The Rise and Fall of Duress: How Duress Changed Necessity Before Being Excluded by Self-Defence

Steve Coughlan*

Taken separately, the Supreme Court of Canada’s clarification of the defence of duress in R v Ryan and Parliament’s new statutory version of self-defence appear to have brought more clarity to the law of defences. However, taken together, they may have three unintended consequences. First, the Court’s introduction of moral involuntariness as a principle of fundamental justice in Ryan requires that the defence of necessity be revised to bring it into line with section 7 of the Canadian Charter of Rights and Freedoms. Second, the new self-defence provision expands the scope of that defence to cover situations now covered by the defence of duress. Third, this expansion of self-defence to cover some situations previously covered by duress serves to narrow common law duress to the point of extinction, even though statutory duress remains. In the result, the bifurcation of duress ended by the Supreme Court in Ryan has been revived by Parliament: principals must rely on statutory duress, while parties must rely on self-defence. The author highlights the intricate relationships between duress, self-defence and necessity, calls for a re-evaluation of the constitutionality of the defence of necessity, and draws attention to the potential implications of the new self-defence provision for the law of defences.

* Schulich School of Law, Dalhousie University.
Introduction

I. The “Essentially Similar” Defences of Necessity, Self-Defence and Duress

A. The Trigger and the Response
B. The Filter

II. The Evolution of Duress and the Implications of Ryan for Necessity

A. The Evolution of Duress
   (i) The Origin of the Problem: Carker
   (ii) The Second Maneuvre: Ruzic
   (iii) The Current Approach: Ryan
   (iv) Moral Involuntariness Pre-Ryan to Ryan
   (v) The Mental Element of Duress
   (vi) The Common Law Defence of Duress
   (vii) The Differences Between Common Law Duress and Statutory Duress

B. The Implications of Ryan for Necessity
   (i) The Existing Test
   (ii) The Fourth Element of Necessity: The Accused Did Not Create the Urgent Situation
   (iii) The Fifth Element of Necessity: A Reasonable Belief that the Situation Exists
   (iv) The Three Dissimilarities Between Duress and Necessity
      a. The First Difference: An Urgent Situation of Imminent Peril
      b. The Second Difference: The Mental Element of Proportionality
      c. The Third Difference: Creating One’s Own Compulsion
   (v) The Final Test Post-Ryan

III: Duress and Self-Defence

A. The Trigger, Filter and Response in the New Section 34
B. A Note on Assumptions
C. Should Self-Defence and Duress Be Kept Distinct?
D. If Ryan Was Considered Today

Introduction

The first three months of 2013 saw two significant events for the law of criminal defences: the Supreme Court of Canada’s decision in R v Ryan1 clarifying the statutory and common law defences of duress, and the proclamation of Bill C–26 (the Citizen’s Arrest and Self-defence Act2) bringing in a new statutory version of self-defence. In this paper, I want to discuss the significance of those specific occurrences in their own right and consider the impact they have on defences more generally.

In particular, I suggest two separate conclusions should be drawn. First, one implication of the decision in Ryan is that the defence of necessity

2. SC 2012, c 9 [Citizen’s Arrest].
must be re-evaluated. The Supreme Court has long held that duress and necessity are essentially similar: a more sophisticated understanding of duress therefore casts light on how to better understand necessity. Second, I argue that although Ryan satisfactorily resolved the difficulties that had plagued duress for forty-six years, the solution came just in time to be undone by the passage of Bill C–26.

My argument will proceed in three parts. In Part I, I will consider the Supreme Court’s observation that self-defence, necessity and duress are “essentially similar”.3 I will point out broadly why this is true and what it means for the relationship between those three defences. In dealing with self-defence, I will consider the statutory provisions which have just been repealed.

In Part II, I will look in detail at the changes that have been made to the defence of duress over the past forty-six years, and most importantly, in the very sensible decision in Ryan. I will do so in part because understanding duress in its own right is necessary. However, the discussion will also be significant for the defence of necessity. I will argue that the developments in the law of duress, as explained in Ryan, demand similar changes to the defence of necessity.

Finally, in Part III, I will turn to the new statutory provision on self-defence and discuss how it changes the relationship between these three defences. I will argue that the new provision expands self-defence to render common law duress largely irrelevant. As a result, the central distinction which was sensibly eliminated in Ryan has been, only weeks later, re-introduced by Parliament.

I. The “Essentially Similar” Defences of Necessity, Self-Defence and Duress

Necessity, self-defence and duress all have the same overall structure: because of some trigger, the accused responds by committing an offence. As the Supreme Court has observed, “self-defence, necessity and duress all arise under circumstances where a person is subjected to an external threat” at para 17).

3. R v Ryan, supra note 1 (“[a]ll three apply in ‘essentially similar’ situations: each is concerned with providing a defence to what would otherwise be criminal conduct because the accused acted in response to an external threat” at para 17).
danger, and commits an act that would otherwise be criminal as a way of avoiding the harm the danger presents”.4

In the case of necessity, the Court has held that if an accused faced urgent circumstances of imminent peril, had no reasonable legal alternative, and the harm caused by the accused was proportional to the harm avoided, then the accused is entitled to a defence for committing an offence.5

The recently repealed version of self-defence is difficult to encapsulate simply—which is one of the reasons it was repealed.6 As an example, however, one might consider section 34(2) of the Criminal Code, the provision most often applied in cases involving deadly force in self-defence.7 It provided that if a person was unlawfully assaulted, reasonably feared death or grievous bodily harm from that assault, and reasonably believed that there was no other method of self-preservation, then that person could cause death or grievous bodily harm in repelling the assault.

Finally, consider duress. It exists in statutory and common law form, and one of the points to be pursued in this paper is how the Court has recently harmonized the two versions in Ryan. For the moment it is enough to note, in essence, that if a person is ordered to commit an offence under threat of serious harm, has no reasonable way to avoid committing it, and the harm caused will be proportional to the harm avoided, the person has a defence for committing the offence in question.8

A. The Trigger and the Response

The primary difference between these three defences is the limitation (or lack of limitation) on the trigger and the response. In the case of necessity, the trigger is any urgent situation of imminent peril, and the response is to commit some offence. In the case of the recently repealed

---

6. See R v McIntosh, [1995] 1 SCR 686, 36 CR (4th) 171 [cited to SCR] (the Supreme Court observed that sections 34–37 were “unbelievably confusing” at para 16).
7. RSC 1985, c C–46, s 34(2), before re-enactment by Citizen’s Arrest, supra note 2 [Criminal Code (pre-2012)].
8. This was the point that was authoritatively determined in Ryan: “the defence of duress is available when a person commits an offence while under compulsion of a threat made for the purpose of compelling him or her to commit it”. Supra note 1 at para 2 [emphasis in original].
versions of self-defence, the trigger was an assault by another person, and the response was specifically to assault that person. In the case of duress, the trigger is a threat from another person made for the specific purpose of compelling the accused to commit an offence, and the response is to commit that offence.

Let us look at that same information presented in chart form:

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Necessity</td>
<td>An urgent situation of imminent peril</td>
</tr>
<tr>
<td></td>
<td>Commit some offence</td>
</tr>
<tr>
<td>Old Self-defence</td>
<td>An assault</td>
</tr>
<tr>
<td></td>
<td>Commit assault</td>
</tr>
<tr>
<td>Duress</td>
<td>A threat of death or bodily harm unless one commits a particular offence</td>
</tr>
<tr>
<td></td>
<td>Commit the particular offence</td>
</tr>
</tbody>
</table>

Presenting the requirements for the defences in this way helps to clarify the relationship between them. The potential triggers for necessity are broad and open-ended and the response is equally open-ended. Self-defence and duress, on the other hand, have a specific trigger and require a specific response. This gives us a sense that necessity is the general case, while self-defence and duress are both special instances of necessity.

This is a reflection of what the Supreme Court meant when it observed that “[d]uress is . . . merely a particular application of the doctrine of ‘necessity’”. Duress is a subset of necessity because its trigger and its response both fall within the general categories defined by necessity. Subject to an elaboration below, the same could be said of the old self-defence provisions: the urgent situation was the assault by another person, and the offence committed was an assault on that person.

---

9. See Criminal Code (pre-2012), supra note 7, ss 34–37. This is something of an oversimplification. However, sections 34(1), 34(2) and 35 all required that the accused had been assaulted; section 37 also allowed for the possibility that the accused was using force to “prevent the assault”. Ibid.

To illustrate this with a diagram, one might tentatively represent the relationship between the three defences as follows:

![Diagram showing the relationship between self-defence, duress, and necessity]

This shows that if self-defence and duress did not exist, the factual circumstances to which they now apply could still be dealt with as cases of necessity. Saying that duress and self-defence are contained within necessity shows the interplay between statutory and common law defences. Section 8(3) of the *Criminal Code* preserves common law defences “except in so far as they are altered by or are inconsistent with this Act or any other Act of Parliament”. As necessity is a purely common law defence, it is preserved by section 8(3). Section 8(3) also prevents necessity from intruding on self-defence and the statutory version of duress in section 17 of the *Code*, since those are circumstances which have been altered by the *Criminal Code*, and since common law duress is a more specific rule, it takes priority over the general defence of necessity when it applies. It is the recognition of this relationship which has caused the Court to conclude that it would “be highly anomalous if the common law defence of duress were to be understood as based on substantially different juridical principles from the common law defence of necessity”.11

However, the diagram is a slight oversimplification. To really see the relationship between the three defences, we need to acknowledge the trigger, the response and a third element: the filter.

---

B. The Filter

In addition to the trigger and the response, for a defence to succeed the crime committed must be an acceptable response to the trigger. What makes a response “acceptable” depends on the defence—this is the filter. In the case of necessity and duress, a demanding standard is set: the accused’s response must be “morally involuntary”. In contrast, self-defence requires only that the accused’s response be “reasonable”. Moral involuntariness requires that the accused’s will was overborne, that “normal human instincts cry out for action”. Reasonableness does not limit the choices nearly so much: a response can fall within the range of reasonable choices without being, in essence, the only imaginable one.

This difference in filters reflects the different characterization of necessity and duress on the one hand and self-defence on the other. Necessity and duress are excuses: we do not applaud the accused’s behaviour, but we do forgive it. In contrast, self-defence is a justification: we do not merely excuse the behaviour, but actually approve of it. As the Court put it in Ryan, “while in a case of duress we excuse an act that we still consider to be wrong, the impugned act in a case of self-defence is considered right”. It makes sense, therefore, that there are fewer circumstances (and so a narrower filter) in which we are willing to say “you are not guilty because your actions were wrong but excusable” than in which we say “you are not guilty because you acted rightly”. Again, as the Court notes in Ryan: “Given the different moral qualities of the acts involved, it is generally true that the justification of self-defence ought to be more readily available than the excuse of duress.”

---

12. R v Perka, supra note 5; R v Hibbert, supra note 4; R v Ruzic, 2001 SCC 24, [2001] 1 SCR 687.
13. R v Perka, supra note 5 at 251.
14. Supra note 1 at para 25.
15. Ibid at para 26.
In that light, a more accurate chart would look like this:

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Filter</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Necessity</strong></td>
<td>An urgent situation of imminent peril</td>
<td>Moral involuntariness</td>
</tr>
<tr>
<td><strong>Old Self-defence</strong></td>
<td>An assault</td>
<td>Reasonableness</td>
</tr>
<tr>
<td><strong>Duress</strong></td>
<td>A threat of death or bodily harm unless one commits a particular offence</td>
<td>Moral involuntariness</td>
</tr>
</tbody>
</table>

The effect of a more relaxed filter for self-defence is that, unlike duress, self-defence is not entirely contained within necessity. In some cases of self-defence (under the old provisions) an accused will have had no choice but to react as she did, meaning that her response would have met the moral involuntariness standard. In those cases, the accused could have successfully pleaded the common law defence of necessity had it not been replaced by statutory self-defence. However, in other cases, a response could be reasonable without being morally involuntary. In such cases, although self-defence is available, the strict filter for necessity would not have been passed.

Therefore, the relationship between the three defences would more accurately be represented in this way:

16. That seems to have been the governing theory behind the use of deadly force in section 34(2)—the accused was required to believe that “he cannot otherwise preserve himself from death or grievous bodily harm”. In section 35, the accused had to believe that the deadly force was “necessary in order to preserve himself from death or grievous bodily harm”. See Criminal Code (pre-2012), supra note 7, ss 34–35.
II. The Evolution of Duress and the Implications of *Ryan* for Necessity

A. The Evolution of Duress

To understand why *Ryan* has implications for the defence of necessity, we have to consider the evolution of the defence of duress over the past forty-six years. The way duress has changed, and the way *Ryan* solves a long-standing problem, are both interesting in themselves and instructive for necessity.

The most sensible way to understand the development of the law of duress is as a response to the ill-advised decision by the Supreme Court of Canada in 1966 in *R v Carker*. Before *Carker* we had a single understandable defence of duress that provided an excuse to a person who was ordered to commit a crime under threat. *Carker* undermined that situation, and courts have been working to undo its effect ever since. It took forty-six years, but *Ryan*, in effect, realized that goal.

(i) The Origin of the Problem: *Carker*

*Carker* was a prisoner in an institution. His fellow inmates were engaged in a protest and were destroying the plumbing in their individual cells, in which they were locked at the time. Carker did not take part until, in the Supreme Court’s words, “a substantial body of prisoners, shouting in unison from their separate cells, threatened . . . that if he did not break the plumbing fixtures in his cell he would be kicked in the head, his arm would be broken and he would get a knife in the back at the first opportunity”.18

*Carker* raised section 17 of the *Criminal Code*, the statutory defence of duress, but to no avail. Section 17 provides a defence to a person who is faced with “threats of immediate death or bodily harm from a person who is present when the offence is committed”.19 *Carker* was found to fall short in two ways. First, the Supreme Court upheld the trial judge’s

17. (1966), [1967] SCR 114, 60 WWR 365 [cited to SCR].
18. *Ibid* at 117 (note that he was literally threatened with harm to life and limb).
19. RSC 1985, c C-46, s 17.
decision that because the threats were of future harm, they were not “immediate”\(^2\). Second, because the other prisoners were in their own cells, they were not “present” when Carker committed the offence.\(^2\) As a result, the defence of duress was not available to him.

In retrospect, it is easy to see this conclusion as ill-conceived. Perhaps the Court had concerns about the defence getting out of hand in prison populations, or perhaps some other policy concern motivated its decision. If so, it was not apparent from the decision. The Court contented itself merely with adopting a narrow and literal interpretation of the words “immediate” and “present” and allowing that to dictate the result.

At a policy level, this very technical approach missed the point of the defence of duress. As the Court has observed in \textit{Ryan}, “the issue is not the immediacy or imminence of the threat, but whether ‘the accused failed to avail himself or herself of some opportunity to escape or to render the threat ineffective’”\(^2\)

There seems little question that if the Supreme Court of Canada had, in 1967, asked itself that question, Carker would have been acquitted.\(^2\) It did not, and we were left with a highly restrictive precedent. And as I have said, the history of the law of duress for the past forty-six years has been an attempt to undo it.

The first manoeuvre to avoid the implications of \textit{Carker} came in \textit{R v Paquette}.\(^2\) Paquette was forced by his two co-accused to drive them to the place where they planned to commit a robbery, and those co-accused in fact killed someone in the course of that robbery. Paquette, along with the other two, was charged with murder. The evidence showed that he had initially refused to drive the others and only did so at gun point. He was threatened with “revenge” if he did not wait, though in fact he drove

\(20.\) \textit{R v Carker}, supra note 17 at 118.

\(21.\) \textit{Ibid.}

\(22.\) \textit{Supra} note 1 at para 58, citing \textit{R v Langlois} (1993), 80 CCC (3d) 28 at 50, 19 CR (4th) 87 (Que CA). Similarly, the Court had earlier observed, “s. 17’s reliance on proximity as opposed to reasonable options as the measure of moral choice is problematic”. \textit{R v Ruzic}, \textit{supra} note 12 at para 88.

\(23.\) At a minimum, the questions asked would have been quite different, such as whether protective custody was available and a reasonable option given the length of the accused’s sentence.

\(24.\) (1976), [1977] 2 SCR 189, 70 DLR (3d) 129 [cited to SCR].

(2013) 39:1 Queen’s LJ
around the block and tried to prevent the two other accused from getting back into the car when they emerged from the store.

Had he been obliged to rely on section 17, Paquette would have had no defence, essentially for the reasons that Carker did not. He was threatened with “revenge”, which is harm at a later date and therefore not “immediate”. Further, the people who threatened him were not in his presence when the offence of murder was committed. However, the Court rejected the proposition that “the principles of law applicable to the excuse or defence of duress or compulsion are exhaustively codified in s. 17 of the Criminal Code”. Specifically, they found that section 17 referred to a person who “commits” an offence, which they interpreted to mean the principal offender but not a party. This meant that the common law defence of duress had not been altered by statute insofar as parties were concerned, and so by virtue of section 8(3) of the Criminal Code (section 7(3) at the time), the common law defence continued to apply. The Court specifically observed that this meant the result was not determined by Carker.

As a means of avoiding the unfairness of Carker, this was an effective manoeuvre. On the other hand, in the abstract, it is not terribly persuasive to claim that section 17 only intended to deal with principals, not parties. Paquette created the anomalous situation that two versions of the same defence were operating in the criminal law. This is not a model of clarity, particularly when one takes into account that the Crown might be able to argue for an accused’s guilt either as a principal or as a party on the very same facts. Similarly, it is hard to reconcile with the fact that the Crown can put forward a theory of the case which allows the jury to accept either of two versions of events: one with the accused as the principal and another with the accused as a party. In the absence of a desire to avoid

25. Ibid at 193.
26. Ibid at 194.
27. See e.g. R v Harbottle, [1993] 3 SCR 306, 24 CR (4th) 137. The accused held the legs of a murder victim while another person strangled her. The accused might have been a party to the offence by aiding the person who killed her (as was conceded in the Court of Appeal), or his actions might themselves have been a cause of her death and he was therefore a principal offender. Ibid.
28. See e.g. R v Thatcher, [1987] 1 SCR 652, 39 DLR (4th) 275 (the Crown’s theory in a murder case was that the accused had either shot his wife or had hired someone to shoot his wife, but in either instance was guilty of murder).
Carker, the bifurcation Paquette created would have been unnecessary and undesirably complex.

This lack of clarity is only exacerbated by the fact that although Paquette confirmed the existence of the common law defence of duress, it did not enumerate the requirements of that defence. Nor, indeed, did any case set them out fully. In R v Hibbert, the Court did conclude that the common law defence of duress had to be very similar to the common law defence of necessity, and that it would be incoherent and anomalous to treat it in any other way. However, the Court specifically articulated only one element of the defence: that the accused must have had no safe avenue of escape.

(ii) The Second Manoeuvre: Ruzic

Paquette had given up some clarity in order to attain some greater fairness. The next significant step in the law of duress, the decision in R v Ruzic, bought more fairness at the cost of even greater uncertainty. Ruzic was a citizen of the former Yugoslavia just as it was becoming “the former Yugoslavia”. She smuggled drugs into Canada as a result of credible threats directed against her and her mother, who remained

29. R v Hibbert, supra note 4 at para 54.

As I noted earlier, the common law defences of necessity and duress apply to essentially similar factual situations. Indeed, to repeat Lord Simon of Glaisdale’s observation, “[duress] is . . . merely a particular application of the doctrine of ‘necessity’”. In my view, the similarities between the two defences are so great that consistency and logic requires that they be understood as based on the same juristic principles. Indeed, to do otherwise would be to promote incoherence and anomaly in the criminal law.

Ibid.

30. Ibid at para 55. Hibbert also reversed a finding in Paquette which is relevant to duress, but is not actually one of the elements of the defence. Paquette had held that a person who only cooperates in the commission of an offence does not have the mens rea for party liability under section 21(2). This conclusion would make it unnecessary to resort to the defence of duress in the first place. In Hibbert the Court reversed that conclusion, holding that a person cooperating only under threat of force does have the mens rea for party liability (under either sections 21(1)(b) or 21(2)). In other words, only a successful defence of duress could save such a person from conviction. Ibid at para 43–44; R v Paquette, supra note 24.

31. Supra note 12.
behind. Unlike Paquette, she was the principal offender and so was forced to use section 17. As with Carker, the threats were of future harm and were made by someone not present when she committed the offence. The Court had to either convict her or find a new way around the rigours of Carker. The Court opted to do the latter, this time making use of the Charter.

Ruzic combined the results in two earlier cases, R v Daviault and R v Perka, to find that the Carker restrictions violated the Charter. In Daviault, the Court had concluded that it was a principle of fundamental justice that an accused cannot be convicted for physically involuntary behaviour. In Perka, in fully establishing the defence of necessity, the Court had found that that defence was based on the notion of “moral involuntariness”, which was meant to be closely analogous to physical involuntariness. In Ruzic then, the Court raised moral involuntariness to the status of a principle of fundamental justice, concluding that an accused could not be convicted for morally involuntary behaviour. Because section 17 permitted conviction for morally involuntary action, it was held to create a section 7 violation which could not be saved by section 1.

However—and this is the point that created confusion—the Court did not find that section 17 as a whole should be struck down. The only question in issue was whether an accused’s behaviour could be found to be morally involuntary due to threats, even where the threat was not one of immediate harm and the threatener was not present. As the answer to that question was yes, the only specific elements which violated section 7 were the immediacy and presence requirements. The Court therefore did not strike down the entire section, but the particular portions that imposed those requirements. This partial excision created further confusion in the law of duress.

Striking down the strict immediacy and presence requirements did not mean that nothing was left in their place. Rather, it meant that they were

33. Ibid.
34. Supra note 12 at para 47. The Court rejected the argument which had succeeded in the Ontario Court of Appeal, that an accused like Ruzic would not be morally blameworthy. In part because of Hibbert’s reversal of Paquette on the question of whether mens rea was proven, the Court concluded that an accused acting under duress was blameworthy. The point, as established by the distinction discussed at length in Perka, was that such a person’s actions should be excused, not justified. Ibid at paras 41–47.
replaced by their old common law equivalents. The requirement that the threatener be present was replaced by the “safe avenue of escape” rule which had been articulated in *Hibbert*. In addition, the Court in *Ruzic* elaborated on another aspect of the common law rule: the immediacy requirement. After considering the common law on duress in Canada, England, Australia and the United States, the Court concluded that the common law did not impose the strict “immediacy” requirement of section 17, but a slightly less rigid rule: the “strict temporal connection” requirement.

Post-*Ruzic*, then, it was fair to say that the law around duress was in a significant state of confusion. For pragmatic rather than principled reasons, two versions of the defence existed. Part of the statutory version had been struck down, but another part remained, and the common law version had never been fully articulated. This was the problem that the court set about to remedy with its decision in *Ryan*. It did so admirably.

(iii) The Current Approach: *Ryan*

Strictly speaking, the facts and legal issues in *Ryan* did not require the Court to do anything about this confusion. The accused was a battered spouse who had attempted to hire a hitman to kill her abusive husband from whom she was separated.\(^{35}\) The lower courts had found that the accused could not rely on self-defence, but that the statutory defence of duress could be extended to her circumstances, since she had committed the offence in response to threats. The Supreme Court of Canada very sensibly rejected this argument, which had misconceived the distinct purpose of duress. The Court concluded:

As we see it, the defence of duress is available when a person commits an offence while under compulsion of a threat made for the purpose of compelling him or her to commit it. That was not Ms. Ryan’s situation. She wanted her husband dead because he was threatening to kill her and her daughter, not because she was being threatened for the purpose of compelling her to have him killed. That being the case, the defence of duress

---

\(^{35}\) These were the facts accepted at trial. There is controversy over whether it accurately depicted the relationship between the accused and her husband, but that dispute is not important to the general development of the law of duress. See Archie Kaiser, “*Ryan*: A Troubling and Doctrinally Meandering Case Sets the Stage for the Law Reform Process and an Independent Inquiry” (2013) 98 CR (6th) 261.
was not available to her, no matter how compelling her situation was viewed in a broader perspective.36

As I have observed elsewhere, the Court could have stopped at that point.37 However, it went on, not only to explain its conclusion (and, controversially, to direct a stay of proceedings), but also to “bring more clarity to the law of duress”. In its words: “The patchwork quilt nature of the present law has given rise to significant uncertainty about the parameters of both the statutory and common law elements of the defence and the relationship between them.”38

First, the Court noted that post-Ruzic, four aspects of section 17 remained in place:

1. there must be a threat of death or bodily harm directed against the accused or a third party;
2. the accused must believe that the threat will be carried out;
3. the offence must not be on the list of excluded offences; and
4. the accused cannot be a party to a conspiracy or criminal association such that the person is subject to compulsion.39

In addition, Ruzic had made it clear that two aspects of the common law defence—the safe avenue of escape test, and a close temporal connection—had to be added to the statutory test. Interestingly, however, the Court in Ryan described Ruzic as having added not two but three elements: “(1) no safe avenue of escape; (2) a close temporal connection; and (3) proportionality.”40

One could be excused for not having seen Ruzic as reading a proportionality requirement into statutory duress. There is brief passing reference to proportionality as a factor in the common law version of the defence, but little elaboration on it.41 In considering whether the trial judge had correctly instructed the jury at Ruzic’s trial, the Court

36. R v Ryan, supra note 1 at para 2.
38. R v Ryan, supra note 1 at para 3.
39. Ibid at para 43.
41. See Kent Roach, Criminal Law, 5th ed (Toronto: Irwin Law, 2012) (the Court “hinted that that may be a requirement of proportionality” at 373).
noted only that he had properly described the safe avenue of escape and immediacy requirements, and pointed to those two elements as the only matters that trial judges should in the future include in their charges.\footnote{R v Ruzic, supra note 12 at 96. In the Court’s words:}

Whether this point was clear in \textit{Ruzic} is not important. What matters is that proportionality is now clearly part of the statutory defence. Even more important is the reason why. \textit{Ryan} clarifies that proportionality is not a requirement in addition to moral involuntariness, but is an aspect of moral involuntariness. This had been left ambiguous in previous discussions of both duress and necessity.

(iv) Moral Involuntariness Pre-\textit{Ryan} to \textit{Ryan}

The Court’s primary discussions of moral involuntariness before \textit{Ruzic} (and \textit{Ryan}) had come in cases about the defence of necessity. Necessity has been described as having three elements: an urgent situation of imminent peril, no reasonable legal alternative, and proportionality between the harm inflicted and the harm avoided.\footnote{R v Perka, supra note 5 at 251–52. For a more thorough encapsulation of the test, see \textit{R v Latimer}, 2001 SCC 1 at paras 28–29, [2001] 1 SCR 3.} Many commentators understood this formulation of necessity to mean that “moral involuntariness” was captured by the first two elements and that proportionality was a requirement in addition to moral involuntariness.

Although this was taken to be the Court’s meaning, commentators generally did not see it as a sensible approach. Trotter, for example, argued:

Proportionality is somewhat out of place in the conceptualization of necessity as an excuse because it speaks more to considerations of justification, rather than excuse. As the \textit{Latimer}}
Court said, it involves an evaluation of society’s values as to what is appropriate and what represents a transgression, and alludes to a comparison of the harm inflicted and the harm avoided. This language is perilously close to the “lesser of two evils” conceptualization of necessity, which the Court had previously rejected in *Perka*. More perplexing is that the *Latimer* Court has given proportionality pre-eminent status by recommending that, where it is apparent that the accused’s acts are disproportionate, it is preferable for the judge to rule out necessity without considering the other parts of the necessity test.  

The problem that Trotter identifies is exacerbated by the conclusion in *R v Latimer*. In *Latimer*, the Supreme Court of Canada held that the first two elements of the defence are to be assessed on a modified objective standard (by how the situation would appear to a reasonable person with the relevant characteristics of the accused), but that proportionality is to be measured on a purely objective standard that takes no characteristics of the accused into account. The Court’s rationale for this difference in *Latimer* was that the proportionality requirement is “fundamentally a determination reflecting society’s values as to what is appropriate and what represents a transgression”.

This difference in approach between moral involuntariness and proportionality seemed to drive a wedge between them. As Zoë Sinel pointed out:

> [B]ecause moral involuntariness does not speak to proportionality, only to reasonableness, several unpalatable consequences arise. Arguing from a reasonableness criterion, we would say that a reasonable person in the “clothes” of the accused would have acted in a similar fashion. Proportionality, on the other hand, is objectively measured: the act committed under duress must be proportionate to the threat the accused was under.

The problem with separating proportionality from moral involuntariness became even more acute when *Ruzic* (decided a few

---


45. *Supra* note 43.

46. *Ibid* at para 34.

months after *Latimer*) raised “moral involuntariness” to the status of a principle of fundamental justice. Given that conclusion, it would violate section 7 of the *Charter* to impose any additional requirements beyond moral involuntariness itself. In that event, if the proportionality requirement were an additional requirement it would have to be struck down.48

*Ryan* solves these problems. The Court made it clear that proportionality is not to be seen as an additional requirement, but as “inherent in the principle of moral involuntariness.”49 They explained that “only an action based on a proportionally grave threat, resisted with normal fortitude, can be considered morally involuntary”.50 Or elsewhere:

[T]he “moral voluntariness” of an act must depend on whether it is proportional to the threatened harm. To determine if the proportionality requirement is met, two elements must be considered: the difference between the nature and magnitude of the harm threatened and the offence committed, as well as a general moral judgment regarding the accused’s behaviour in the circumstances.51

In *Ryan*, then, it was established that proportionality is not, as most commentators had assumed, a utilitarian balancing of harms.52 Rather, proportionality is itself a moral judgment about how much harm we expect people to suffer before they succumb to the temptation to pass that harm along to another.

For the defence of necessity, this means that its three elements—an urgent situation of imminent peril, no reasonable legal alternative and proportionality—jointly capture moral involuntariness. With regard to duress, because proportionality is a part of moral involuntariness and *Ruzic* incorporated moral involuntariness into the statutory defence, section 17 now includes a proportionality requirement.

49. *R v Ryan*, *supra* note 1 at para 70.
50. *Ibid* at para 54.
(v) The Mental Element of Duress

The incorporation of moral involuntariness into section 17 also had an effect on the second requirement in section 17—namely, that the accused believes that the threat will be carried out. Although the statute sets out a subjective standard, the Court in Ryan concluded that it must actually be an objective one.

In part, the Court justifies this as simply better policy: the subjective standard, which is favourable to an accused, made sense given the very strict immediacy and presence requirements created in Carker. However, once the immediacy requirement was relaxed by Ruzic, the subjective standard had to be “tightened up” to an objective one. More importantly, the Court held that moral involuntariness required no less: “Considering that society’s opinion of the accused’s actions is an important aspect of the principle, it would be contrary to the very idea of moral involuntariness to simply accept the accused’s subjective belief without requiring that certain external factors be present.”

Ultimately, then, the Court’s new articulation of the statutory defence of duress in Ryan incorporates seven elements:
1. a threat of death or bodily harm directed against the accused or a third party;
2. a reasonable belief that the threat will be carried out;
3. no safe avenue of escape;
4. a close temporal connection;
5. proportionality;
6. the accused is not a part of a conspiracy or criminal association such that the person is subject to compulsion, and;
7. the offence is not on the list of excluded offences.

(vi) The Common Law Defence of Duress

After establishing the new statutory law of duress, the Court in Ryan turned its attention to the common law defence of duress. Given that the heavy lifting had already been done with the statutory defence, this task was relatively easy. Much of the statutory defence had been established by incorporating aspects of the common law defence. As a result, the first

53. R v Ryan, supra note 1 at para 52.
five elements of the common law defence are the first five elements of the statutory version.

The Court then considered the sixth statutory requirement: that the accused is not a member of a conspiracy subjecting him or her to compulsion. They noted that some courts have incorporated this element into the common law defence and, more importantly, that moral involuntariness requires its inclusion: “An accused that, because of his or her criminal involvement, knew coercion or threats were a possibility cannot claim that there was no safe avenue of escape, nor can he or she truly be found to have committed the resulting offence in a morally involuntary manner.”

Accordingly, the common law defence of duress is fully described by the first six of the seven elements that make up the statutory version.

Thus, with mutual adjustments to the statutory and common law versions of duress, the Court managed to turn them into essentially identical defences. Further, the unified defence did not incorporate the unreasonably strict immediacy and presence standard of Carker. Although it took forty-six years, we were back to the pre-Carker situation: a single, understandable defence of duress which was meant, in limited but appropriate circumstances, to provide an excuse to a person who was ordered to commit a crime and was threatened for that purpose.

(vii) The Differences Between Common Law Duress and Statutory Duress

The Court identified two remaining differences between common law duress and statutory duress. The first is that, as was established in Paquette and confirmed in Ruzic, the statutory defence applies to principals, while the common law defence is available to parties to an offence. The second is that the statutory version of the defence has a lengthy list of exclusions, whereas it is unclear in the Canadian common law of duress whether any offences are excluded.

Really these two differences should be regarded as one, because it is only the statutory list of exclusions that matters. If that list were gone, the statutory and common law defences would be identical, and it would

54. Ibid at para 77.
55. See ibid at para 83.
be of no consequence whatsoever that principals relied on one and parties relied on the other.

In fact, there is every reason to think that the list will go. The constitutionality of the list of exclusions did not arise on the facts of either Ruzic or Ryan, and so was not addressed in either case. However, it seems pretty clear that the Ruzic moral involuntariness argument, understood to incorporate proportionality in the way that Ryan dictates, leaves no room for the list. First, there are circumstances in which an accused could commit a listed offence under the threat of death—robbery, for example\textsuperscript{56}—and society would not expect the person to suffer death rather than commit the offence. In this situation, so long as the other elements of duress are met, the accused’s behaviour would be morally involuntary. Since the list of exclusions would deny the accused the defence (despite his morally involuntary act), the enumerated exclusion—in this case, robbery—would violate section 7 and have to be struck down.

Second, the list is not necessary to achieve the goal behind having a list. There is an unresolved dispute, for example, over whether common law duress (which does not have a list of exclusions) could ever be successfully pleaded as a defence to a charge of murder.\textsuperscript{57} In essence, the reasoning on this issue amounts to questioning whether we would expect the accused to suffer the harm threatened rather than commit the offence demanded. In other words, the argument revolves around whether the accused’s behaviour is proportional, and by extension, whether it is morally involuntary. Therefore, even if duress should be excluded as a defence to murder (or sexual assault, or high treason), the statutory list of exclusions is not necessary to achieve that result: the behaviour will not have been morally involuntary.

Really, the statutory list of exclusions is already meant to be a proportionality requirement, though a rough and ready one: “it is never

\textsuperscript{56} See e.g. R v Fraser (2002), 3 CR (6th) 308 (available on WL Can) (NS Prov Ct) [cited to CR].

\textsuperscript{57} For a case holding that it cannot, see R v Sandham (2009), 70 CR (6th) 203 (available on WL Can) (Ont Sup Ct). For a case holding it can, see R v C(P), 2012 ONSC 5362, 99 CR (6th) 116.
proportional to commit this offence” is what inclusion on the list means.\footnote{One could reach this same conclusion in an alternative way. If the inclusion of an offence on the list is found to violate section 7 of the Charter, then a remedy is to apply the common law defence of duress to that offence. See e.g. R v Fraser, supra note 56 at para 27. In that event (since the other elements of the statutory and common law defences are the same), a court would be required to ask whether committing the offence was a proportional response to the threat, instead of asking whether it was on the list.} \textit{Ryan}, however, has now explicitly incorporated a more sophisticated proportionality analysis into the statutory defence. There is no reason to do the same thing twice.

Once proportionality is understood to be an aspect of moral involuntariness, the offences on the list can be divided into two groups: those whose inclusion violates section 7 based on moral involuntariness, and those whose inclusion is redundant.\footnote{Properly, of course, any given analysis will be fact-dependent. One cannot say that committing arson is never morally involuntary, nor that it always is: the result will depend at a minimum on the nature of the threats.} There seems little reason to doubt, then, that the list of exclusions, which serves as the only distinction between statutory and common law duress, will eventually disappear. When it does, clarity in the law of duress will have been fully restored.\footnote{There is little reason to doubt that the list of exclusions will vanish. However, as will be argued in Part III, consideration of the defence of common law duress might vanish given the new self-defence provisions, which would actually re-introduce a lack of clarity.}

\textbf{B. The Implications of Ryan for Necessity}

I will now discuss the impact of the decision in \textit{Ryan} on the defence of necessity. As demonstrated in Part I, duress is simply a special case of necessity. Therefore, at a policy level, one would expect the defences to proceed along the same lines, or at a minimum, any differences would have sufficient justification. In fact, there is more than just that. Since changes to duress were dictated by an elaboration of moral involuntariness, which is the founding principle of both duress and necessity, necessity needs to be brought into line with duress to remain constitutional.

Post-\textit{Ryan}, the contours of duress (both statutory and common law) are almost entirely dictated by the Charter. In section 17, the Charter was the basis for replacing the strict immediacy and presence requirements, for changing the subjective standard to an objective standard, and for
explicitly including proportionality. Further, the Charter will likely be the basis for removing the list of exclusions. In the common law version of duress, the Charter principle of moral involuntariness dictates literally everything but the fact that, in step one, the trigger is a threat of death or bodily harm made for the purpose of compelling an accused to commit a crime, rather than some other sort of threat. And even with regard to that step, the Court observed: “Although, traditionally, the degree of bodily harm was characterized as ‘grievous’, the issue of severity is better dealt with at the proportionality stage, which acts as the threshold for the appropriate degree of bodily harm.”61 Since proportionality is an aspect of moral involuntariness, even the nature of the threat is in part dictated by the Charter.

As discussed in Part I, the defences of necessity and duress follow the same structure: a trigger, a filter and a response. Both share the same filter—namely, moral involuntariness. Building on Ruzic, the Court in Ryan explained the effects of moral involuntariness on duress. In essence, then, the law of duress is now what it must be. But since virtually all of those elements are dictated by a principle of fundamental justice which governs both duress and necessity, it is not simply good policy to conform the latter to the former: it is likely a matter of constitutional necessity.

(i) The Existing Test

The Court articulated the defence of necessity in Latimer as:
1. an urgent situation of imminent peril;
2. no reasonable legal alternative; and,
3. proportionality.
This formulation of necessity is much shorter than the newly articulated, six-element common law defence of duress. This is already a bit surprising, since duress is just a special case of necessity. In fact, I want to argue that this formulation of the defence is incomplete even on the existing necessity case law. Further, I will argue that the decision in Ryan requires some further changes.

61. R v Ryan, supra note 1 at para 55.
(ii) The Fourth Element of Necessity: The Accused Did Not Create the Urgent Situation

The defence of necessity was firmly established by the Court in Perka. In Perka the accused were in international waters with a shipload of marijuana they were taking from California to Alaska. When bad weather combined with various mechanical problems led them ashore in British Columbia, they were charged with importing cannabis into Canada and with possession for the purpose of trafficking. The Crown argued that as a matter of policy the accused should not be able to plead necessity because they were engaged in illegal activity at the time their emergency arose.62

The Court rejected that argument for two reasons. The first was that being in international waters while intending to smuggle drugs into the United States was not actually illegal under Canadian law. Second, and more importantly, the Court rejected the argument that illegality at the time the emergency arose was an automatic bar to claiming necessity.

However, this issue led the Court to observe that there was a somewhat analogous bar to claiming necessity: an accused cannot create his or her own emergency:

If the necessitous situation was clearly foreseeable to a reasonable observer, if the actor contemplated or ought to have contemplated that his actions would likely give rise to an emergency requiring the breaking of the law, then I doubt whether what confronted the accused was in the relevant sense an emergency. His response was in that sense not “involuntary”.63

The Court was very clear in Perka that this is a prerequisite to successfully pleading the defence of necessity. In fact, it ought to be explicitly recognized as a fourth requirement:

1. urgent situation of imminent peril
2. no reasonable legal alternative
3. proportionality
4. the accused did not create the urgent situation.

An accused who does create the emergency will not succeed in the defence. For that matter, if there is an air of reality to the first three

62. R v Perka, supra note 5 at 253.
63. Ibid at 256.
elements but no air of reality to the fourth, the defence ought not to go to a jury. However, on the current three-part formulation of the defence from *Latimer*, the air of reality test would have been passed, which is clearly an inappropriate result. A better description of the current law is this four-part articulation.

Note, of course, that this amounts to incorporating into necessity (the general defence) a general case version of the specific requirement in duress that “the accused is not a part of a conspiracy or criminal association such that the person is subject to compulsion”. In *Perka*, the Court specifically observed that the two requirements were equivalent. Further, the rationale offered in *Perka*—that an accused who has chosen to create a risk cannot later claim to have acted involuntarily because of that risk—is exactly the rationale offered in *Ryan* for incorporating the “no conspiracy” rule into common law duress. Both rules amount to “you can’t create your own compulsion”, whether that compulsion is by circumstances or by threats.

(iii) The Fifth Element of Necessity: A Reasonable Belief that the Situation Exists

A further way that necessity and duress should be brought more into line, which also does not require any change to the existing law, relates to the mental aspects of the defences. That is to say, does the accused subjectively have to believe that the particular state of affairs is true, or is it necessary that a reasonable person would believe it? Or (as is more commonly the case) is the test a modified objective standard that “takes into account the particular circumstances and human frailties of the accused”?66

The Court often talks about the mental aspects of defences, but does not routinely list them when enumerating the requirements of a defence. In *Latimer*, for example, the Court stated that the first two elements of necessity should be judged on a modified objective test while the third should be based on a purely objective test. Nonetheless, it did not list these requirements when enumerating the elements of the defence.

64. *Supra* note 43 at para 28.
65. *Supra* note 5 at 256.
Similarly, most of the mental aspects of duress have been discussed but tend not to be listed. In *Hibbert*, the Court spent some time discussing the mental state appropriate for judging the “safe avenue of escape” requirement, and eventually rejected both subjective and purely objective standards and adopted the modified objective standard. In *Ryan*, the Court considered the mental state that should accompany three elements: the threat, proportionality and participation in a conspiracy. The Court concluded that the mental elements are modified objective, modified objective, and subjective, respectively. Perhaps somewhat surprisingly, neither *Ruzic* nor *Ryan* explicitly discussed the type of belief which must accompany the “strict temporal connection”, though it seems likely that, like the safe avenue of escape, it should be judged on a modified objective test.

Of these mental aspects, only one is specifically enumerated when the Court sets out the two versions of the defence of duress in *Ryan*: “the accused must reasonably believe that the threat will be carried out”. The reason it is explicitly included in the test is because it was changed in the

---

67. *Hibbert* is significant on this point. Through a series of cases in the late 1980s and early 1990s, the Court engaged in a debate over whether to incorporate personal characteristics into the objective standard (a modified objective test) or to set the same standard for everyone (a purely objective test, or uniform objective test). The Court authoritatively determined that point, deciding that a modified objective test should not be used. See *R v Creighton*, [1993] 3 SCR 3, 105 DLR (4th) 632. However, throughout that same period the Court had been routinely incorporating personal characteristics into the objective test in the context of defences. The Court concluded that the proper question was not whether a reasonable person in the abstract would be provoked, but whether a reasonable person with the relevant characteristics of the accused (in that case young and male) would be provoked. See e.g. *R v Hill*, [1986] 1 SCR 313, 27 DLR (4th) 187. The Court decided that the proper test of reasonableness with regard to self-defence was what “the accused reasonably perceived, given her situation and her experience”. See *R v Lavallee*, [1990] 1 SCR 852 at 883, [1990] 4 WWR 1. In Lavallee’s case, someone with “the heightened sensitivity of a battered woman to her partner’s acts”. *Ibid* at 882.

The Court acknowledged that these two lines of authority ran contrary to one another, but resolved to continue both approaches nonetheless. That is, *Creighton* established that a modified objective test would not be used for offences, but *Hibbert* in effect held that the modified objective test should be the norm for defences. See supra note 4 at para 61.


69. *Ibid* (“[t]he first purpose of the close temporal connection element is to ensure that there truly was no safe avenue of escape for the accused” at para 68).

70. *Ibid* at para 81.
statutory version. The wording of section 17 created a subjective standard but the Court concluded that the proportionality requirement dictated that a modified objective standard must be used instead. It was therefore necessary to explicitly make the observation that the modified objective standard is required for both the statutory and common law versions of duress.

It is worth noting that the Court in *Latimer* had already found that the modified objective standard applied to the belief of an urgent situation of imminent peril.\(^71\) Writing the defence in a way which makes these existing parallels more apparent, we could say that necessity requires:

1. an urgent situation of imminent peril;
2. a reasonable belief that that situation exists;
3. no reasonable legal alternative;
4. proportionality; and,
5. the accused did not create the urgent situation.

Adding the second and fifth elements is simply pointing out parallels that already exist between duress and necessity. Indeed, a further parallel is that the equivalent elements of “no safe avenue of escape” and “no reasonable legal alternative” are both decided on a modified objective standard. Thus, necessity and duress are very similar not only in their general approach, as discussed in Part I, but in their specific requirements.

(iv) The Three Dissimilarities Between Duress and Necessity

There are only three dissimilarities between duress and necessity. First, duress lists a “close temporal connection” as a separate element, but that requirement is folded into the “urgent situation of imminent peril” element in necessity. Second, proportionality is based on a modified objective standard in duress but on a purely objective standard in

\(^71\). *Supra* note 43 at para 33.

The accused person must, at the time of the act, honestly believe, on reasonable grounds, that he faces a situation of imminent peril that leaves no reasonable legal alternative open. There must be a reasonable basis for the accused’s beliefs and actions, but it would be proper to take into account circumstances that legitimately affect the accused person’s ability to evaluate his situation.

necessity.72 Third, “not part of a conspiracy” is said in Ryan to be based on a subjective standard in duress, while “did not create the urgent situation” was said in Perka to be an objective test in necessity.

My argument is that for all three of these differences, the particulars of the element in duress were dictated by the moral involuntariness requirement. Therefore, in each of these three cases, the defence of necessity should be amended to conform to the particulars of the defence of duress.

a. The First Difference: An Urgent Situation of Imminent Peril

Let us first consider necessity’s requirement of “urgent situation of imminent peril”. In Latimer, this element was framed as a single requirement.73 It is less clear that the original intention of the Court in Perka was for those two issues to be a single consideration, since Perka refers to “[t]he requirement that the situation be urgent and the peril be imminent”.74 I argue that it is better to treat whether the peril is imminent as a separate question from whether the situation is urgent. Indeed, because the immediacy requirement in duress violated moral involuntariness, it seems that the Charter demands this approach.

Because both Perka and Latimer—the leading cases on necessity—predate Ruzic in articulating necessity, they predate the consideration of whether behaviour could only be morally involuntary if the threat was imminent. Ruzic established that the answer was no: threats of future harm could also cause a person’s behaviour to be morally involuntary.75

That argument should remain true whether the threats arise from another person, as with duress, or from circumstances, as with necessity.

72. See R v Ryan, supra note 1 (“[t]he evaluation of the proportionality requirement on a modified objective standard differs from the standard used in the defence of necessity, which is purely objective” at para 74).
73. Supra note 43 “[t]o begin, there must be an urgent situation of ‘clear and imminent peril’” at para 29).
74. Supra note 5 at 251 [emphasis added].
75. On the facts of Ruzic, the Court noted that the threat to the accused’s mother could only be given effect if the person who was to meet her discovered that she had not arrived, got in contact with the conspirator in Belgrade, and that conspirator then went to the location where the accused’s mother was to be found. Nonetheless, though not imminent, the threat made her behaviour morally involuntary.
The real question is not whether the harm will occur immediately, but whether “it was realistically unavoidable”.76 As a general proposition, the more imminent an event, the more difficult it will be to avoid, but imminence is not an absolute requirement.77

By way of analogy, consider the Court’s hypothetical in R v Lavallee: a person taken hostage and told she will be killed in three days’ time could reasonably kill her captor in self-defence if the opportunity arose on the first day.78 The very point of the analogy was to show that imminence, despite being generally a good indicator that the defence should succeed, was not an actual requirement of that defence.

The situation should be no different for a person who is in danger due to natural events rather than human intervention. Jeremy Horder, for example, imagines a group of people exploring a cave—“pot-holers”, as he refers to them—who become trapped when the fattest of their group becomes stuck in the only exit. If rising waters threaten to inevitably drown them all in three days’ time, and they use dynamite on the first day to blast the exit wider but kill the fat pot-holer in the process, should the defence of necessity be available to them?79 It is not necessary to settle this particular question to see that circumstances could become urgent before the peril becomes imminent. The certainty of death in three days time could make action on the first day not only reasonable, but morally involuntary.80

76. R v Ruzic, supra note 12 at para 29, citing R v Perka, supra note 5 at 250.
77. In Perka itself, the Court described the “imminence” requirement in this way: “until the time comes when the threatened harm is immediate, there are generally options open to the defendant to avoid the harm, other than the option of disobeying the literal terms of the law”. Ibid at 251, citing Wayne R LaFave & Austin W Scott, Criminal Law, 1st ed (St. Paul, Minn: West Publishing Co, 1972) at 388 [emphasis added]. My argument simply amounts to not ignoring the word “generally” in that sentence.
78. Supra note 67 at 889.
80. Consider as well the circumstances of Aron Ralston, a climber who amputated his own arm as the only method of escape after becoming trapped by a rock in an isolated location. The title of the movie recounting his circumstances—127 Hours—is an indication of how long he waited until responding to his urgent circumstances. Had he waited until his death was actually imminent, however, he certainly would not have survived. This is not an example in which a necessity claim arises, but it shows the distinction between
Just as in *Ruzic*, this does not mean that imminence is completely irrelevant to necessity. Clearly, the more immediate a threat, the more likely it is that a person’s response to it will be morally involuntary. But that should not make imminence a strict requirement, a *sine qua non* for necessity, any more than it was for common law duress. As the Court concluded in *Ruzic*, “A requirement that the threat be ‘imminent’ has been interpreted and applied in a more flexible manner.”

A more sensible interpretation of necessity would not make “urgent situation of imminent peril” a single element. Instead, it would treat “urgent situation” as one requirement, and incorporate a slightly relaxed “strict temporal connection” requirement as a separate step. Indeed, since the principle of moral involuntariness demanded this flexible approach in *Ruzic* and *Ryan*, such flexibility is probably required by the *Charter*.

The Court was aware of this difference between the two defences: it notes in *Ryan* that “[t]he strict immediacy or imminence requirement found in the defence of necessity was not imported into the common law defence of duress.” However, the Court did not engage in an analysis of whether the necessity standard remained constitutional. As I have argued, it clearly does not.

b. The Second Difference: The Mental Element of Proportionality

The second difference is the use of a purely objective assessment of proportionality in necessity but a modified objective standard for proportionality in duress. In this case, the Court was not only aware of, but also re-affirmed, that difference in *Ryan*. It held:

The evaluation of the proportionality requirement on a modified objective standard differs from the standard used in the defence of necessity, which is purely objective. While the defences of duress and necessity share the same juristic principles, according to Lamer C.J. in *Hibbert*, this does not entail that they must employ the same standard when evaluating proportionality. The Court in *Ruzic* noted that the two defences, although both categorized as excuses rooted in the notion of moral or normative involuntariness, target different types of situations. Furthermore, the temporality requirement for necessity

---

“urgent” and “imminent”. See *127 Hours*, 2011, DVD: (Los Angeles, Cal: Fox Searchlight Pictures, 2011).
81. *Supra* note 12 at para 86.
82. *Supra* note 1 at para 58.
is one of imminence, whereas the threat in a case of duress can be carried out in the future. It is therefore not so anomalous that the courts have attributed differing tests for proportionality, especially when we consider that the defences may apply under noticeably different factual circumstances.\textsuperscript{83}

The Court said that the two defences need not employ the same standard when evaluating proportionality, which would certainly be true if that standard were simply a matter of positive law. However, it was section 7 of the \textit{Charter} that dictated the result for duress. Proportionality is an aspect of moral involuntariness, and the Court observed in \textit{Ryan} both that “the ‘moral voluntariness’ of an act must depend on whether it is proportional to the threatened harm”,\textsuperscript{84} and that “[g]iven that the defence of duress ‘evolved from attempts at striking a proper balance between those conflicting interests of the accused, of the victims and of society’ . . . proportionality measured on a modified objective standard is key.”\textsuperscript{85} It is therefore hard to see how this leaves room for a stricter, yet still constitutional, standard for proportionality in necessity.

Further, the two rationales offered by the Court to explain the difference are not persuasive. The first was that the defences target different types of situations. While this is true—everything is different some way—it is not true in a relevant way. As the Court has noted, and as was discussed at length in Part I, duress is only a special case of necessity. If the defence of duress did not exist, necessity would in fact apply to those situations, by virtue of section 8(3) of the \textit{Code}. Therefore, it is hard to see how the “types of situations” targeted by the two defences could be different in a meaningful way.

The Court’s second rationale for the different standards is that duress allows for future threats, while necessity requires imminence. I have just argued that the imminence requirement in necessity must, as a result of the Court’s own section 7 argument, be adjusted to allow for future threats. And if the imminence requirement of necessity is brought into step with duress, as I argue it must, this rationale is also not persuasive.

Finally, note that assessing proportionality on society’s purely objective standard, but “urgent situation” and “no reasonable legal alternative” on a modified objective standard, made sense when it seemed as though

\textsuperscript{83.} \textit{Ibid} at para 74. \\
\textsuperscript{84.} \textit{Ibid} at para 72. \\
\textsuperscript{85.} \textit{Ibid} at para 73, citing \textit{Ruzic}, supra note 12 at para 60 [citation omitted].

S. Coughlan 113
proportionality was an additional requirement meant to provide a check on the moral involuntariness claim. Now that Ryan has made clear that all three elements are part of moral involuntariness, it is difficult to see how one would judge some aspects of moral involuntariness on one standard and some on the other. Accordingly, I suggest that based on Ryan, this aspect of necessity ought also to be brought into line with duress.

c. The Third Difference: Creating One’s Own Compulsion

Finally we come to the third difference between duress and necessity, the standard for the “no creating one’s own compulsion” requirement in the two defences. Although the Court set out different standards for the two defences—subjective for duress but objective for necessity—it is not clear that they are actually committed to this view. In Perka, when discussing the “create your own emergency” criterion for necessity, the Court spoke of it in objective terms, but never explicitly discussed whether an objective or subjective standard is appropriate. In Ryan, on the other hand, the Court engaged directly in an analysis of the competing arguments for the two standards, and concluded that “the subjective standard is more in line with the principle of moral involuntariness”.86 Considering that the Court only really analyzed the issue with regard to duress, and that it reached a conclusion based on the Charter, the case for the same standard applying to necessity is compelling.87

(v) The Final Test Post-Ryan

In conclusion, a more accurate articulation of the defence of necessity post-Ryan should incorporate requirements that the Court has already included in the defence, and should also reflect the now more fully understood demands of the principle of moral involuntariness. It would accordingly be as follows:

1. urgent situation;
2. a reasonable belief that that situation exists;
3. no reasonable legal alternative;

86. Ibid at para 80.
87. This is especially true as Perka is a pre-Charter case, and as Latimer never considered the standard for this consideration since it had not been listed as an element of the defence.
4. close temporal connection;
5. proportionality; and
6. the accused did not create the urgent situation.

The modified objective standard would apply to all elements but the sixth, which would be judged based on the accused’s subjective belief.

In essence, this formulation of necessity amounts to concluding that the decision in *Ryan* is a very good one, and we ought to do more of what it says. Further, if necessity is not brought in line with duress, it risks violating the *Charter* by denying a principle of fundamental justice: moral involuntariness. This proposed improvement to the defence of necessity is especially important because of the second major issue to be discussed in this paper: the possible demise of common law duress. As I discuss next, there is a good argument that duress has been clarified just in time to be rendered partly redundant by legislative changes to self-defence.

### III: Duress and Self-Defence

In Part I, I used charts and diagrams to represent the relationship between necessity, duress and self-defence. As the Venn diagram showed, duress was entirely contained within necessity, which was significant to the argument in Part II. However, the discussion in Part I was based on the recently repealed self-defence provisions in the former sections 34 to 37 of the *Criminal Code*. In this part, I will turn to discuss the significance of the newly-enacted section 34 of the *Code*, which is applicable in a broader range of circumstances. In particular, I will argue that the statutory defence of self-defence has now been broadened to encompass all of the situations that would otherwise fall within duress. Since self-defence is a statutory defence and statutes take priority over the common law, it must exclude at least the common law defence of duress.

Ironically, then, the legislative action which bracketed *Ryan* (since the provision was passed before *Ryan* but proclaimed after) precludes common law duress for reasons which are explained in *Ryan* itself. Before turning to that explanation, however, let us consider the new self-defence provision in its own right.
A. The Trigger, Filter and Response in the New Section 34

I argued in Part I that each of necessity, duress and self-defence consist of a trigger, a response and a filter. That is, in fact, exactly how the new section 34(1) is structured. It reads:

34. (1) A person is not guilty of an offence if
   (a) they believe on reasonable grounds that force is being used against them or another person or that a threat of force is being made against them or another person;
   (b) the act that constitutes the offence is committed for the purpose of defending or protecting themselves or the other person from that use or threat of force; and
   (c) the act committed is reasonable in the circumstances.88

The trigger is the use or threat of force (section 34(1)(a)). The response is committing some act which constitutes an offence (section 34(1)(b)). The filter is reasonableness (section 34(1)(c)).

The filter of reasonableness is, for the most part, unchanged from the previous version of self-defence. Section 34(2) lists a number of factors which “shall” be considered, many of which seem to have been specifically adopted from previous case law. For example, section 34(2)(e) and (f) call for consideration of:

   (e) the size, age, gender and physical capabilities of the parties to the incident; and
   (f) the nature, duration and history of any relationship between the parties to the incident, including any prior use or threat of force and the nature of that force or threat.

These factors seem to be adopted straight from the reasoning in Lavallee.89 Likewise, other factors in section 34(2) raise considerations which are quite familiar to us from the preceding discussion, such as these:

   (b) the extent to which the use of force was imminent and whether there were other means available to respond to the potential use of force; or
   (g) the nature and proportionality of the person’s response to the use or threat of force.

For that matter, the fact that section 34(1)(a) explicitly sets a reasonableness standard for the accused’s belief, while section 34(2)(e) and (f) incorporate

88. *Criminal Code*, supra note 19, s 34.
89. *Supra* note 67.
personal characteristics, seems to make clear that this defence is to be assessed on a modified objective standard.

No doubt there will be some teething pains while courts work out precisely how to use the new section. For example, is the reference in section 34(2)(c) to “the person’s role in the incident” meant to capture the former provisions’ distinction between those who face unprovoked assaults and those who are initial aggressors? On the whole, however, the “reasonableness” issue, and therefore the filter in self-defence, does not seem to be radically changed.\textsuperscript{90}

It also appears that the trigger has not changed, although it might seem so at first glance. The new provision allows an accused to respond not only to “force” but also to a “threat of force”, whereas the previous sections 34 and 35 provided that the defence only kicked in if the accused was “assaulted”. However, the old provisions were triggered more easily than their wording might immediately suggest. First, section 37 broadened self-defence to include preventing an assault, not merely responding to one. Further, under section 265(1)(b) an assault includes the behaviour

\textsuperscript{90} The complete list is:

In determining whether the act committed is reasonable in the circumstances, the court shall consider the relevant circumstances of the person, the other parties and the act, including, but not limited to, the following factors:

(a) the nature of the force or threat;
(b) the extent to which the use of force was imminent and whether there were other means available to respond to the potential use of force;
(c) the person’s role in the incident;
(d) whether any party to the incident used or threatened to use a weapon;
(e) the size, age, gender and physical capabilities of the parties to the incident;
(f) the nature, duration and history of any relationship between the parties to the incident, including any prior use or threat of force and the nature of that force or threat;
(f.1) any history of interaction or communication between the parties to the incident;
(g) the nature and proportionality of the person’s response to the use or threat of force; and
(h) whether the act committed was in response to a use or threat of force that the person knew was lawful.

\textit{Criminal Code, supra} note 19, s 34(2). Precisely why an addition to this list had to be labelled “f.1” when the section as a whole had not yet been proclaimed is something of a mystery.
of a person who “attempts or threatens, by an act or a gesture, to apply force to another person, if he has, or causes that other person to believe on reasonable grounds that he has, present ability to effect his purpose”. 91

That too allowed for self-defence by someone who was threatened rather than actually assaulted. Most importantly, in Lavallee, the Court specifically rejected the requirement that an assault must always be imminent before self-defence could be argued, finding that that accused’s claim could be based on the victim’s “threat to kill her when everyone else had gone”. 92 Accordingly, although the new section 34 explicitly mentions “a threat of force”, the trigger might be no broader than it was before. 93

However, it is no longer true that “self-defence is based on the principle that it is lawful, in defined circumstances, to meet force (or threats of force) with force”, as the Supreme Court characterized it in Ryan. 94 Although the trigger and filter are essentially unchanged, it is very clear that the response covers a much wider ground under the new self-defence provision. The old sections 34 to 37 justified one thing only: using force. An accused could only commit an assault in response to an assault. The new section 34 is not limited to an accused who has committed an assault. Rather, section 34(1)(b) says that a person is not guilty for performing “the act that constitutes the offence”. This allows for a completely open-ended range of possible responses.

For example, an accused threatened by a mob might steal a car and drive away quickly. Any of the potential offences—smashing a window, taking the vehicle, driving dangerously, driving while impaired—might be “the act that constitutes the offence”. Consequently, section 34 might provide a defence to each offence. There are other examples: an accused might break into a building to avoid an attack; an accused might discharge a firearm, as a way of attracting attention or as a distraction; or an accused

91. Ibid, s 265(1)(b).
92. Supra note 67 at para 40. Note that the new provision retains imminence as merely a consideration, not a requirement.
93. Note that, as observed above in the discussion of imminent peril in necessity, that Lavallee’s relaxation of “imminence” in self-defence is parallel to Ruzic’s relaxation of the immediacy requirement in duress. In other words, for both defences the more immediate the threat the more likely it is that the defence will succeed, but in neither is an absolute strict temporal connection always required.
94. Supra note 1 at para 20.
might forcibly confine an attacker without actually applying force. Again, these acts would constitute offences other than assault, but could fall within section 34 as now phrased.

Let us return to the chart of defences from Part I, and include the new self-defence provision.

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Filter</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Necessity</strong></td>
<td>An urgent situation of imminent peril</td>
<td>Moral involuntariness</td>
</tr>
<tr>
<td><strong>Old Self-defence</strong></td>
<td>An assault</td>
<td>Reasonableness</td>
</tr>
<tr>
<td><strong>Duress</strong></td>
<td>A threat of death or bodily harm unless one commits a particular offence</td>
<td>Moral involuntariness</td>
</tr>
<tr>
<td><strong>New Self-defence</strong></td>
<td>Force or a threat of force</td>
<td>Reasonableness</td>
</tr>
</tbody>
</table>

It is readily apparent from the last two rows that duress is now contained within self-defence, just as it was contained within necessity. The trigger for duress is a threat of death or bodily harm made for the purpose of compelling a particular offence. A threat of death or bodily harm will necessarily be a threat of force, one of the triggers for self-defence. Further, any choice which is morally involuntary, the filter for duress, will at a minimum be a reasonable choice, therefore satisfying the filter for the new self-defence provision. Finally, the new self-defence provision allows any offence to be committed for the purpose of protecting oneself from the threat of force. That will be exactly the reason why an accused will have committed the offence in duress: another person said to the accused, “rob that store or I will assault you”, and the “act which constitutes the offence” is robbing the store. Every element of duress is contained within its corresponding element in self-defence.
In that event, to return to the Venn diagram, the situation has now become this:

\[ \text{NECESSITY} \]

\[ \text{NEW SELF-DEFENCE} \]

\[ \text{DURESS} \]

B. A Note on Assumptions

We all continue to think of section 34 as “the new self-defence provision”. This causes us to make certain assumptions about the circumstances in which it is meant to be used. However, many of those assumptions are not actually reflected in the language of the section, and judges must apply the law as it is written. There are, of course, some circumstances in which a change in statutory language is not meant to change the approach to the law in any fundamental way. It can become necessary for courts to decide whether a particular amendment is intended only to codify existing principles or is meant to be remedial and to sound a fresh start.\(^95\) It is very clear, however, that the new self-defence provision is a response to the many implicit and explicit calls on Parliament from the courts to amend the “unbelievably confusing” provisions which have

---

95. For an example of the Court finding that changes to the sentencing provisions in the Code were not just a codification of existing principles but instead a change which had to be read remedially, see *R v Gladue*, [1999] 1 SCR 688, 171 DLR (4th) 385. See also *Interpretation Act*, RSC 1985, c I-21 (“[e]very enactment is deemed remedial, and shall be given such fair, large and liberal construction and interpretation as best ensures the attainment of its objects”, s 12).
been replaced.96 There can be no room for doubt that courts are meant to start afresh with the new provision—the only question is in what way.

The scope of the new provision will be defined by the arguments made by defence counsel, and the extent to which those arguments succeed with judges. Judicial inertia could no doubt play a role. However, in principle there is no reason why section 34 must end up limited to the same factual circumstances as the old provisions.

C. Should Self-Defence and Duress Be Kept Distinct?

In Ryan the Court explained why self-defence and duress should be kept distinct. It was concerned with this distinction because the accused was arguing to expand duress into self-defence’s territory. The Court explained why that could not be allowed. Its reasons, however, are either no longer persuasive because they are based on the old provisions, or are not reasons why self-defence cannot encroach on duress’ territory.

The first argument was that the self-defence provisions allowed an accused “to meet force (or threats of force) with force”.97 While that was true at the time, I have just observed that this is no longer what the provision does. Second, the Court said that “self-defence is an attempt to stop the victim’s threats or assaults by meeting force with force; duress is succumbing to the threats by committing an offence”.98 Once again, this distinction no longer holds true, since succumbing to the threat by committing a crime could be, under section 34(1)(b), a person’s “act committed for the purpose of . . . protecting themselves . . . from that use or threat of force”. Certainly, these are arguments that one can expect defence counsel to make on appropriate facts, but those “appropriate facts” will be ones that previously would have supported a finding of duress.

---

96. See e.g. R v Eggleston (1997), 94 BCAC 241, 117 CCC (3d) 566 [cited to BCAC] (“I express again as many have before me, my hope that these self defence sections will be reconsidered by Parliament and proper amendments put in place” at para 32).

For use of the “unbelievably confusing” language by the courts, see also R v McIntosh, supra note 6; R v Pintar (1996), 30 OR (3d) 483, 2 CR (5th) 151 (CA); R v Lei (1997), [1998] 5 WWR 134, 120 CCC (3d) 441 (Man CA); R v Grandin, 2001 BCCA 340, 152 BCAC 228.


98. Ibid at para 20.
The Court’s other bases in *Ryan* for distinguishing duress from the old self-defence provisions actually reinforce the argument that the new provision has overtaken duress. Self-defence is exhaustively codified, while duress is partly a common law defence. The common law cannot be used to cover supposed “gaps” in an exhaustive codification without frustrating Parliament’s intention.99 That is an entirely sensible justification for not allowing a common law defence of duress to encroach on the statutory defence of self-defence. However, if the statutory defence is amended to govern the situations formerly governed by the common law, then that argument (and section 8(3) of the *Code*) says that the common law will give way.

Finally, the Court said that self-defence is a justification based on reasonableness, while duress is an excuse based on moral involuntariness. The latter must therefore be a narrower defence, available in fewer situations: “the justification of self-defence ought to be more readily available than the excuse of duress”.100 But that rationale—not to let the narrower defence apply where the wider one should govern—is not a rationale for limiting the application of the wider defence.

One might object that since self-defence is a justification but duress is an excuse, it is anomalous to allow potential duress claims to be assessed by section 34. In other words, we should hesitate before saying that a person who commits an offence against some innocent third party is “justified” simply because it was done under duress. That might be so. However, unlike the repealed self-defence provisions, the new section 34 makes no reference to the accused being justified. The language of the new provision only provides that if its conditions are met, the person “is not guilty of an offence”. This wording leaves it ambiguous as to whether the person is justified or merely excused. That the new self-defence provision provides a justification is merely an assumption based on the repealed provision.

Even if, on some particular set of facts, an accused’s behaviour does not meet the reasonableness standard, that does not mean that duress


would be available. When self-defence, as now articulated by section 34, is the governing law in an area, then the result, whether success or failure, is dictated by that law. This is exactly the point explained by the Court in *Ryan*:

These distinctive underlying principles of self-defence and duress take on added significance when we remember that in Canadian law, self-defence is exhaustively codified, whereas duress is an amalgam of statutory and common law elements. This means that the courts must take care not to use the flexibility of the common law to develop duress in ways that circumvent limitations and restrictions imposed by Parliament on the defence of self-defence. This would amount to judicial abrogation of parts of the *Criminal Code*.101

With section 34 of the *Code* broadened, the result of this reasoning is that the common law defence of duress is narrowed—narrowed, I argue, to the point of exclusion.

Conceptually, it is equally true that self-defence now overlaps with the statutory defence of duress. An offence committed under threat of force will be an offence committed for the purpose of protecting oneself. This overlap could have practical consequences. Consider, for example, an accused who commits murder under threat. Statutory duress is not available because murder is in the list of exclusions, but self-defence can be pleaded.102 In that event such an accused would have the incentive to look to section 34 rather than section 17.

However, it seems unlikely that section 17 has been rendered redundant. Although a new statutory provision can exclude the common law, two statutory provisions have equal status and must both be given meaning. Under old rules of statutory interpretation one might argue that the particular takes priority over the general.103 More probably, however, the Court will apply a modern approach: “the words of an Act are to

102. As noted earlier, there is some question whether the list can survive *Charter* scrutiny, but for the moment it is still in place.
103. See *Pretty v Solly* (1859), 53 ER 1032 at 1034, cited in *James Richardson & Sons, Ltd v MNR*, [1984] 1 SCR 614, 9 DLR (4th) 1 [cited to SCR] (“[t]he rule is, that wherever there is a particular enactment and a general enactment in the same statute, and the latter, taken in its most comprehensive sense, would overrule the former, the particular enactment must be operative, and the general enactment must be taken to affect only the other parts of the statute to which it may properly apply” at 621). See also *Charlebois v Saint John (City)*, 2005 SCC 74 at para 46, [2005] 3 SCR 563, Bastarache J, dissenting.
be read in their entire context . . . harmoniously with the scheme of the Act”.\textsuperscript{104} The Court could then give section 17 meaning by concluding that it applies to a subset of those circumstances where section 34 would also apply. For example, the Court may decide that threats which are contingent on failing to commit an offence are to be singled out in the Code from other threats and subject to the special requirements of section 17. As statutory duress and self-defence are both statutory provisions, spheres will need to be found for each. Common law duress will disappear, but statutory duress will remain. As a result, the progress made in Ryan to unify duress will be undone. Principals who are compelled to commit an offence under threat of harm will continue to rely on section 17, while parties, who would have relied on common law duress, must now raise section 34.

\textbf{D. If Ryan Was Considered Today}

As a thought experiment, it is worth considering what would have happened had the new section 34 been in place at the time Nicole Ryan was first brought to trial. Would her counsel have had any incentive to argue for an unprecedented and unusual interpretation of duress?

Had she argued section 34, she would have had to prove three elements: that she believed on reasonable grounds that a threat of force was being made against her or another person; that the act constituting the offence was committed for the purpose of defending herself; and that the act was reasonable in the circumstances. First, all three levels of court accepted that she reasonably believed that her husband “was threatening to kill her and her daughter”\textsuperscript{105} Second, all levels of court believed that she hired the hitman for the purpose of defending herself. But was that act reasonable in the circumstances? There are many factors listed under section 34(2), and room for argument, but the claim made by the accused and accepted by the courts was that her behaviour was morally involuntary—that she had no other choice about how to behave. If that is so, it is hard to see how her choice could have been an unreasonable one.

\begin{itemize}
\item \textsuperscript{104} \textit{Bell ExpressVu Ltd v Rex}, 2002 SCC 42 at para 26, [2002] 2 SCR 559, citing Elmer A Driedger, \textit{Construction of Statutes}, 2d ed (Toronto: Butterworths, 1983) at 87.
\item \textsuperscript{105} \textit{R v Ryan}, \textit{supra} note 1 at para 2.
\end{itemize}
It seems that the accused in *Ryan* could have used the new section 34. Indeed, one could make the same argument for the accused in *Hibbert*—he called his friend Cohen to the lobby (which amounted to committing the offence of aiding an aggravated assault) for the purpose of protecting himself from Quasi’s use of force—or the accused in *Paquette*—who drove his two co-accused to the Pop Shoppe (thereby aiding in a murder) for the purpose of protecting himself from their threats of revenge.

The irony here is palpable. For forty-six years courts struggled with the impact of *Carker*, which led to the bifurcation of duress—one applying to principals and the other applying to parties. Finally, in the very well-thought-out and sensible decision in *Ryan*, the Court solved all but a small residual issue. *Ryan* restored us to the desirable situation of having, in effect, one defence of duress that applied to everyone. And then, less than two months later, Parliament proclaimed a law which, because of the reasoning in *Ryan* itself, puts us back in the position of having one law for principals and another for parties.

*Plus ça change, plus c’est la même chose.*