REPLY

Paul H. Robinson

As a preliminary matter, consider John Mikhail's comment, in which he agrees with the general approach to justification proposed but objects to one aspect: what he says is its "assumption that objective justification should be equated, simply and without further qualification, with conduct that actually avoids the greater harm." He criticizes the Model Penal Code for taking the same approach.

Unfortunately, his comment is based on a false premise. As my treatise and other writings make clear, it is not my view that the balance of interests in justification defenses is limited to a comparison of the tangible harms. Indeed, it is commonly the intangible interests that are determinative.¹ In the standard self-defense case, for example, the competing tangible harms are a life for a life. It is the intangible interest of society's abhorrence of aggression that tips the balance in favor of allowing force against the aggressor. Mikhail does the same disservice to the Model Penal Code. The drafters have no intention of limiting the balance of interests to tangible harms, even under the broad balancing of Section 3.02.² That is why the provision uses the phrase "harm or evil" (indeed, the section is titled "Lesser Evils") and why the commentary speaks of "weighing conflicting values" — and why I was careful to use the phrase "harm or evil" in the first paragraph of my Core Text.

Mikhail may be forgiven his error, however, because it is a common one, even among serious criminal law theorists, including here both Mitch Berman (his last paragraph) and Peter Westen (his third paragraph).

Turning to the main points of attack of the objective theory of justification, consider the responses in terms of the two kinds of arguments the Core Text makes in support. First, it argues that the distinction between actual, objective justification and subjective, mistaken justification is an important one. This is so because ignoring the distinction, as the subjective theory of justification does, makes it more difficult both (1) to announce clear ex ante rules of conduct and (2) to define the conditions triggering a right to use defensive force.

Shlomit Wallerstein appears to agree that the actual-mistaken distinction is an important one and that mistaken actors ought to be excused, not justified. Westen also agrees but rightly points out that the resulting benefits can be realized only if the system adopts special verdicts that advertise the distinction in case adjudications. I have urged such special verdicts.³

Berman disagrees that the objective theory makes easier the law's ex ante function of announcing rules of conduct. He makes the standard Type 1 error, discussed in Section III.A. —

¹ See, e.g., Paul H. Robinson, Criminal Law 408, 422-3, 428, 443 (1997). And some would argue that some intangible interests are so important that they can never be outweighed.

² See Model Penal Code Section 3.02 comment, esp. 6, 17 (1985).

he fails to appreciate that future actors will not know of their errors at the time they act, thus will use the objective, actual justification rules as their guide, not rules relating to mistaken actors. As to the problem of an effective triggering condition for defensive force, he appears to concede that the Model Code's subjective approach has a problem, but argues that it can be improved, offering a rewrite. But remember that this rule must serve as a rule of conduct for laypersons facing threats, thus tolerates no complexity or intricacy. His rewrite is as ineffective in performing this function as Model Code language. What is required here is the simple, almost intuitive rule that only the objective theory provides: "one can defend against unjustified conduct."

The second claim of the Text is that the objective theory of justification ought to be preferred over the subjective theory because the former gives correct liability assessments — attempt liability for the unknowingly justified actor, no defense for knowingly resisting an unknowingly justified actor, and no liability for knowingly assisting an unknowingly justified actor — while the latter gives incorrect results. Its results are "correct," the Text argues, because they are consistent with community intuitions and provide the correct incentives to advance societal interests.

Westen agrees that the objective theory results are preferable. His complaint is that I don't give enough credit to moral philosophy, where some have long supported these results, such as Plato.

Wallerstein, who supports the "mixed" theory of George Fletcher, in which justification requires that both objective and subjective requirements be satisfied, prefers contrary results but offers no explanation for why they ought to be preferable. She would have the law authorize the farmer to prevent the firebreak that the farmer knows is needed to save the townspeople, a result contrary to common sense, if not more.

Berman concedes that the objective theory has the right results with regard to assisting and resisting the unknowingly justified actor, but argues that the subjective theory also could come to these results. The Text, he complains, "misdescribes the outcomes a subjective formulation directs." This seems something less than fair, however, given that the Text describes the results that people have always assumed the subjective Model Code provides. It turns out that what Berman is really suggesting is that one could construct a special rule to alter the Model Code's results to match those of the objective theory.

With regard to liability for the unknowingly justified actor, in contrast, Berman would create no special rule and defends the Model Code's full liability. However, he offers no explanation for why people's intuitions and society's interests should be sacrificed. Indeed, even under the deontological analysis that he seems to prefer, the unknowingly justified actor is in a morally analogous position to the impossible attempter. Neither actor has injured society (indeed, the unknowingly justified actor has benefitted it) but both mistakenly believe that they have. On what grounds can one justify punishing the unknowingly justified actor more?

Note also the peculiarity of Berman's position here: the unknowingly justified actor is not justified, while the person assisting him is justified. So much for Berman's earlier claim that his theory would do better in announcing an ex ante rule of conduct. What rule of conduct is one to draw from these conflicting dispositions about the permissibility of the burning?

A broad view of the debate suggests that the best critics can do is to argue that the
subjective theory of justification is not quite so bad as claimed or that, with some special rules, its problems can be minimized or fixed. But why bother?

What is missing in the attacks is an argument for why the subjective theory should be preferred despite its problems. In the end, I suspect that support for the subjective theory lies in the Type 3 error, discussed in Section III.C — that people commonly start with a premise that justification must be subjectively defined because it has been so in some previously-existing literature, then set themselves to defend that unreasoned premise.