *ex ante* mutually beneficial to all parties, which means that it would (again, *ex ante*) benefit all civilians. Hence, civilians authorize their states to give up, on their behalf and on the basis of mutuality, their right to protection from collateral harm in times of war – except in cases in which the harm is truly excessive, which, in STUDY's terms, would mean cases where there is a reasonable convergence among (academic) experts.

In response, one might argue that even from a non-pacifist (contractarian or other) view there is good reason for extra restraint when it comes to (collaterally) harming civilians. Since warring parties tend to assign too *little* weight to the rights of enemy civilians, what they need to do in order to correct this tendency is to err, so to speak, in the other direction and use *extra* restraint when they foresee harm. What this advice would mean

in practice is that after making a judgment that N civilian casualties is proportionate relative to the value of the attack, the parties should revise the judgment to $N + X$, undertaking stricter self-restraint than that required by their initial judgment. But erring on the side of caution is not always the right policy, especially not when extra caution might interfere with a credible and effective defense or with the protection of rights. In individual self-defense, the demand for extra caution which is not mandated by the facts and the accepted doctrines might mean that a would-be victim is prevented from effectively protecting herself from unjust aggression. This would also be true at the collective level. Demanding that armies reduce the risk to civilians beyond what they deem as proportionate when executing attacks on valuable military targets, such as their enemy's headquarters, comes at a significant price. It compromises their ability to conduct the war and their duty to protect their fellow citizens. It also gives ruthless parties a huge advantage because by locating their soldiers or military bases among civilians they could make their enemies hesitate to attack.

The *in bello* permission to collaterally kill enemy civilians does not mean a sweeping permission to do so any time the predicted harm is not truly excessive. As just mentioned, in every case soldiers are required to strike a balance between the military value of the attack and the lives of the enemy civilians that are put at risk. The problem is that after STUDY, it is no longer clear what this requirement means in practice. If even experts do not have a reliable way of figuring out what *in bello* proportionality means in practice, what do we mean when we impose such a requirement on officers and soldiers? We seem to have reached a deadlock.

We believe the way out of the deadlock is to think about such balancing through the prism of virtue ethics. In contrast to rule-centered or act-centered ethics, virtue ethics offers a different way of thinking about the problem. For virtue ethics, (a) judgments about the value of character traits are independent of judgments about the rightness or wrongness of actions, (b) such judgments about character (“aretaic judgments”) are explanatory prior to judgments about right conduct (see Statman 1997), and (c) there is often more than one set of behaviors that is compatible with being a person of virtue (or with having a particular virtue, such as benevolence, compassion or courage). Two virtuous people could, therefore, respond differently to similar circumstances without this showing that either of them was ill-informed or was guilty of an error in moral judgment. The notion of virtue is thus largely indeterminate when it comes to guiding behaviour.¹⁶

According to virtue ethics, however, this indeterminacy does not mean that virtues are of no use when making practical decisions. What will and should shape our behavior is our moral character, which will rule out many courses of action. As Rosalind Hursthouse argues in her influential paper on virtue ethics and abortion (1991), while some cases of abortion are clearly vicious, callous or selfish, others manifest benevolence and caring. In between are other situations where different (virtuous) women could make different decisions without any of them being morally mistaken. According to this view, morality is not about developing a decision procedure that could yield a correct answer to all practical dilemmas, but about inculcating virtue so that people feel the moral emotions “at the right times on the right grounds towards the right people for the right motive and in the right way [...] Similarly, there are excess and deficiency and a mean in the case of action” (*Nicomachean Ethics*, Book 2, 6, 1106b 22-24; in Aristotle 2004, 41). For those who regard

this advice as insufficient and dream of a Dworkin-like Hercules (Dworkin 1977, 105) who could offer them a more precise and implementable device for moral decision, Aristotle would say that we should not look for precision in all things alike, but “only for as much as the subject-matter in each case allows and so far as is appropriate to the investigation” (*Nicomachean Ethics*, Book 1, 7, 1098a 26–27; in Aristotle 2004, 17). The expectation that moral theory provides a way of determining what exactly is required by courage in every situation where there is great risk reveals a deep misunderstanding of the subject-matter of ethics.

Turning to soldiers in decent armies, Nancy Sherman (2005) has shown the Stoic roots of martial virtues such as resilience, courage and self-sufficiency, and has explained how, for the Stoics, these virtues were balanced by respect and empathy towards all human beings. In Cicero’s words, “we must exercise a respectfulness towards men, both towards the best of them and also towards the rest” (*On Duties* 1.99; cited by Sherman 2005, 171). Such respectfulness, Sherman argues, should be part of the honor code of any warrior (Sherman 2005, 179). The correct and efficacious ethical training of soldiers does not consist of equipping them with principles that they could then “apply,” but in inculcating in them two kinds of virtues: virtues of resilience that will help them to overcome fear and act jointly to defeat their enemies, and virtues of respectfulness that will help them not to lose sight of the fact that those who are harmed by them are fellow human beings. If this education succeeds, soldiers will be in a position to better handle the dilemmas concerning collateral harm virtuously. Yet again, what this exactly means cannot be determined by some independent criterion, and might differ from one (virtuous) soldier to another.¹⁷

This ethical education might be seen by some as insufficient for providing protection to civilians in times of war. We do not share this worry because, in our view, ethical education is much more effective in influencing behavior than the teaching of abstract principles. But for those who do have this worry, we raise two points that might lessen it:

First, STUDY deals only with collateral harm to civilians and not with deliberate harm. Nothing in what was argued above undermines the validity or the practicality of the blanket prohibition against *intentional* attacks on civilians. Although there are some grey zones with regard to this prohibition as well – for instance, in cases of “supreme emergencies” – it is, by and large, clear and easy to implement.

Second, the criterion of *in bello* proportionality is not the only restriction on unintentionally yet foreseeably harming civilians. Walzer famously adds another restriction, namely, that the attackers select the least harmful measure required to achieve their legitimate aims, and make a genuine effort to minimize harm to civilians:

Simply not to intend the death of civilians is too easy ... What we look for in such cases is some sign of a positive commitment to save civilian lives ... War necessarily places civilians in danger; that is another aspect of its hellishness. We can only ask soldiers to minimize the dangers they impose. (Walzer 1977, 155–156)

We should add in the light of the insights from STUDY, one cannot take for granted that there will be a better convergence of views concerning necessity judgments among experts any more than we find with proportionality judgments. The perception that the necessity condition is more reliable relies on the mistaken thought that what it requires is to make sure that the means selected for the attainment of some military

goal are the least harmful, namely, that if some goal, g , can be achieved in N ways, and $N1$ is the least harmful, then $N1$ ought to be selected. But in real life the dilemma is often not between different ways of achieving the same goal, but between different ways of action that are predicted to achieve different goals. Often the number of collateral civilian casualties cannot be reduced if one wishes to effectively attack some military target $T1$, but if a different target is selected, $T2$, which is militarily less valuable than $T1$, that *would* result in fewer casualties. In dilemmas of this kind which are not uncommon in war, comparing these two packages is no less challenging than making proportionality judgments.¹⁸ We would not be surprised if a future study, analogous to STUDY, demonstrated that some *in bello* necessity judgments suffer from a similar lack of convergence to that demonstrated in STUDY with regard to the proportionality condition.¹⁹

Finally, what can STUDY teach us about the widespread practice of condemning soldiers, armies and countries for launching attacks that are purportedly disproportionate? What follows from STUDY is that such condemnation is no more reliable than the military judgment of the *in-bello* proportionality in question. If one's belief that p is close to a mere guess, there is no basis for one to condemn others for holding not- p and for acting accordingly. Justifiable condemnation of disproportionate attacks exists only in extreme cases in which the harm to civilians is *consensually excessive*.

The reliability of condemnations on the basis of assumed disproportionality is further undermined by a recent empirical study that demonstrated how one's view on the *in bello* conduct of soldiers is strongly influenced by one's view on the *ad bellum* justification of the war the soldiers are fighting (Watkins and Goodwin 2020). This strengthens the suspicion that when critics complain about *in bello* violations of the proportionality condition, they are biased by their general opposition to the war, or by the fact that the assumed disproportionate attacks were taken against their own national group. If they had sympathized with the general goals of the war, or if the attacks had been carried out by *their* side to the conflict, the critics would have probably regarded them as proportionate.

6. Concluding remarks

- 1 STUDY demonstrates the lack of reasonable consensus among moral and legal experts concerning what proportionality in war requires in practice, outside of extreme cases. Since there is no further information the discovery of which could lead to a stronger consensus, and since there seem to be no arguments that could establish a more precise way of balancing military value and respect for the lives of enemy civilians, this lack of consensus seems intractable.
- 2 This lack of convergence is a strong case of peer disagreement for which the Equal Weight View seems most fitting. If this is correct, it seems that STUDY should encourage those who make judgments about *in bello* proportionality either to abstain from such judgments, or at least to question seriously their confidence in them. This moderated confidence should lead to the recognition that, in many cases, condemning states or armies for assumed disproportionate use of force is unwarranted and should be significantly reduced. *A fortiori* bringing legal charges against officers or soldiers for attacks on military targets where the harm caused to

- civilians did not cross a threshold where there is a consensus that this was extreme is rarely appropriate or fair.
- 3 If even experts in just war theory cannot know what proportionality demands in practice in the “in-between zone,” their proportionality judgments are probably often influenced by their prior political/ideological commitments and interests. If some attack is perceived by them as proportionate and therefore legitimate, this is often because of their general sympathy for the attacking side and for the morality of its cause. Similarly, if an attack is conceived as disproportionate, this often has to do with their general opposition or hostility to the attacking party. In both cases, those offering proportionality judgments should acknowledge that they are on shaky grounds. Among other things, such acknowledgment might contribute to moderating the combative nature of most debates about the morality of armed conflicts.
 - 4 Despite the unreliability of proportionality judgments, this *in bello* requirement ought not to be abandoned. It is best understood in the context of virtue ethics rather than in terms of act- or rule-ethics.²⁰ Proportionality is not a decision procedure that can be taught in a way that would yield clear answers to dilemmas concerning collateral harm to civilians, but a martial *virtue* that must be inculcated among soldiers. A virtuous soldier is (among other things) one who does not lose sight of the fact that enemy civilians are human beings just like herself and their lives need to be taken into consideration when planning attacks. Like with other virtues, what exactly this means in practice cannot be spelled out precisely and depends on the specific details of the case at hand. To follow Aristotle’s formulation about justice in *Nicomachean Ethics* (Book 2, 4), acts are called proportionate when they are such as the “proportionate” soldier would do, not the other way round. This understanding of proportionality takes the sting out of the peer disagreement problem because two agents might both be virtuous, yet respond differently to dilemmas they face. There is more than one way to be compassionate or courageous, and there’s more than one way to express one’s respectfulness for enemy civilians while attempting to win the war.
 - 5 A further advantage of a virtue ethics reading to proportionality is that it helps to block an overly permissive interpretation of STUDY, along the lines that since even experts do not know what proportionality requires in practice, officers may ignore the potential harm to enemy civilians and – with the exception of truly excessive cases – do whatever it takes to fulfill the military mission and to safeguard the lives of their soldiers. The virtue ethics response to this approach would be that this view is *callous*; it expresses a lack of appropriate sensitivity to the lives of innocent people on the enemy side. Although virtue ethics offers no algorithm to calculate the exact maximum number of civilians that may be killed in military attacks, it mandates taking such dilemmas seriously, not lightheartedly.
 - 6 Finally, STUDY raises serious doubts about the idea that law professors or moral philosophers should be deferred to on questions concerning *in bello* proportionality. In doing so, STUDY strengthens a prominent line of research in the last decade that challenges the view that, due to their professional expertise, the moral intuitions of philosophers are more trustworthy than those of non-philosophers.²¹ To be sure, philosophers (and law professors) might bring to the table useful tools for analyzing the dilemmas at hand, but when it comes to making *judgments*, such as those concerning *in bello* proportionality, they fare no better than laypeople.²²

Notes

1. The literature on the topic is extensive, part of it dealing with *in bello* proportionality in general, and part dealing with the proportionality of specific tactics or operations. For a selective sample, see Hurka 2005; Haque 2015; Tadros 2018; and Tomlin 2019.
2. Protocol I, Article 57(2) (a)(iii).
3. See, for instance, Andreson 2014, 18–32; Franck 2010, 230–242; and Haque 2019, 118–160.
4. To date, we are not familiar with any methodological or philosophical criticisms of STUDY.
5. The debate about the rationality and usefulness of the proportionality principle is rich and complex. See most recently Chang and Dai 2021 and the responses to them by Peters 2021 and Caviedes and Urbina 2021.
6. 47% of the experts who responded were from philosophy, 46% from Law, 33% from Political Science and 20% from other fields. Some were from multiple fields.
7. Australia, Canada, France, Germany, Israel, Netherlands, Singapore, Sweden, United Kingdom, and United States.
8. The study among experts was conducted in 2015.
9. In a recent empirical study among more than 4,000 philosophers from 96 countries, Niv & Sulitzeanu-Kenan (2022) found that approximately 50% of them support 'the broad view', according to which philosophers have an enhanced ability to know what is morally right. If our interpretation of STUDY is correct then this self-assessment – at least regarding proportionality judgments – is an illusion.
10. For the general phenomenon of experts being unaware of the limits of their expertise, see Kahneman 2011, ch. 22.
11. See, for example, Schwitzgebel and Cushman 2012; Mizrahi 2015; Nado 2015; Rini 2015; Schwitzgebel and Cushman 2015; and Drodowicz 2018.
12. For discussions of this conclusion, see see Rini 2015 and Demaree-Cotton 2016
13. For defenders of the Equal Weight View see, for instance, Bogardus 2009 and Matheson 2015.
14. See also Cross 2016. According to Cross, to the extent that moral philosophers deeply disagree among themselves, especially as members of various ethics committees, "the fact of disagreement entails that there are no good reasons for people to defer to even the most highly credentialed moral philosophers" (188).
15. Given the tendency of people to overestimate their qualities, achievements, and value more generally, some philosophers have suggested that modesty consists not in holding a lower self-assessment than one merits (which would make modesty a form of ignorance), but in having an *accurate* self-assessment, i.e., in overcoming the temptation of overestimation. See Flanagan 1990.
16. According to Adams, this indeterminacy has to do with viewing virtue as a kind of goodness rather than a kind of rightness. In his view, this "makes it easier to see how there can be quite different alternative ways of being genuinely virtuous" (2006, 11).
17. The virtue approach we are proposing is already incorporated in military training, for instance, in the British army with its emphasis on the displaying of six virtues; courage, discipline, respect for others, integrity, loyalty, and selfless commitment (<https://www.army.mod.uk/who-we-are/our-people/a-soldiers-values-and-standards/>) and the American army, emphasizing loyalty duty, respect, selfless service, honor, integrity, and personal courage (for the claim that VE is needed in the US military, see <https://blog.usejournal.com/virtue-ethics-and-the-united-states-army-5702fef93990>). For a recent helpful explication of military virtues, see Skerker, Whetham, and Carrick 2019.
18. Some have suggested, on the basis of such considerations, that the necessity condition, properly understood, can be incorporated into the proportionality one. See Hurka 2005, 37 ("The last resort condition is in effect a comparative version of the initial, simple proportionality condition").
19. For the philosophical debate about necessity in war, see, e.g., Statman 2011; Lazar 2012; and Steinhoff 2019.

20. For other proposals to use (or not use) virtue ethics in order to shed light on the ethics of war, see Chan 2012 and Schulzke 2016.
21. See note 8 above.
22. For a contrary view, according to which philosophers' enhanced ability to analyze, understand, and conceptualize moral problems tends to lead to an enhances ability to make true moral judgments, see Niv (2022). Again, in our view, STUDY undermines this assumed connection.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Stephen de Wijze is Senior Lecturer in Political Theory at the University of Manchester. His recent publications are “Small Scale Evil”, *Journal of Value Inquiry*, 2018, *The Routledge Handbook on the Philosophy of Evil* (co-edited with Thomas Nys, 2019) and “Are Dirty Hands Possible”, *The Journal of Ethics*, 2022. He is currently completing a book on the problem of dirty hands.

Daniel Statman is head of the philosophy department at the University of Haifa and former chair of the Israeli Philosophical Association. In his co-authored book, *War by Agreement* (Oxford University Press, 2019), he defends a contractarian understanding of just war theory. He served on the committee that revised the ethical code for the IDF.

Raanan Sulitzeanu-Kenan is an Associate Professor of Political Science and Public Policy at the Hebrew University, and former Head of the Federmann School of Public Policy. His research interests include political behavior, behavioral public administration and policy, and empirical legal studies.

ORCID

Raanan Sulitzeanu-Kenan  <http://orcid.org/0000-0001-9663-8954>

References

- Adams, Robert M. 2006. *A Theory of Virtue*. Oxford: Oxford University Press.
- Andreson, Joshua. 2014. “Challenging the Perplexity Over Jus in Bello Proportionality.” *European Journal of Legal Studies* 7: 18–32.
- Aristotle. 2004. *Nicomachean Ethics*, Translated by J. A. K. Thomson, Revised by Hugh Tredennick. London: Penguin.
- Benbaji, Yitzhak, and Daniel Statman. 2019. *War by Agreement*. Oxford: Oxford University Press.
- Bogardus, Tomas. 2009. “A Vindication of the Equal Weight View.” *Episteme: A Journal of Social Epistemology* 6 (3): 324–335.
- Caviedes, Cristóbal, and Francisco J. Urbina. 2021. “Cost-Benefit Analysis in Rights Adjudication – An Assessment in Light of the Proportionality Debate: A Reply to Yun-chien Chang and Xin Dai.” *International Journal of Constitutional Law* 19 (3): 1146–1153.
- Chan, David. 2012. *Beyond Just War: A Virtue Ethics Approach*. New York: Palgrave.
- Chang, Yun-chien, and Xin Dai. 2021. “The Limited Usefulness of the Proportionality Principle.” *International Journal of Constitutional Law* 19 (3): 1110–1134.
- Cross, Ben. 2016. “Moral Philosophy, Moral Expertise, and the Argument from Disagreement.” *Bioethics* 30 (3): 188–194.
- Demaree-Cotton, Joanne. 2016. “Do Framing Effects Make Moral Intuitions Unreliable?” *Philosophical Psychology* 29 (1): 1–22.

- Drodowicz, Anna. 2018. "Philosophical Expertise Beyond Intuitions." *Philosophical Psychology* 31 (2): 253–277.
- Dworkin, Ronald. 1977. *Taking Rights Seriously*. London: Duckworth.
- Elga, Adam. 2007. "Reflection Disagreement." *Noûs* 41 (3): 478–502.
- Feldman, Fred. 2006. "Reasonable Religious Disagreements." In *Philosophers Without Gods: Meditations on Atheism and the Secular Life*, edited by L. Antony, 194–215. New York: Oxford University Press.
- Flanagan, Owen. 1990. "Virtue and Ignorance." *Journal of Philosophy* 87 (8): 420–428.
- Franck, Thomas M. 2010. "Proportionality in International Law." *Law & Ethics of Human Rights* 4 (2): 230–242.
- Haque, Adil Ahmad. 2015. "A Theory of Jus in Bello Proportionality." In *Weighing Lives: Combatants & Civilians in War (ch. 8)*, edited by Jens David Ohlin, Larry May, and Claire Finkelstein. Oxford: Oxford University Press.
- Haque, Adil Ahmad. 2019. "Indeterminacy in the Law of Armed Conflict." *International Law Studies* 95: 118–160.
- Hurka, Thomas. 2005. "Proportionality in the Morality of War." *Philosophy and Public Affairs* 33 (1): 34–66.
- Hursthouse, Rosalind. 1991. "Virtue Theory and Abortion." *Philosophy & Public Affairs* 20 (3): 223–246.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Lazar, Seth. 2012. "Necessity in Self-Defense and War." *Philosophy & Public Affairs* 40 (1): 3–44.
- Matheson, Jonathan. 2015. "Disagreement and the Ethics of Belief." In *The Future of Social Epistemology: A Collective Vision*, edited by James Collier, 139–148. Lanham: Rowman and Littlefield.
- McGrath, Sara. 2008. "Moral Disagreement and Moral Expertise." In *Oxford Studies in Metaethics, Volume 3*, edited by Russ Schafer-Landau, 87–108. Oxford: Oxford University Press.
- Mizrahi, Moti. 2015. "Three Arguments Against the Expertise Defense." *Metaphilosophy* 46 (1): 52–64.
- Nado, Jennifer. 2015. "Philosophical Expertise and Scientific Expertise." *Philosophical Psychology* 28 (7): 1026–1044.
- Niv, Yarden. 2022. "Beyond All-or-Nothing Approaches to Moral Expertise." *Journal of Applied Philosophy* 39: 282–296.
- Niv, Yarden, and Raanan Sulitzeanu-Kenan. 2022. "An Empirical Perspective on Moral Expertise: Evidence from a Global Study of Philosophers." *Bioethics*. doi:10.1111/bioe.13079.
- Parfit, Derek. 2011. *On What Matters*. Oxford: Oxford University Press.
- Peters, Anne. 2021. "A Plea for Proportionality: A Reply to Yun-Chien Chang and Xin Dai." *International Journal of Constitutional Law* 19 (3): 1135–1145.
- Rattan, Gurpreet. 2014. "Disagreement and the First-Person Perspective." *Analytic Philosophy* 55 (1): 31–53.
- Rini, Regina. 2015. "How Not to Test For Philosophical Expertise." *Synthese* 192 (2): 431–452.
- Schafer, Karl. 2015. "How Common is Peer Disagreement? On Self-Trust and Rational Symmetry." *Philosophy and Phenomenological Research* 91 (1): 25–46.
- Schoenfield, Miriam. 2016. "Moral Vagueness Is Ontic Vagueness." *Ethics* 126 (2): 257–282.
- Schulzke, Marcus. 2016. "Rethinking Military Virtue Ethics in an Age of Unmanned Weapons." *Journal of Military Ethics* 15 (3): 187–204.
- Schwitzgebel, Eric, and Fiery Cushman. 2012. "Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers." *Mind & Language* 27 (2): 135–153.
- Schwitzgebel, Eric, and Fiery Cushman. 2015. "Philosophers' Biased Judgments Persist Despite Training, Expertise and Reflection." *Cognition* 141: 127–137.
- Sherman, Nancy. 2005. *Stoic Warriors: The Ancient Philosophy Behind the Military Mind*. New York: Oxford University Press.
- Skerker, Michael, David Whetham, and Don Carrick. 2019. *Military Virtues*. Hampshire, UK: Howgate Publishing.

- Statman, Daniel. 1997. "Introduction." In *Virtue Ethics*, edited by David Statman, 1–41. Edinburgh: Edinburgh University Press.
- Statman, Daniel. 2011. "Can Wars Be Fought Justly? The Necessity Condition Put to the Test." *Journal of Moral Philosophy* 8 (3): 435–451.
- Statman, Daniel, Raanan Sulitzeanu-Kenan, Micha Mandel, Stephen de Wijze, and Michael Skerker. 2020. "Unreliable Protection: An Experimental Study of Experts' In-Bello Proportionality Decisions." *European Journal of International Law* 31 (2): 429–453.
- Steinhoff, Uwe. 2019. *Self-Defense, Necessity, and Punishment: A Philosophical Analysis*. New York: Routledge.
- Sud, Rohan. 2019. "Moral Vagueness as Semantic Vagueness." *Ethics* 129 (4): 684–705.
- Tadros, Victor. 2018. "Past Killings and Proportionality in War." *Philosophy & Public Affairs* 46 (1): 9–35.
- Thaler, Richard, and Cass Sunstein. 2008. *Nudge*. New York: Penguin.
- Tomlin, Patrick. 2019. "Subjective Proportionality." *Ethics* 129 (2): 254–283.
- van Inwagen, Peter. 1996. "It is Wrong, Always, Everywhere, and for Anyone, to Believe Anything, Upon Insufficient Evidence." In *Faith, Freedom, and Rationality*, edited by J. Jordan, and D. Howard-Snyder, 137–154. Lanham: Rowman and Littlefield.
- Walzer, Michael. 1977. *Just and Unjust Wars*. New York: Basic Books.
- Watkins, Hanne, and Geoffrey Goodwin. 2020. "A Fundamental Asymmetry in Judgments of Soldiers at War." *Journal of Experimental Psychology: General* 149 (3): 419–444.
- Wedgwood, Ralph. 2010. "The Moral Evil Demons." In *Disagreement*, edited by R. Feldman, and T. Warfield, 216–246. Oxford: Oxford University Press.