7

Why Victims' Rights are Irrelevant to Paradigmatic Justifications

MARK DSOUZA

1. Introduction

Responding to reports of a dangerous gunman in a residential area, Beckford, a police officer, saw Barnes fleeing with what appeared to be a gun. Beckford gave chase and then shot Barnes dead. When he was shot, Barnes was on his knees, unarmed, with his hands in the air, begging to be spared. But Beckford believed (let us accept, honestly) that Barnes was about to shoot him dead, and that he therefore had to shoot Barnes to save himself.[1] Should Beckford be entitled to a justificatory plea of self-defence?

What if we slightly modified the facts? Say, Barnes-2 was indeed armed and, had he not been shot first, would have fired at Beckford-2. Should Beckford-2 be entitled to a justificatory plea of self-defence? (I deliberately leave open the question of whether Beckford-2 was aware of these facts about what Barnes-2 was about to do.)

In both these 'Beckford variations', the defendant commits a *pro tanto* offence (the *actus reus* performed with its *mens rea*), but there are also some key differences. Consider in particular, the issue of whether, objectively and all things considered, the respective victims suffer a wrong, i.e. a violation of their rights. It is plausible to think that because Barnes posed no actual threat to Beckford, objectively speaking, Barnes suffered an all-things-considered wrong. His right to life was violated insofar as, objectively, there existed no all-things-considered-wrong-denying reasons for Beckford to shoot Barnes. Barnes-2, on the other hand, did pose an actual threat to Beckford-2, and so objectively speaking, he was liable to Beckford-2's defensive force. Therefore, all things considered, he was not

[1] *R v. Beckford* [1988] AC 130 (PC). On these facts, the Privy Council ruled that Beckford was entitled to succeed in his plea of self-defence; even though he was mistaken – even unreasonably so – about whether Barnes posed a threat to him, the facts as he perceived them did entitle him to defend himself, and the force he used was not disproportionate on those (perceived) facts.

objectively wronged by being killed by Beckford-2. Objectively, there were all-things-considered-wrong-denying reasons for Beckford to shoot Barnes. If you subscribe to what I call the 'wrongness hypothesis' (Dsouza 2017: 3–6), this difference dictates whether a justificatory defence to criminal liability is available. Here's what the wrongness hypothesis states:

> Justifications exculpate by negating (at least) the objective all-things-considered wrongness of the defendant's *pro tanto* offence whereas excuses leave this undisturbed, and instead exculpate the defendant based on personal blameworthiness-denying factors. Where the defendant's *pro tanto* offence is a victimising one (as in the Beckford variations), the wrongness hypothesis implies that if, objectively, a victim's rights were violated by the *pro tanto* offence, then a plea of justification should never succeed.

In other words, if you subscribe to the wrongness hypothesis, you think that Beckford's plea of justification should not succeed, because Barnes' rights were violated. Beckford-2's identical plea is not disqualified for that reason, though you may perhaps think it is disqualified for other reasons (the one typically cited is that to successfully plead a justificatory defence in respect of a *pro tanto* offence, the defendant needs to have been motivated by justifying reasons). Thus understood, the wrongness hypothesis is one feature of criminal defences that a majority of theorists, including George Fletcher (1975: 320), John Gardner (2007a: 92–5), Andrew Simester (2021: 401–10), Benjamin Sendor (1990: 766), Douglas Husak (1989: 516–7) (1999: 52–3, 55), Miriam Gur-Arye (2003: 21), Albin Eser (1976: 621–3, 635, 637), Paul Robinson (1975: 272–3) (1990: 749–50) (1997: 394–9), Peter Westen (2006: 306), Suzanne Uniacke (1994: 12, 14–22), and Antony Duff (2007: 264–6, 270, 273–6, 281–4), agree upon.

But there is reason to doubt the central place that the wrongness hypothesis occupies in theories of defences.[2] In previous work, I have argued that what I call the 'quality of reasoning' hypothesis is a plausible, well-founded alternative to the wrongness hypothesis and that it gives rise to desirable liability outcomes (Dsouza 2017). But even while raising doubts about wrongness hypothesis-based models of justification, I had stopped short of suggesting that they should be rejected. I take up that task here. Specifically, I argue that we should reject the wrongness hypothesis when theorising criminal justifications. In the context of victimising offences, this means that a defendant, D's, entitlement to a justificatory defence should not depend on whether the putative victim, V, of a *pro tanto* offence (i.e. the *actus reus* performed with *mens rea*) suffered an all-things-considered wrong at D's hands. So, the fact that, objectively, Barnes' rights were violated, should not disqualify Beckford from succeeding in his plea of self-defence. Conversely, it also means that the fact that a *pro tanto* offence happened not to violate the

[2] Greenawalt and Baron separately theorise the availability of justificatory defences based broadly on subjective perceptions of facts, and therefore hold that even someone who commits an objective all-things-considered wrong can be justified. See Greenawalt 1986: 91–99, 102; 1984: 1903, Baron 2005: 393, 396–8; 2009: 124–30. I agree.

apparent victim's rights does not contribute to the success of a justificatory defensive plea. The fact that, objectively speaking, Barnes-2's rights were not violated, does not help Beckford-2 in his plea of self-defence. (That said, as will become clear presently, in no way do I deny the importance of victims' rights to questions of criminalisation, or to exercises of discretion before and during a trial, and to sentencing.) I defend these claims by showing that basing one's theory of defences on the wrongness hypothesis limits the ability of these defences to perform their distinctive role within the criminal justice system, whereas basing it on the quality of reasoning hypothesis facilitates defences in their performance of that distinctive role.

To this end, I begin in Section 1 by identifying the distinctive role that justifications play within the criminal justice system, and demonstrating, in Section 2, that incorporating the wrongness hypothesis into a theory of justification limits the ability of justifications to perform that distinctive role. In Section 3, I survey the arguments made in support of the wrongness hypothesis and argue that they do not convince. Then, in Sections 4 and 5, I illustrate how a theory of justification based on the quality of reasoning hypothesis can allow justifications to essay their distinctive role, while generating plausible liability outcomes. Section 6 concludes.

2. The Distinctive Role of Supervening Defences

Not all defences are alike. Pleas such as infancy, automatism, and insanity deny D's status as a responsible moral agent at the time of the offence; pleas of alibi, inadvertence, consent, and intoxication deny the elements of an offence's *actus reus* or *mens rea*; and pleas such as limitation, diplomatic immunity, and double jeopardy invoke procedural objections to trying D. None of these defences are justificatory. Justificatory defences belong to the set of 'supervening' defences – defences that do not dispute D's responsible moral agency, or that she committed the charged offence's *actus reus* with its *mens rea*, but instead invoke additional factors that 'block the presumptive transition from responsibility to liability' (Duff 2007: 263; Simester 2021: 400; Dsouza 2017: xv). The set of supervening defences includes justifications such as self-defence, the defence of property, and necessity (which can, in some forms, be justificatory), and some excusatory defences like duress. Accordingly, let us focus on identifying the distinctive role of supervening defences in substantive criminal law.

The distinctive (though not necessarily sole) role that supervening defences play in the criminal justice system seems to relate to conduct evaluation (as contrasted with conduct guidance). This is most clearly true of supervening excuses. It is generally accepted that supervening excuses are not conduct-guiding – they do not tell us what we should or may do; instead, they tell the decision maker what to tolerate from us (Thorburn 2008: 1095; Greenawalt 1984: 1899–1900;

Dsouza 2017: 87–92). Merely tolerated behaviour is not encouraged or permitted behaviour, and although we can probably infer from previous grants of excuses what behaviour we are likely to get away with (Duff 2002: 61–8; Lee 2009: 137–8), excuses are not meant to offer conduct guidance as to what we should, or may, do. Their distinctive role is in helping with conduct evaluation.

Justifications are different. They do seem to offer some conduct guidance – they tell us what behaviour is permissible (Stewart 2003: 333–6; Gardner 2007a: 106; Dsouza 2017: 88). So, Robinson treats them as being 'rules of [conduct] guidance' (along with offence stipulations) rather than 'principles of adjudication' (along with doctrines of excuse) (Robinson 1990; Robinson 2013: 237). But that is too quick. Both justifications and offence stipulations do offer conduct guidance, but both also influence adjudication; adjudicators must refer to them in deciding whether their guidance was followed. So, the real issue is whether the *distinctive* roles of justifications and offence stipulations are the same, *va.*, offering conduct guidance. Justifications and offence stipulations differ enough to suggest that that's not obviously the case (see for instance, Tadros 2007: 103–15). Offence stipulations are imperative – 'Do this; don't do that', whereas justifications are permissive – 'You may do this, you may omit to do that'. Additionally, the imperative guidance in offence stipulations narrows the set of conduct choices available to us, while justifications expand them by creating exceptions to guidance contained in offence stipulations (Dsouza 2017: ch 4). And these differences matter: on any plausible theory of justification, claiming a justification is not the lawmaker's 'Plan A' for how things should go; it is a contingency plan – 'Plan B'. Plan A is to prevent *pro tanto* offences from being committed, via the guidance contained in offence stipulations. When that plan fails because a specified contingency has arisen, then Plan B – the justificatory guidance comes into play. All things considered, when planning ahead, the lawmaker would rather that Plan A was followed (i.e. that nobody committed a *pro tanto* offence), than that Plan A was abandoned and Plan B was followed (i.e. a *pro tanto* offence was committed with justification).[3] This may even be the case when the contingency has arisen – even though D is permitted to use fatal force to defend herself against a child innocently aiming a gun at her, there is no reason to assume that the lawmaker *prefers* that D kill the child, rather than vice versa. So, the conduct guidance contained in offence stipulations applies *imperatively* to *all persons* from the time the offence is created, and it continues to apply until and unless the offence is repealed. But when an agent who was following Plan A encounters the specified contingency, the guidance in Plan B *also* becomes applicable, as *optional* guidance, for *that particular agent*, from that moment, and it ceases to apply as soon as the justificatory option is either adopted or precluded. In sum, a justification's guidance has a limited audience – it speaks only to agents

[3] Gardner put it thus: 'Legal justifications are not there to be directly followed by potential offenders. They merely permit one to follow reasons which would otherwise have been pre-emptively defeated.' (Gardner 2007a: 117).

that find themselves in the situation identified as the justification-triggering contingency; and it has only a brief conduct guiding lifespan – starting from when the specified contingency is encountered, and ending when the justified option is exhausted or ruled out. Typically, this lasts just a few moments, though exceptionally, it can be longer. This contrasts sharply with not only an offence stipulation's conduct guiding lifespan (which I have previously discussed), but also with a justification's conduct *evaluating* lifespan, which begins from the time a potentially justified action is performed and can last as long as it takes to exhaust the agent's final appeal against a conviction. At the very least, it lasts until an investigating body decides that the justification so clearly applies that further investigation is unnecessary.

On balance then, although like offence stipulations, justifications assist with both, conduct guidance and conduct evaluation, unlike offence stipulations, the distinctive role of justifications seems to relate to conduct evaluation (Dsouza 2017: 85–8; Duarte d'Almeida 2015: 77; Hart 1948).[4] In this, they are more like supervening excuses; they guide those tasked with evaluating an agent's conduct after it has been performed.

The particular conduct evaluation with which supervening defences assist is performed at trial.[5] The aims of a criminal trial shape the sort of conduct evaluation that is performed during the trial, and not all matters that could possibly be relevant to conduct evaluation are in fact relevant to conduct evaluation during a trial. There are several things a criminal trial aims to do, but probably the most important in the present context, is reaching a just verdict of 'guilty' or 'not guilty'. The central judgement in these verdicts is what Michael Zimmerman (2002: 554) calls 'hypological' – a judgement *about the laudability, culpability, or moral neutrality of the defendant* in respect of her conduct. Hypological judgements are a type of agent evaluation that can be contrasted with 'deontic' judgements, which are often thought to be a type of act evaluation involving judgements about moral rightness or wrongness of the agent's act (Zimmerman 2002: 554).[6] They can also be contrasted with 'aretaic' judgements, which are judgements about the agent's character – her moral virtue and vice (Zimmerman 2002: 554). Whereas a deontic judgement may be concerned with whether the agent's *act* was morally good, and an aretaic judgement may be concerned with whether the agent has the *character* trait of bravery, a hypological judgement is concerned with whether an *agent acted* praiseworthily or blameworthily in respect of *this* particular conduct token.

[4] Hart later retracted his 1948 essay, but Duarte d'Almeida has sought to reinvigorate Hart's claims from that essay in his monograph.

[5] Even when investigators and prosecutors consider the availability of a defence prior to trial, they do so in order to predict the defendant's chances of success should the defence be raised at trial.

[6] Although Zimmerman expresses some doubt about the precision of this characterisation of deontic judgements, the proposition that they are a type of act evaluation captures their essence adequately and distinguishes them from those that are hypological.

The central judgement in a verdict is this last sort of judgement – one that is hypological.[7] It follows therefore, that the distinctive role of supervening defences is to assist with the making of hypological judgements at trial. Any consideration that interferes with the ability of supervening defences to perform this distinctive role ought to be excluded.

3. The Distorting Effects of the Wrongness Hypothesis

In what follows, I argue that one such consideration is the content of deontic judgements about the objective all-things-considered wrongness (or rightness, or neutrality) of D's act, which, proponents of the wrongness hypothesis, insist shapes justificatory defences. But first, let me clarify that I am not suggesting that the substantive criminal law should be *uninterested* in whether we judge what happened as to be welcomed, lamented, or greeted with indifference. It seems plausible to think that such deontic judgements can matter at stages other than when applying supervening defences. For instance, it seems likely that they matter when deciding what sorts of conduct to make criminal in the first place. Perhaps they matter pre-trial, in the exercise of any discretion an investigator or prosecutor has as to which matters to investigate, or prosecute, respectively. They may also matter during the trial, but prior to applying the supervening defences, when establishing whether D is *eligible* for a criminal law hypological judgement. Apart from showing that D performed her conduct as a responsible agent, this may entail showing that her conduct was deontically objectionable (by showing that a *pro tanto* offence was actually committed). Additionally, deontic judgements may well matter for post-conviction sentencing decisions. At all these stages, it may well be appropriate to be guided by concerns about the victims' rights. At least some scholars would dispute the relevance of victims' rights in at least some of these domains (see e.g. Alexander and Ferzan 2009), but here, I need not take a firm stance on those debates. My immediate focus will be to show that they should be excluded when applying supervening defences.

Note first that our hypological judgement about an agent (let's call this judgement$_a$), and our deontic judgement about the act (let's call this judgement$_e$) can diverge. One clear example is when an agent performs the *actus reus* of an offence without its *mens rea*. Consider someone who has non-consensual sex with another,

[7] No doubt, a victim (assuming there is proved to have been one), will have an interest in knowing the verdict, and especially if the verdict is 'not guilty', she will have an interest in knowing *why*. (Interestingly, some common law systems – specifically those in which juries try criminal cases and are not required, or indeed permitted, to give reasons for their verdicts – will be unable to cater to this latter interest in knowing why a verdict was returned.) But this does not mean that the judgement contained in the verdict itself is a judgement about whether she was recognised as having suffered a wrong, i.e. a deontic judgement about the rightness or wrongness of the defendant's act. The verdict remains a hypological judgement. As such, it need not be determined by a deontic judgement about the rightness or wrongness of the defendant's act. My thanks to the editors for pushing me to clarify this point.

while reasonably believing that the latter was consenting. Here, judgement$_e$ is adverse but judgement$_a$ is not. Gardner explains that while we are tempted to think that we can avoid authoring any wrongs if only we are careful and clever enough, this temptation stems from the overly optimistic assumption that the perfection of our rational faculties will allow us to perfect our lives (by, in this context, never authoring a wrong) (Gardner 2007a: 81). But we are only human; we have limited epistemic access to objective truths, and limited control over the uncertainties of the physical world. So, despite our best efforts, we will inevitably author tragedies and fortuitous events. But the example above shows that even when objectively terrible things happen, an agent may not commit a *pro tanto* criminal offence, and so may not be liable to the adverse hypological judgement contained in a conviction.

Judgement$_a$ and judgement$_e$ can also diverge when a *pro tanto* offence has been committed. Here's how: The criminal law works in stages. *Ex ante*, it identifies conduct that should be avoided so as to prevent some specific *pro tanto* wrong from occurring. Then, *ex post*, it evaluates cases in which the proscribed conduct was performed anyway, to determine whether the agent was blameworthy for having performed the conduct (Simester and von Hirsch 2011: 3; Dsouza 2017: 8). The major part of the *ex-ante* conduct guidance is found in offence stipulations, which are typically, a combination of *actus reus* and *mens rea* elements. But some *ex-ante* conduct guidance is contained in justifications, rather than offence stipulations. So, an agent may commit a pro tanto offence, but still have chosen her actions entirely in line with the criminal law's advance conduct guidance, if she acted in a manner that was open to her under some justificatory criminal law rule. In this sort of case, the agent has, after all, behaved as the criminal law told her she was required or permitted to behave, and so judgement$_a$ is not adverse (Dsouza 2015). But again, because the agent has limited epistemic access to objective truths, she may nevertheless have authored a wrong such that judgement$_e$ is adverse. *Beckford*[8] was a case in which this happened. Here's another example:

> D reasonably believes that V is threatening her life, and so, she chooses to defend herself by pushing V. This is the minimum amount of force that would be effective against the threat thought to be posed by V. Choosing in this way is permitted by the justificatory rule governing self-defence. But it transpires that despite appearances, she was not being attacked, and therefore her 'defensive' action victimises an innocent V.

Since D has chosen her actions entirely in line with the criminal law's advance guidance, we have no reason to make an adverse judgement$_a$ about her. Yet, since V has been wronged, our deontic judgement$_e$ about what happened is adverse.

In short, persons who perform deontically problematic conduct need not have behaved in a manner that calls for an adverse hypological judgement. To be clear, adverse deontic judgements are strongly correlated with adverse hypological judgements insofar as when we set our minds to performing a task (including one that merits a deontically adverse judgement), we usually succeed. But even

[8] *Beckford* (n 1).

the most skilled, careful, and clever of us will sometimes fail, because we do not control all of life's variables. In those cases, judgement$_a$ and judgement$_e$ will pull in opposite directions.

Now recall that all wrongness hypothesis-based theories of defences reserve the label 'justification' for instances in which judgement$_e$ and judgement$_a$ are both at least neutral and use the label 'excuse' for instances in which judgement$_e$ is adverse, while judgement$_a$ is neutral or positive. But if the distinctive role of supervening defences is to assist with making hypological judgements, then it is inappropriate to deny a justification when judgement$_a$ is neutral or positive. To do so conveys false information about our hypological judgement of the agent, and we do not want the criminal law to speak untruths (Simester 2021: 10–1). What's more, wrongness hypothesis-based theories of defences also use the label 'excused' for agents who offend under duress. In duress cases, judgement$_e$ is clearly adverse; where D commits a victimising *pro tanto* offence, an innocent V has suffered a wrong. Additionally, insofar as D (knowingly) does not act as the criminal law permits,[9] D deserves an adverse judgement$_a$ as well. This means that some agents who respond perfectly to the criminal law's guidance (i.e. those who deserve a positive or neutral judgement$_a$, although judgement$_e$ is adverse) are clubbed with others who do not (i.e. those who successfully plead duress). Again, wrongness hypothesis-based theories deploy the terminology available to convey misleading, or at least, insufficiently nuanced, information about defendants. In fact, any theory of defences that tries to use the labels 'justification' and 'excuse' to accurately state both hypological and deontic judgements asks too much of these labels.

Duff's solution is simply to expand our vocabulary (2007: 264–6, 270, 273–6, 281–4). He says that people for whom judgement$_e$ is adverse and judgement$_a$ is not, should be said to have acted with 'warrant'. I have no especial theoretical objection to this general approach, but I doubt it is a good idea because:

(a) labels like 'warranted' do not carry the same social significance as the labels, 'justification' and 'excuse', so they may be less able to convey our hypological approval of an agent who responded perfectly to the criminal law's guidance;
(b) introducing new terminology increases the complexity of our system of defences solely to convey information about whether an all-things-considered wrong occurred. In victimising offences, this amounts to information about whether the victim, in some sense, deserved what happened to them. We should be uncomfortable with embedding this sort of victim-evaluation (and potentially, victim-blaming) into our criminal justice system. Besides, information about whether something objectively wrong happened is gratuitous; the verdict is meant to state a hypological judgement about the defendant. The added complexity associated with creating new labels is not justified by any added value.

[9] As distinct from tolerated, or 'not censured or punished'. As I use these terms, one may tolerate impermissible action, but if action is permissible, there is no need for anyone to display toleration.

In sum, the wrongness hypothesis, with its insistence on considering the deontic merits of D's conduct, hinders supervening defences in the performance of their distinctive role by distorting the labels 'justification' and 'excuse'. Rejecting it (and connectedly, the idea that justifications should depend on deontic judgements about D's conduct), may enable us to deploy these labels, with their existing social significance, to convey accurate hypological judgements about D.

4. Arguments for the Wrongness Hypothesis

So why do so many people, theorists and lay, have such strong intuitions in favour of the wrongness hypothesis? One reason might be that in ordinary speech, it seems natural to want to communicate whether a wrong happened even if it was occasioned non-culpably. A person suffering a wrong may also want that fact to be recognised in an authoritative judgement issued by a court considering the facts. It is possible to communicate this by reserving the language of justification for instances where the agent acted non-culpably, *and* no wrong happened. Doubtless, the ease with which we can slip between two senses of a statement like, 'D pushed V with justification' – the first, a judgement about whether the agent was justified, and the second, a judgement about whether the pushing was justified – also helps. But *pace* Gardner (2007a: 95), I doubt that the rationale for using the term 'justification' like this survives the transition from ordinary speech into speech made in the context of a criminal verdict, for reasons that are hopefully clear by now. While in ordinary speech, we can afford to be promiscuous about the sorts of judgements we try to convey by our usage of terms like 'justification', in the context of a criminal verdict, our focus is unambiguously on making hypological judgements – judgements about the merit *of the defendant* in light of her conduct. When we use terms like 'justification' and 'supervening excuse' in this context, we should focus only on judgements about the agent.

Nor does the sheer ubiquity of the supposed intuition favouring the wrongness hypothesis offer any support for its truth. Constant repetition of a plausible proposition (which the wrongness hypothesis undoubtedly is) can convert the proposition into received 'wisdom', and then, into gospel. As Ralph Waldo Emerson (2015: 136) observes, when that happens, the proposition can masquerade as intuition when actually, it is an instance of a 'tuition' – a learned belief that circumvents scrutiny. In other words, the familiarity of the wrongness hypothesis may not evidence its correctness so much as reflect its pervasiveness. So, discard intuition. Consider instead the occasional attempts to argue for the wrongness hypothesis from logic.

For Robinson (1975: 271–2; 1990: 740–2, 749–52; 1997: 399–400), justifications are simply negative offence elements that cannot conveniently be included in the offence stipulation. Since we criminalise *actual* wrongful harms rather than merely intended or risked ones, justifications must work by negating the objective

wrongness (or harmfulness) of what happened. It will be clear from my arguments in Section 1, that I reject the claim that justifications can, in principle, be collapsed into the same category as offence stipulations. If my arguments for doing so convince, then the premise of Robinson's argument falls, and with it, his support for the wrongness hypothesis.[10]

A different strategy is to start with some intuited claims about how justifications and excuses should function and argue back from those features to the wrongness hypothesis. Those arguing in this way often identify one or more of the following intuited claims about justifications and excuses:

(a) justifications cannot conflict with each other;
(b) third parties should be permitted to assist justified, but not excused, agents;
(c) nobody should be permitted to resist justified action (Robinson 1996: 51–54, 59–60) (Robinson 1997: 404–6) (Fletcher 1978: 759–69).

Additionally, some people using this argumentative strategy also make assumptions about which defensive claims can properly be called justifications. For instance, Robinson (1996: 45; 1997: 391) assumes that both, police officers making arrests, and persons acting in self-defence, make justificatory defensive pleas. That assumption is at least controversial – Gur-Arye (2011; 2021) and I (Dsouza 2024) have argued that police officers making ordinary arrests exercise a power or act within the realm of a rights-displacement. They do not commit a *pro tanto* offence. Therefore, they need offer no supervening defence to a criminal court. But my concern here is about the broader tactic of arguing backwards from expected features of the justification/excuse divide to its structure.

However deployed, this tactic is weak. There is considerable disagreement on the features that justifications and excuses should have, so this sort of argument cannot convince anyone with different intuitions. Besides, this style of argument arguably proceeds in the wrong direction – our theory of justification and excuse should decide further questions about third-party consequences, conflicts of justification, and the right to resist; not the other way around (Husak 1999; Tadros 2007: 119–20, 281–2; Greenawalt 1984: 1919–27).

Another approach, which Fletcher (1974: 1272, 1274-6, 1304; 1985: 958) introduced into Anglo-American thinking from German criminal law theory (Eser 1976: 626–30) appeals to the pleasing symmetry that the wrongness hypothesis sets up between the *actus reus*/*mens rea* distinction and the justification/excuse distinction. On this view, just like the *actus reus* refers to the objective elements of a *pro tanto* offence and the *mens rea* to its subjective elements, justifications negate the objective wrongness of the *pro tanto* offence and excuses negate the subjective

[10] Incidentally, Alexander and Ferzan (2009) also think that there is no substantive difference between justifications and offence stipulations but instead argue that the criminal law should actually focus on culpability rather than wrongs and harms. They therefore argue that neither offence stipulations nor justifications, should focus on the fact-relative wrongness of what happened. So, even if we agree with Robinson's initial premise, his conclusion does not necessarily follow.

wrongness (or culpability) of what D did. Now, undoubtedly this elegant symmetry does lend some appeal to the wrongness hypothesis and, additionally, it offers a straightforward way to distinguish between justifications (which must, at a minimum, deny that D authored an all-things-considered wrong) and excuses (which do no such thing). However, the logical leap in assuming that thinking of offences and defences symmetrically will help us work out how criminal laws ought to function is unsubstantiated. There is no particularly good reason to think that the claimed *actus reus*-justification and *mens rea*-excuse symmetry can offer any normative support for a theory of how the criminal law should function. True, offence and defence definitions (and within them, *actus reus*, *mens rea*, justification and excuse stipulations) perform different functions within the criminal law, but they are not *how* the criminal law functions – or at least, they do not set up the only account of how the criminal law functions. I described one plausible alternative in Section 1. On this account, the criminal law functions by *ex-ante* stipulating conduct that should be avoided, and *ex post* evaluating cases in which the conduct nevertheless occurs, to determine the actor's blameworthiness. The distinctive role of offence stipulations (i.e. the *actus reus* and *mens rea*) relates to the *ex-ante* stage of criminal law, and so the *ex-ante* function of the criminal law, viz. providing prior conduct guidance with a view to avoiding identified harms, shapes their features. Conversely, since the distinctive role of supervening defences relates to evaluating conduct after it contravenes an offence stipulation, the *ex-post* function of the criminal law, viz. evaluating D's blameworthiness, shapes the features of supervening defences. Accordingly, all supervening defences should work by negating (at least some aspects of) the defendant's blameworthiness, and none of them should work by negating the occurrence of a wrong (Dsouza 2017). This would mean we need a new way of distinguishing between justifications and excuses, but plausible options exist (e.g. Dsouza 2017: 109–20). In sum, the argument from symmetry too does not give us compelling reasons to accept the wrongness hypothesis.

5. Outlining an Alternative

Rejecting the wrongness hypothesis simplifies our task; if information about judgement$_e$ is superfluous to a criminal verdict, we need not convey it. We can reserve the label 'justification' for all and only cases in which our hypological judgement (i.e. judgement$_a$) is positive or neutral, such that justifications turn on the defendant's culpability/blameworthiness, and not (at all) on the wrongness of the defendant's deed. The flip side is that now we need a new way to distinguish between justifications and supervening excuses. Elsewhere (Dsouza 2017), I have defended one model, based on what I call the 'quality of reasoning' hypothesis, that rejects the wrongness hypothesis, but can still distinguish between justifications and excuses. I use it here to demonstrate that viable alternatives to wrongness

hypothesis-based theories of supervening defences do exist. In outline, my quality-of-reasoning-based model of defences:

- is limited to supervening defences. Nobody suggests that the wrongness hypothesis explains the exculpatory pull of other defensive claims, including those that deny the court's jurisdiction, D's responsible agency, the *actus reus*, or the *mens rea*, (Dsouza 2017: xv) and so they can be defended on their own terms;
- distinguishes between the quality of D's 'functional-reasoning', i.e. how well she meets expected standards of acuity in perceiving facts and drawing appropriate conclusions from them, and her 'norm-reasoning', i.e. whether she chooses her conduct in a manner that aligns with the normative guidance applicable to her volitional conduct, including normative guidance as to how much care she should take when exercising her capacities to perceive facts and draw appropriate conclusions from them, where such guidance is provided (Dsouza 2017: 24–8). In general, the criminal law's hypological judgements should track the quality of D's norm-reasoning; poor norm-reasoning should attract 'norm-blame'. The quality of D's functional-reasoning (and correspondingly, her desert of 'functional-blame') is relevant primarily outside the criminal law, where a finding of liability does not automatically suggest any morally adverse judgement about D;
- distinguishes between norms contained in the criminal law's system of conduct-guiding rules (criminal conduct norms) and broader societal conduct-guidance (societal conduct norms) (Dsouza 2017: 110–1). These often overlap substantially, but occasionally societal conduct-norms require something that criminal conduct norms do not (e.g. politeness), and vice versa;
- accepts that since some criminal conduct norms are contained in justifications rather than offence stipulations, D may deliberately violate an offence stipulation but still have chosen her conduct in line with the criminal law's (justificatory) conduct norms. In such cases, D is entitled to a justification. Having conducted herself in line with the criminal law's advance conduct guidance, she has displayed no culpable norm-reasoning in respect of criminal conduct norms. In short, justifications negate any blameworthiness flowing from poor norm-reasoning in relation to criminal conduct norms (Dsouza 2017: 96–9). But (assuming perfect knowledge of the law's conduct guidance[11]) when D's choice of conduct in violating an offence stipulation reveals an inappropriate attitude towards the criminal law's (overall) conduct guidance, she is not entitled to a justification. Having deliberately chosen to act contrary to criminal conduct norms, D deserves, pro tanto, to be called a criminal;

[11] This is so unlikely an assumption that it is better to think of it as a deeming fiction. See Dsouza 2017: 26. Insofar as it simplifies the analysis, it will serve for the present purposes. The (de)merits of this fiction, and what should replace it, are questions for another piece.

- allows an unjustified D to nevertheless have a claim to a supervening excuse, also based on a denial of blameworthiness; albeit blameworthiness flowing from poor norm-reasoning in relation to societal conduct norms. This may happen when, say, she deliberately contravenes the criminal conduct norms because she is wrongfully subjected to a serious threat. On such occasions, D's conduct, notwithstanding its deviance from criminal conduct norms, may be in line with societal conduct norms governing how one may respond in light of the threat. She will not therefore, by her conduct, have shown herself to have norm-reasoning that compares poorly to the normative gold standard for societal norm-reasoning. It would therefore be hypocritical for a criminal court, as representative of that society, to single D out as being especially deserving of a criminal conviction. This objection to hypocritical blaming is an objection to the blamer, rather than the content of the blame – a blaming judgement may be both hypocritical, and true. Therefore, the hypocrisy-based objection to a conviction is not conclusive – hypocritical, but true, blaming decisions may well be justifiable despite their hypocrisy, by reference, for instance, to overriding policy considerations. These overriding considerations may also permit the imposition of additional conditions (which may well be motivated by a consideration of victims' rights, and may, for instance, require that the agent's functional reasoning have met a specified standard of quality) for the grant of an excuse. In sum, when D claims a supervening excuse, she has a good blameworthiness-denying case for (but no entitlement to) the defence (Dsouza 2017: 109–20); and
- treats necessity as a non-paradigmatic form of justification, with its own exculpatory logic. Briefly, on this view, 'best interests intervention' necessity[12] is not a supervening defence at all; it negates the (often implicit) non-consent element of an offence's *actus reus* by legally recognising a substituted consent (Dsouza 2017: 123). Lesser-evils justificatory necessity[13] is a supervening justification, that functions like an excuse, except that, because the evil averted was so much greater than the evil caused, the law deems D to have acted with justification.

The differences between the outcomes generated by this model of defences and wrongness hypothesis-based models are set out below.

[12] This is the form of necessity in which an agent (D) appears to victimise another (V), but does so in the best interests of V, in circumstances in which it was impracticable to ascertain whether V consented to D's intervention, because, for instance, V was temporarily or permanently incapacitated, or there was no time to ask. See Winnie Chan and Simester 2005: 124–7; Stark 2013: 950, 952–56; Dsouza 2017: 122–3.

[13] This form of necessity is premised on an act-utilitarian comparison of the evil brought about by an agent's *pro tanto* criminal intervention, as compared to the evil that would have occurred in a counterfactual world in which the agent did not intervene. See the judgment of Brooke LJ in *Re A* [2001] Fam 147; Stark 2013: 957–9; Dsouza 2017: 123.

Table 7.1 Supervening Defence Outcomes: Wrongness Hypothesis-Based Models v The Quality of Reasoning Model

#	All-things-considered wrong? (A)	Followed criminal conduct guidance believed applicable?[14] (B)	Mistaken about what criminal conduct guidance was applicable?[15] (C)	If mistaken re (C), reasonable mistake? (D)	Followed societal conduct guidance believed applicable?[16] (E)	Mistaken about what societal conduct guidance was applicable?[17] (F)	If mistaken re (F), reasonable mistake? (G)	Outcome: Wrongness hypothesis (H)[18]	Outcome: Quality of reasoning (I)
1	N	Y	N					Justified	Justified
2	N	Y	Y	Y				Justified/ Excused	Justified
3	N	Y	Y	N				Excused/No Defence	Justified
4	N	N			Y	N		Excused	Can be Excused

[14] If the answer here is 'N', then no justification is available even on (most) wrongness hypothesis-based models.

[15] I refer only to cases where D is mistaken about the facts, so believes some inapplicable legal guidance applies. I exclude mistakes as to the content of the law.

[16] I refer here to potentially exculpating reasons, whether or not they are accepted in a legal system. Where the answer is 'N', no defence is available even on (most) wrongness hypothesis-based models.

[17] I refer only to cases where D is mistaken about the facts, and so believes some inapplicable societal norms apply. I exclude mistakes as to the content of societal norms.

[18] Alternative entries in this column reflect differences in the details of the main wrongness hypothesis-based models. These details are not material to my argument here.

5	N	N			Y	Y	Y	Excused	Can be Excused
6	N	N			Y	Y	N	No Defence	Can be No Defence[19]
7	N	N			N			No Defence	No Defence
8	Y	Y	N[20]					Excused	Justified
9	Y	Y	Y	Y				Excused	Justified
10	Y	Y	Y	N				Excused/No Defence	Justified
11	Y	N			Y	N		Excused	Can be Excused
12	Y	N			Y	Y	Y	Excused	Can be Excused
13	Y	N			Y	Y	N	No Defence	Can be No Defence
14	Y	N			N			No Defence	No Defence

[19] I use different words here (and in Row 13) than the words in Rows 4, 5 etc, because although nothing in my view commits us to treating these cases differently from others in which there is a *pro tanto* reason to excuse, I recognise that matters external to my view (i.e. policy considerations) may favour the withdrawal of a defence here.

[20] Such wrongs can occur because we humans are not in perfect control of all variables that determine the outcomes of our actions – e.g. *R v Scarlett* [1993] 98 Cr App 290.

6. Comparing the Models of Defences

So, how does the quality of reasoning model's way of using the terms 'justification' and 'excuse' compare with that of wrongness hypothesis-based models? Consider the outcome columns in Table. 7.1. In a full model of defences, the outcome column(s) can contain three possible definitive values – Justified, Excused, or No Defence.[21] With just those options, there is only limited information one can convey.

On the quality of reasoning model:

(1) A 'Y' (for 'yes') in column (B) means that D displays the proper attitude towards the criminal law's guidance, and so D acted with justification. Of course, since norm blame depends on D's attitude towards the norms, I refer here to the criminal law norms applicable in the facts as D perceives them. Being justified in the criminal law does not preclude non-criminal liability, for instance, for poor functional reasoning.

(2) A 'Y' in column (E) indicates that D has a defeasible case for an excuse based on having responded appropriately to the societal (though not criminal) conduct norms applicable in the facts that she perceived. Since her conduct shows her normative reasoning to be no worse than that of the normatively acceptable member of society, it would be hypocritical for the criminal justice system representing that society to single out D as meriting the stigmatising label, 'criminal' (Dsouza 2017: ch 6). However, this argument may be defeated by, for instance, overriding policy considerations. Moreover, a criminal law excuse does not preclude non-criminal liability, for instance, for poor functional reasoning.

(3) We always convey accurate information about criminal norm-blameworthiness when granting D a justification. The grant of an excuse is less definitive, though even there, since one would rather be justified than excused (Gardner 2007a: 108–13, 133; Baron 2005: 389; and for a survey of the literature on this claim, see Husak 2005a: 557), the model virtually guarantees that an excused D's conduct did not conform to criminal conduct norms, although it conformed to societal conduct norms. Despite conforming to societal conduct norms, some people may not be excused from criminal liability, since there is no entitlement to being excused – while one *asserts* a justification, one only *asks* to be excused (and so, can be refused). Of course, the state cannot refuse an excusatory defence groundlessly – there must be reasons. But these reasons flow not from the structure of criminal law defences, but

[21] In Table 7.1, there are also two non-definitive values in these columns, viz. 'Can be Excused' and 'Can be No Defence'. I have used these non-definitive values to recognise that the outcomes in these cases are under-determined by our theory of the structure of defences. Each criminal law system will need to take a definitive stance on these cases, and its stance will be dictated by the matters that fall under the broad heading of public policy. But whatever the dictates of public policy in any given system, the available definitive stances remain these three identified in the main text.

rather, from broader policy considerations, like the seriousness of the offence, the standards of epistemic reasoning we expect people to display, and the need for judicial consistency.

(4) whether an all-things-considered wrong occurred is extraneous to deciding what type of supervening defence D can plead. This considerably simplifies the analysis, allowing us to drop column (A), and retain what remains of Rows 1 to 7 to address all relevant variations of cases.

On standard wrongness hypothesis-based models:

(1) The labels 'justification' and 'excuse' attempt to convey information about both, judgement$_e$ and judgement$_a$. The most prominent such models allow a justification when there was no all-things-considered wrong (i.e. judgement$_e$ is at least neutral), no criminal norm-blame, and no functional-blame. They grant an excuse when a justification is unavailable, but some other factor negates either criminal norm-blame, or societal norm-blame (coupled, on some accounts, with an absence of functional-blame).

(2) We need twice the number of rows, and one extra column, in our table to describe the outcomes in various factual permutations because of the added complexity.

(3) The message conveyed by the verdict is distorted by obscuring information about the hypological judgements we can make about the agent in respect of her conduct. The labels 'justification' and 'excuse' cannot reliably convey information about either D's criminal norm-reasoning, or her societal norm-reasoning. Instead, some agents who respond appropriately to criminal norms (Rows 2 and/or 3) are clubbed with agents who do not (Rows 4, 5, 8, 9, 11, 12 and possibly also 10). While this problem is starkest in relation to agents who complied with norms applicable to the facts as they reasonably, but mistakenly, believed them to be, the Table shows that it also affects several *other cases*.

One apparent weakness of the quality of reasons model vis-à-vis wrongness hypothesis-based ones is that is seems insensitive to the fault associated with being unreasonably mistaken as to facts; it does not distinguish between Rows 2 and 3 (and correspondingly, Rows 9 and 10) and Rows 5 and 6 (and correspondingly, Rows 12 and 13). But on closer examination, this objection rarely bites. Consider first, Rows 2 and 3 (and correspondingly Rows, 9 and 10). The objection does not apply whenever the law has a criminal conduct norm guiding agents, in advance, to take care that their functional reasoning conforms to a specified standard (usually, reasonableness) when they choose to commit a *pro tanto* offence and then plead justification. Such a norm is not built into the structure of the criminal law, but lawmakers can choose to impose it (Dsouza 2017: 40–45).[22] Where they do, and

[22] I take no stand on whether, and in what circumstances, lawmakers should require agents to take care to ensure that their functional reasoning conforms to a specified standard. My thesis here is not

D does not choose to take the requisite care, she displays poor criminal conduct norm-reasoning. Her case is a Row 4 (or correspondingly, Row 11) case, rather than a Row 2 or 3 case. Now consider Rows 5 and 6 (and correspondingly, Rows 12 and 13). A similar argument can be made here. The objection does not apply whenever social morality contains advance guidance requiring agents to take care that their functional reasoning conforms to a specified standard when they choose to commit a *pro tanto* offence and then plead a defence. Again, while this guidance is not a necessary feature of societal morality, a given society's morality may include it. Where that is the case and D does not choose to take the requisite care, she displays poor societal conduct norm-reasoning, and so there is no reason to excuse her. Her case is a Row 7 (or correspondingly, Row 14) case instead of a Row 5 or 6 case.

The objection still bites in respect of the few cases in which

(a) there was no advance guidance in criminal or societal conduct norms requiring the agent to take special care in exercising her capacities for functional reasoning when she chooses to commit a pro tanto offence and then plead a supervening defence, and the agent reaches conclusions that we, *post facto*, judge to have been unreasonable; or
(b) the agent tried to obey the relevant advance conduct guidance but fell short where a reasonable person would not have, because she has certain personal traits that the law refuses to attribute to the 'reasonable person'. I refer here to cases in which D is less able to meet the standard of the reasonable person because she is on the autism spectrum or has a lower-than-normal IQ (but is ineligible for an irresponsibility defence). D may even be undiagnosed at the time of her alleged offending, and therefore unaware that she ought to calibrate her efforts to account for her divergent attributes.[23]

I have argued elsewhere (Dsouza 2015) that such agents do not exhibit shortcomings of a nature that merit criminal consequences (though perhaps non-criminal law responses may be appropriate). For that reason, I consider it a strength of the quality of reasoning-based model that, unlike wrongness hypothesis-based models, it does not impose criminal law consequences in these cases.

7. Conclusion

The wrongness hypothesis commits us to the view that justifications are unavailable whenever an all-things-considered objective wrong has occurred (or in the

about what norms lawmakers ought to include in the criminal law system, but rather, about the proper way to internally organise those norms that they do include.

[23] See *R v B (MA)* [2013] EWCA Crim 3, in which the defendant argued that the facts were precisely of this sort. The court disbelieved the defendant but ruled, obiter dicta, that even if it had believed him, it would have been appropriate to convict him.

context of a victimising offence, someone's rights have been violated. But this consideration obstructs supervening defences (including justifications) in their performance of their distinctive role, viz., contributing to making fair hypological judgements about defendants who have committed pro tanto offences. We should therefore abandon the wrongness hypothesis, and accordingly, treat any consideration of whether, objectively, an all-things-considered wrong occurred (or whether a victim's rights were violated), as irrelevant to whether the defendant acted with justification. Doing so considerably simplifies our analysis of which defence is available and in what circumstances, while helping us to convey more accurate hypological verdicts about the defendant.