Proportionality review in appellate courts: a wrong turning?

Lord Sales*, Justice of the UK Supreme Court

ALBA annual lecture, November 2020

Introduction

In this lecture I want to ask whether appellate courts are at risk of making a wrong turn in their approach to reviewing proportionality assessments by first instance courts. Lines of authority are diverging, and it is unclear how they should be reconciled. The courts following one line do not refer to cases in the opposing line in order to interrogate what they are doing. We need to think more carefully about how the authorities can be reconciled and about what the underlying principles ought to be.

In some cases the Court of Appeal appears to be moving towards a proposed general position, applicable across the whole field of proportionality review, that they will uphold the assessment made by the first instance court provided it has directed itself correctly and has reached a result which is reasonable and cannot be said to be wrong. According to this approach the appeal court will not make its own proportionality assessment unless the first instance court has stepped outside the generous parameters allowed to it. I will call this the proposed standard approach.

Whilst this approach seems appropriate in many cases, I suggest that there are strong reasons to do with the rule of law and the proper role of appellate courts in our system - to provide

---

* With thanks to Anisa Kassamali for research assistance and to Timothy Endicott for comments on a draft. The usual caveat applies.
guidance and authoritative determinations of law - why the approach is wrong in some cases. In some situations, it would be an unjustified abdication of responsibility for an appellate court to confine itself to review of what the first instance court has decided, rather than making its own judgment and pronouncing on the legal position in an authoritative way which is applicable across a range of cases. In turn, this raises the question of how to formulate a test to delineate the boundary between the two classes of case.

The first line of authority is exemplified by the *Antique Cultural Treasures* case in 2018.\(^1\) It concerned the Ivory Act 2018, which introduced wide ranging prohibitions on the domestic and international trade in ivory. The judge at first instance in the High Court rejected a claim that the Act constituted a disproportionate restriction on EU free trade rights and rights to protection of property under the EU Charter of Fundamental Rights and Article 1 of Protocol 1 to the ECHR. This was so even though the effect of the legislation included preventing trade in existing ivory artefacts, thereby depriving them of all commercial value in their owner’s hands, without putting any compensation scheme in place. The Court of Appeal dismissed the appeal. This was on the basis that it was not for the Court of Appeal to re-perform the proportionality assessment, but to assess whether the judge’s assessment was justified.\(^2\) The court found that it was. The judge had directed himself correctly as to the law and his assessment could not be said to be unreasonable or wrong.

Obviously, on this approach whoever wins at first instance will be delighted. The very limited form of review by the court on an appeal gives them a huge advantage. However, in many

---

\(^1\) *R (Friends of Antique Cultural Treasures Ltd) v Secretary of State for the Department of Environment, Food & Rural Affairs* [2020] EWCA Civ 649. It was followed in *Independent Workers Union of Great Britain v The Mayor of London* [2020] EWCA 1046. For another example see *Lord Chancellor v McCloud* [2018] EWCA Civ 2844.

\(^2\) See para [6].
proportionality assessments the conclusion of the court or tribunal at first instance might be reached very much on balance, and in circumstances in which another judge could reasonably have come to a contrary conclusion. Yet, on the proposed standard approach, the appellate courts simply sign off on the proportionality assessment at first instance, provided only that it can be regarded as justifiable, without forming their own independent view on the available material whether a measure is proportionate or not. This puts a huge premium on the judgment of the first instance judge. It makes them the main decision-maker, to whom the appeal court must defer. The appellate court might disagree with the judgment at first instance, but still be bound to endorse it. Is this always appropriate?

Now contrast the judgments in the Supreme Court in R (Nicklinson) v Secretary of State for Justice. The case concerned the question whether the criminalization by statute of giving assistance for suicide by persons with locked-in syndrome, trapped in their bodies in severe distress and unable to terminate their lives on their own, involved a breach of the right to protection for private life under article 8 of the ECHR. This turned on whether the existing law imposed a disproportionate restriction on that right, having regard to the other interests at stake, such as avoiding the risk of vulnerable individuals being put under pressure to end their lives. None of the nine justices on the panel thought that the way to answer this was to defer to the view of the first instance court. This was a case literally involving questions of life and death, which called for a decision which balanced competing fundamental moral values. It also called for a decision which balanced competing fundamental institutional values in terms of whether the court should or should not strike the balance of moral values itself or accept the balance as struck by Parliament. Surely the justices were right that this was a context in which the public had a right to expect them to exercise their own judgment, as members of the highest court, regarding the

legal outcome. It would have been an abrogation of their duty if they had deferred, not to Parliament, but to the views of a court of first instance.

The same pattern appears in a series of cases on social welfare reforms which were highly controversial, in *SG* and *DA* in the Supreme Court on the compatibility of the benefits cap with Convention rights and in *SC* in the Court of Appeal on the compatibility of the two child limit on child tax credits with Convention rights. The proportionality of measures in the application of Convention rights was at the heart of what these cases were about. These are major cases in which the courts pass judgment on the compatibility with human rights of politically controversial legislation of general application, with wide ramifications for the public. They also raise major constitutional issues regarding the relationship between the courts and government and the courts and Parliament. Throughout this litigation, the appellate courts understood that it was for them to give their own judgment on the proportionality of the measures in question. They did not defer to the judgment of the first instance court.

These lines of authority have developed without cross-referencing each other. This is a recipe for confusion. We are at a stage when the appellate courts should be trying to work out in a principled way which sort of approach is applicable. We should not be trying to grasp at a ‘one size fits all’ approach.

Part of the problem arises from the wide variety of cases in which a proportionality assessment is required. Proportionality review covers a wide field and very different sorts of measure. Sometimes, the proposed standard appellate approach seems justified, particularly when the

---

6 *R (SC) v Secretary of State for Work and Pensions* [2019] EWCA Civ 615. The case is currently on appeal to the Supreme Court.
measure in question is a one-off decision affecting only the particular parties in the litigation. But sometimes the approach feels like an abdication of responsibility by an appellate court. I will argue that a more modulated approach to appellate review is justified in principle and is capable of being worked out in practice.

.Features of proportionality assessment

The basic four stage format of proportionality assessment is well established. The fourth stage, proportionality *stricto sensu*, requires there to be proportionality between the effects of the measures on countervailing rights or interests and the objective that is achieved.7

Very often, proportionality assessment involves balancing incommensurable values against each other. There is no simple metric by which to test whether a proper balance has been struck or not. An evaluative judgment has to be made, whether some particular value should be given greater priority than another.8 In the *Antique Cultural Treasures* case the proportionality assessment required weighing up the value of setting a clear rule to ban the ivory trade, providing guidance to other countries as an international standard, against the value of establishing a compensation scheme for those holding ivory items whose value was affected by the ban.

---

7 See e.g. *R (Aguilar Quila) v Secretary of State for the Home Department* [2011] UKSC 45; [2012] 1 AC 621, [45] per Lord Wilson; as regards stage (iv), the question is “whether the impact of the rights infringement is disproportionate to the likely benefit of the impugned measure”: *Bank Mellat v HM Treasury (No. 2)* [2013] UKSC 39; [2014] AC 700, [74] per Lord Reed.

The absence of an agreed metric creates a tension between the proportionality test and the rule of law values of predictability in the application of the law and the wish to avoid arbitrariness in its effect across similar cases.

There are two consequences of incommensurability. First, it expands the area within which a first instance judge will act reasonably when making a decision either that a measure is proportionate or that it is disproportionate. The result in any particular case is under-determined by the proportionality test, with the result that the area within which the judge can legitimately choose different results in applying it is enlarged. Accordingly, where the proposed standard appellate approach is followed, it reduces the range of cases in which an appellate court will intervene to overrule the first instance decision.

Secondly, since the absence of a simple metric of comparison for the balancing exercise means there is no clear right answer, this increases the importance of authoritative judgment within the system. What counts in arriving at a determinate answer when there is a dispute about whether a particular measure is proportionate or not is the authority invested by the system in a particular decision-maker. In other words, whose view about the correct answer does the system say should count most in arriving at the answer in any given case? The effect of the proposed standard appellate approach is to endorse a decision made at first instance, rather than giving authority to the appellate court to make its own judgment.

The issue of authority to decide features in the debate about the extent to which a court should adopt an attitude of “deference” or respect for assessments made by Parliament. Parliament is a

---

9 i.e. the sort of case in which in *In re B (a Child)* [2013] UKSC 33; [2013] 1 WLR 1911, Lord Neuberger said that an appellate court should not intervene: para [93].
body which can reflect a wide range of perspectives and has democratic legitimacy. The debate about the width of the margin of appreciation allowed to Parliament is about the extent to which the views of Parliament regarding the appropriateness or proportionality of a measure should be treated as authoritative and hence should govern the outcome in a particular case. In Nicklinson the Justices were divided about this.\textsuperscript{11} The courts adjust the ambit of the margin of appreciation to reflect a range of factors. In review of the compatibility of primary legislation with Convention rights, an important factor tending to widen the margin of appreciation is the democratic legitimation attached to a decision of Parliament.\textsuperscript{12} The extent of the margin reflects the varying underlying strength of the argument for deference or respect in different contexts.

Widening the margin of appreciation in favour of a decision made by Parliament is a form of adjustment on grounds of relative authority, in the sense of the relative authority of courts and Parliament flowing from democratic legitimation in making the relevant balancing decision. That is, of course, a different form of authority from that which exists between courts at different levels in the court system.\textsuperscript{13} But looking at matters in terms of the allocation of authority on the two dimensions - the democratic dimension and the rule of law dimension - provides insight into a possible approach for regulating the exercise of authority within the court structure, by allocation of different responsibilities between a first instance court and an appellate court.

\textsuperscript{11} One group of Justices was of the view that the statutory criminalisation of assisting suicide was a matter pre-eminently for Parliament to decide, so that there was no disproportionality: see [2013] 1 WLR 1911 at [233]-[234] (Lord Sumption), [267] (Lord Hughes) and [196]-[198] (Lord Reed); a second group deferred judgment on the point because Parliament was on the point of considering a proposal for amending legislation: see [113]-[118] (Lord Neuberger), [188] and [190]-[191] (Lord Mance), [196]-[197] (Lord Wilson) and [293] (Lord Clarke); and a third group considered that the existing legislation was disproportionate on any view: see [299]-[321] (Baroness Hale) and [326]-[361] (Lord Kerr).

\textsuperscript{12} See e.g. R (Humphreys) v Revenue and Customs Comrs [2012] UKSC 18; [2012] 1 WLR 1545 and R (Z) v Hackney LBC [2020] UKSC 40 (in the context of the domestic application of Convention rights); and R (Lumsdon) v Legal Services Board [2015] UKSC 41; [2016] AC 697 (application of EU law).

\textsuperscript{13} For an account of the margin of appreciation as a doctrine which seeks to accommodate the tension in liberal democracies between the democratic principle and the principle of the rule of law, see Sales "Proportionality and the Margin of Appreciation" (n. 8), 182-183.
Similarly, in the case law of the European Court of Human Rights, debate about the margin of appreciation allowed by the Court to national institutions is about the extent to which the views of national authorities about the appropriateness or proportionality of a measure should be treated as authoritative for the purposes of the Court’s own assessment. The margin of appreciation will generally be found to be wide where the Court is examining a choice made by a democratically elected legislature in relation to a topic which is the subject of public debate and one on which opinions may reasonably differ in a democracy. In other cases, different factors may narrow the margin of appreciation, giving greater scope for the Court to rely on its own judgment. Again, modulation in the ambit of the margin of appreciation reflects the strength of the underlying justifications for giving greater or lesser decision-making authority to national institutions, or to the Court.

The debate regarding the role of an appellate court is, in a similar way, about the respective degree of decision-making authority which should be allocated between a first instance court and an appellate court. However, the arguments for according decision-making authority to a first instance court in relation to the assessment of the proportionality of a measure do not arise on grounds of the democratic credentials of that court. Instead, the justification for limiting an appellate court to review should reflect arguments regarding the allocation of decision-making authority between courts within a national legal system.

The force of those arguments is not uniform across all cases. So, in this area as well, it ought to be open to examination whether there are grounds for distinguishing between situations in which

---

16 Sales, “Proportionality and the Margin of Appreciation” (n. 8).
the appellate court should adopt a review role and those in which it should adopt a primary
decision-making role, and make the proportionality assessment for itself.

Experience with the fluctuating margin of appreciation allowed to Parliament in different
contexts suggests that it will not be possible to produce a simple bright-line rule to govern the
demarcation of roles. Instead, it seems more realistic to hope for a gradual clarification through
argument and decision, context by context.

A third significant feature of proportionality assessment is that it is often closely bound up with
factual findings. For example, if the factual findings are that a measure will have only a small
effect in promoting some legitimate interest but that it will have a major detrimental effect on
opposing rights or interests, that will be a strong basis on which it could be argued that it is
disproportionate. Usually fact-finding is a matter for the lower court. 17

A fourth feature of proportionality assessment is that outcomes in the application of a rule which
specifies that a measure shall be proportionate are under-determined by the rule itself. There is a
large area calling for the exercise of judgment on the part of a decision-maker in applying the
rule to the facts of a case. This means that the courts have a function to fill in the gap to provide
guidance as to how the proportionality norm will be applied in similar cases in the future. 18

Within the legal system, as between first instance and appellate courts, it is the function of
appellate courts to take the lead in providing such guidance. That is particularly important where

---

17 See eg United Policy Holders Group v Attorney General of Trinidad and Tobago [2016] UKPC 17; [2016] 1 WLR 3383. The lower court’s proportionality assessment was challenged on the facts, but the Privy Council found that the court had been entitled to make the findings it did: see [74].

18 “… by gradual articulation of authoritative guidance through which expectations can be focused and judicial behaviour managed”: Sales, “Proportionality and the Margin of Appreciation”, n. 11, p 181.
the measure in issue might itself be a rule of law or a public norm applicable to a wide range of cases.

The growing ambit and significance of proportionality review

Proportionality review has a growing ambit and significance. It is familiar to you all as a feature of EU law and ECHR case-law. It has also been explicitly adopted as the relevant legal standard in some domestic statutes, most obviously in various provisions of the Equality Act 2010.¹⁹ Proportionality has also increasingly come to be adopted as the relevant legal standard for a number of common law doctrines. It is the standard according to which the lawfulness of a public authority’s departure from a policy statement which creates a legitimate expectation will be judged.²⁰ It has also recently been incorporated by the Supreme Court into the test for the illegality defence in Patel v Mirza²¹ and into the test for the validity of a penalty or liquidated damages clause in a contract in Cavendish Square v Makdessi.²²

In EU law, proportionality is a concept which affects the interpretation and application of EU Treaties and legislation.²³ Domestic courts therefore have to apply the concept when giving EU law direct effect, and also when the EU Marleasing doctrine²⁴ of sympathetic construction

---

¹⁹ See in particular section 19 (indirect discrimination) and in sections 158 and 193, considered in R (Z) v Hackney LBC [2020] UKSC 40.
requires that they harmonise the interpretation of domestic legislation and EU legislation. In both cases, assessment of proportionality may directly affect the meaning of the legal instrument in question.

Usually, one would expect that interpretation of legislation would be a matter of law, for the court; and that on an appeal, it would again be for the appellate court to determine the meaning of the law in question authoritatively for itself.

However, there is scope for a dividing line to be drawn between the meaning of legislation and its application. If the effect of the legislation is that a measure may be put in place provided that it is proportionate, it might be said that the meaning of the legislation as a matter of law is established at that level of abstraction and that the question whether a measure introduced pursuant to that legislation is in fact proportionate and hence in accordance with the legislation is a matter of the application of the law to the facts. This division might be serviceable in terms of producing an allocation of responsibility between a first instance court and an appellate court. The interpretation of the law at the more general level should be regarded as the primary responsibility of the higher court, on which deference is not due to the lower court, while the application of the general law to the particular facts could safely be left as the primary responsibility of the lower court. There might be no need for the higher court to intervene unless the lower court has misunderstood the law to be applied (and so has misdirected itself) or has reached an unreasonable conclusion.

Under the *Marleasing* interpretive obligation, the dividing line between determining the true meaning of a statute and its application to a particular case on the facts can become blurred. The obligation only applies where, in the case of the claimant, EU law prescribes a particular result
which varies from the result given by the ordinary interpretation of the domestic legislation. The result prescribed by EU law may apply in a wide range of cases, or it may be limited to the case in hand and reflect a more specific proportionality analysis in that case.\(^{25}\) In the latter situation, it may be that the assessment whether the EU rights of the claimant are engaged or not in the particular case in light of a proportionality assessment can safely be regarded as the primary responsibility of the lower court.

Similar comments can be made about application of Convention rights set out in the ECHR, pursuant to the Human Rights Act. Again, there is a wide range of cases in which proportionality review of some kind may be relevant. In the welfare benefits cases, for example, the review was of the proportionality of a legislative scheme. In other cases, proportionality review may be required in relation to decisions taken under such a scheme. Again, the two sorts of cases can blend into each other. The obligation under section 3(1) to interpret legislation compatibly with Convention rights is an interpretive obligation closely similar to the \textit{Marleasing} obligation\(^{26}\) and similar comments apply.

The approach of the Court of Human Rights when conducting a proportionality assessment in relation to a national measure where a national court has itself carried out such an assessment first provides a potential analogy for our inquiry. The Court regulates the intensity of its review using the mechanism of a sliding scale margin of appreciation. Elsewhere\(^{27}\) I have suggested that the three axes along which the margin of appreciation operates are, first, to accommodate the tension between Convention rights and democratic decision-making; secondly, “to provide a space for the determinative application of local expertise or superior knowledge of relevant


\(^{26}\) \textit{R (Z) v Hackney LBC} [2020] UKSC 40, para 112.

\(^{27}\) Sales “Proportionality and the Margin of Appreciation” (n. 8), 184-186
circumstances within the national judicial and regulatory systems”; and thirdly, “to regulate the calls upon the [Court’s] time and attention” in recognition of its role as a sort of constitutional court of the European public order with limited capacity. The second and third of these are factors which are also relevant to the relationship in the domestic legal system between appellate courts and first instance courts. Thus, the Court of Human Rights adopts a non-intensive approach to review of factual findings made by a national court:

“As a general rule, where domestic proceedings have taken place, it is not the Court’s task to substitute its own assessment of the facts for that of the domestic courts and it is for the latter to establish the facts on the basis of the evidence before them. Though the Court is not bound by the findings of domestic courts and remains free to make its own assessment in the light of all the material before it, in normal circumstances it requires cogent elements to lead it to depart from the findings of fact reached by the national judicial authorities.”

Further, where superior national courts have made a sincere effort to follow the caselaw of the Court in applying Convention rights in national law, the Court will not readily intervene to impose a different view.

However, it is revealing that the Court refers to decisions by superior national courts. Its willingness to concede a margin of appreciation in the application of Convention rights depends on the national system bringing the judgment of its senior judges to bear on the issue. It is not obvious that the Court would be so willing to allow such a margin of appreciation to a decision

---

by a first instance judge which has only been subject to limited review by the superior courts.

Part of the justification given for the enactment of the HRA was that the domestic enforcement of Convention rights would give domestic courts the opportunity to make a contribution capable of influencing the development of the Strasbourg case law and to guide the Strasbourg Court in the application of those rights with full knowledge of and sensitivity to the domestic context. Those objectives would be liable to be undermined if within the domestic legal system critical judgments on the proportionality of, say, major legislative initiatives are in substance allocated to first instance courts.

Where a proportionality standard is expressly written into legislation as the standard against which to measure the actions of private persons, such as in the Equality Act 2010, it often seems possible to say that a court applying that standard is applying the law to the facts rather than engaged in a general formulation of a legal rule for general application. But where the assessment is made in relation to norms laid down by public authorities and applicable across a wide range of standard cases it is more difficult to detach the proportionality assessment from a statement of the law itself or an evaluation of its legitimacy as law.

In general terms, the more widely the proportionality standard spreads in our law and the greater the contexts in which it falls to be applied, the greater the danger that the self-denying ordinance

---

30 See the White Paper, Rights Brought Home: The Human Rights Bill, October 1997 (Cm 3782), para 1.14 (“British judges will be enabled to make a distinctively British contribution to the development of the jurisprudence of human rights in Europe”) and para 1.18 (“Enabling courts in the United Kingdom to rule on the application of the Convention will also help to influence the development of case law on the Convention by the European Court of Human Rights on the basis of familiarity with our laws and customs and of sensitivity to practices and procedures in the United Kingdom.”)

31 R (Z) v Hackney LBC [2020] UKSC 40, esp. para 56, approving the judgment of the Court of Appeal on this issue, including its observation at para 68 that application of this approach was particularly suitable where the first instance court had been presented with a mass of demographic and sociological evidence and required to assess its factual significance.
adopted by appellate courts in following the proposed standard approach to appellate review of proportionality assessments will be inappropriate in a range of cases.

The foundations of the proposed standard approach: how we got here

How did the proposed standard approach arise? Section 6(1) of the HRA imposes a duty on public authorities to act compatibly with Convention rights. That duty also applies to courts, including appellate courts. At first the view at Court of Appeal level was that in each case involving Convention rights the appellate court had to carry out a proportionality assessment afresh for itself. This clashed with the domestic procedural rule in England and Wales in CPR Part 52.11 that ordinarily appeals would proceed by way of a review of the first instance decision rather than a rehearing and that the decision at first instance would be overturned only if it was “wrong”.

The Supreme Court confronted and corrected this trend in its majority decision in Re B (A Child). The case concerned an application by a local authority for a final care order in relation to a child, with a view to the child’s adoption. The trial judge conducted a trial with oral evidence and concluded that the order should be made despite the objections of the parents. The judge held that the making of the care order was a proportionate interference with the article 8 rights of the relevant parties. On appeal the Court of Appeal held that it should carry out the proportionality assessment afresh for itself. By a majority the Supreme Court held that this was the wrong approach. The obligation of the appellate court under section 6 of the HRA did

---

32 Section 6(3).
33 Now CPR Part 52.21.
35 [2012] EWCA Civ 1475
36 Lord Neuberger, Lord Wilson and Lord Clarke; Lady Hale and Lord Kerr dissenting.
not require it to depart from its normal appellate function under CPR part 52.11, of secondary review of the trial judge’s decision. An appeal should be dismissed if the first instance judge had reached a conclusion on proportionality which an appellate court could not say was wrong, even if the court might not have come to the same view. However, the reasoning of the justices in the majority reflected the particular facts of the case. The proportionality assessment was in relation to a one-off decision in respect of which the judge was particularly well-placed to make the assessment, depending as it did on evaluation of oral evidence.

Other cases in the Supreme Court took a similar line. DB v Chief Constable of the Police Service of Northern Ireland 37 concerned a judicial review claim in relation to the policing of demonstrations and marches in Northern Ireland. Lord Kerr took the opportunity to make observations about the approach which an appellate court should take to reviewing findings by the first instance court. 38 He emphasised that an appeal court is unlikely to add to the accuracy of fact determination by the trial judge and that the trial on the merits at first instance should be “the main event” rather than a “tryout on the road”. He also explained 39 that the insight of a trial judge who has sat through the whole case and is fully familiar with the evidence may be far deeper than that of an appellate court, whose view of the case may be much more limited and narrow. Then, by extension of this reasoning, 40 Lord Kerr said that these points also had cogency in a case where no oral evidence had been given, including in judicial review proceedings. Again, the proportionality assessment was fact specific. It was not concerned with reviewing the proportionality of a norm of general application.

38 Paras 78-80.
40 at para 80
R (R) v Chief Constable of Greater Manchester\textsuperscript{41} involved judicial review of another one-off decision by a public authority. The first instance judge rejected a complaint of disproportionate interference with rights. The Court of Appeal and the Supreme Court dismissed the claimant’s appeals. The Supreme Court held that the test was whether the judge’s decision was “wrong” under CPR 52.11, but it was not enough that the appellate court might itself have reached a different evaluation.\textsuperscript{42}

Finally, I would mention my recent judgment in R (Z) v Hackney LBC.\textsuperscript{43} My involvement in that case is one reason I have become interested in this question. The case concerned the housing allocation policy operated by a charity, rather than a public authority. Challenges under the Equality Act were dismissed at first instance following a proportionality assessment in favour of the charity. The Court of Appeal applied the proposed standard approach and dismissed the claimant’s appeal. On appeal to the Supreme Court, the claimant did not challenge the application of that approach and we endorsed its application in the case. For good measure, the majority also said that we agreed with the proportionality conclusion reached at first instance. However, it seems to me that an important point about the case is that the measure in question was adopted by a private body in the management of its affairs, involved no deployment of state power or authority and did not have wide ramifications for the public.

In Z, however, we did pick up on one point.\textsuperscript{44} Although the appellate approach in Re B and R (R) v Chief Constable of Greater Manchester had been articulated by application of CPR Part 52.11, this is slightly odd at the level of the Supreme Court. The Civil Procedure Rules do not apply to the Supreme Court. Also, the issues that arise in this area are of general concern across all the

\textsuperscript{41} [2018] UKSC 47; [2018] 1 WLR 4079.
\textsuperscript{42} Paras 64-65 (Lord Carnwath).
\textsuperscript{43} [2020] UKSC 40.
\textsuperscript{44} See para 74.
jurisdictions of the United Kingdom, and the appellate approach should be the same across all those jurisdictions, whether or not the rule for England and Wales in CPR Part 52 applies or not. In other words, the approach to be adopted (and the test for whether the applicable appellate approach should be one of review or primacy decision-making) should be a matter of general legal doctrine, and the test for whether a decision of a court at first instance is “wrong” under CPR Part 52 should reflect the application of that doctrine, rather than the other way around.

In my view, cases like Antique Cultural Treasures indicate that there is a danger that the appellate courts may be over-generalising from these cases in the approach they are adopting. The approach does not appear to take properly into account the function of appellate courts to give rulings on issues of law which are authoritative across the legal system, and as a result the guidance function of senior courts in relation to matters of law has been undermined. In addition, the proposed standard approach seems to give insufficient weight to the significance of the relationship between the courts and Parliament, where the courts make judgments about the compatibility of legislation with human rights standards. Should it not be the judgment of the most senior judges which determines such questions, rather than the judgment of the most junior?

The rule of law perspective

Rule of law values inform the analysis in three ways. First, an important dimension of the rule of law is that the law should fulfil a guidance function, according to which citizens can predict how it will bear upon their activities. As Joseph Raz writes, “the law must be capable of guiding the
The law should provide a reasonably adequate guide by which people can orientate themselves, so far as circumstances allow.\(^{46}\) Secondly, it is also a function of rules to act as a transmission belt for authority, whether from legislature to courts or from higher courts to lower courts. The rule of law implies that there should be rule by bodies which have the authority to state the law. In a well-known essay,\(^{47}\) the US Supreme Court Justice Antonin Scalia wrote that he used to like legal tests stated in terms of standards which conferred discretion on judges, but that he came to prefer rules. Clear general rules promote predictability and provide a more effective check upon arbitrariness in judging.\(^{48}\) Excessive use of discretionary standards may be inappropriate in a legal system in which the Supreme Court can review only a tiny proportion of cases. It sacrifices the uniformity which the exercise of authority by the most senior court is supposed to achieve.\(^{49}\) Putting it another way, to the extent an appellate court treats a first instance judge as exercising discretion in the application of a standard and only conducts a light-touch review of what he or she has done, it dissipates its own authority in stating the law to be applied in the legal system.

This is not something unique to proportionality assessment. But there is something particularly open-ended about proportionality assessment, in that it involves bringing moral and consequential issues into account in what can be a relatively unstructured way at the fourth, \textit{stricto sensu}, stage.\(^{50}\) There is no necessary conclusion about their effect which everyone is simply

\(^{45}\) J. Raz, \textit{The Authority of Law} (1979), 214
\(^{48}\) Ibid., 1180.
\(^{49}\) Ibid., 1179-1179.
\(^{50}\) The appearance of structured determinacy in the framework for proportionality analysis is rather misleading, particularly at the fourth stage of the analysis in the application of proportionality \textit{stricto sensu}. See the discussion and references in Sales, “Proportionality and the Margin of Appreciation”, n 8; also, in relation to the application of proportionality analysis in the field of private law, E. Lim and F Urbina, “Understanding Proportionality in the Illegality Defence” (2020) 136 LQR 575.
bound, as a matter of rationality, to accept. Therefore, what is required in order for a legal system which applies proportionality analysis to operate in a satisfactorily coherent and determinate way is a system of authority in adjudication on the issues governed by such analysis.\(^\text{51}\)

Timothy Endicott writes, “it is a basic duty of courts to decide cases that the law does not resolve: i.e. to impose resolution … Resolution is a basic requirement of the rule of law.”\(^\text{52}\) A first instance judgment does supply resolution; but as Endicott says, the duty of resolution also needs to be married with Raz’s statement that the law should be capable of guiding its subjects.\(^\text{53}\) Further, “the need for resolution never overrides need for justice and the need for legality”.\(^\text{54}\) It is appellate courts which have authority to determine matters of justice and legality.

Thirdly, the rule of law implies reasonable access to the justice system to vindicate rights given by the law without undue delay. The justice system does not have unlimited resources. Further, in order to produce authority at appellate levels within the system, it has a steep pyramid structure. Authority is concentrated in a small cadre of judges at Court of Appeal level and an even smaller cadre at Supreme Court level. This means that access to appellate courts has to be rationed, since otherwise the system would seize up with hopeless delays in the resolution of disputes, thereby undermining the rule of law imperative of resolution mentioned by Endicott.

Appellate courts require protection against being overburdened in the exercise of their appellate function. A requirement of permission to appeal is one mechanism. Legal doctrine also has a role

---


\(^\text{52}\) T. Endicott, n. 46, 186 and 197-203.

\(^\text{53}\) Ibid., 203.

\(^\text{54}\) Ibid, 199.
to play. An appellate court does not have the same time to devote to a dispute as a first instance judge. It is therefore reasonable to think in terms of distributing responsibilities between them. For instance, in deciding on facts or in making highly discretionary decisions, such as when exercising case management functions, it may not add significantly to improvement in decision-making if an appellate court seeks to second-guess the judge, rather than simply checking they have acted reasonably. Further, any small improvement there might be may not be justified by the resources which the parties and the appellate court would have to devote to achieving it.55

However, in deciding when it is right to ration the attention of appellate courts in this way, the wider responsibilities of appellate courts to foster the other rule of law objectives I have mentioned also have to be brought into account. Excessive abdication by appellate courts by deferring to the judgment made by a court at first instance undermines those other objectives. The reason appellate courts have authority is to articulate general norms of law with binding force across many cases. This ensures that like cases are treated alike and helps to secure public confidence in the law. Also, by providing authoritative guidance about how the law should be construed and applied, it helps to secure goals of reasonable determinacy in law so individuals and undertakings can plan their affairs with knowledge of how the law will be applied and of the principles of effective co-ordination between different levels of the court system. This minimises incentives for speculative litigation in search of possible favourable outcomes in individual cases.

By contrast, if an appellate court applies a rule that a first instance court will be allowed to decide some question for itself and it is that court’s decision that, if it has directed itself correctly in general terms, will be respected, cases which appear (and which in fact are) closely analogous may be decided in different ways, in each case by a first instance judge forming their own view of

55 See e.g. Re B, para 89 (Lord Neuberger). One also sees a conception of a distribution of authority between courts, to reflect the institutional context, in the case-law of the ECtHR and CJEU.
the overall merits. The legal results will appear arbitrary and coordinated guidance within the legal system will be lost. Parties will have an incentive to litigate, in the hope of capturing the attention and sympathy of the first instance judge in their case. The benefits associated with having a system of *stare decisis* will be lost.

As Timothy Endicott writes in another essay,\(^{56}\)

> “*Stare decisis* is a good doctrine … because of the need for the rule of law. The doctrine imposes a rule-governed form of control on the wide discretions of judges in applying the common law and statute law. The judgment in a case solves a coordination problem: the problem of resolution in a particular case. And then *stare decisis* solves another coordination problem: how is the law generally to treat cases like this?”

The more that *stare decisis* is undermined through application of the proposed standard approach to appellate review in the area of proportionality assessment, leaving the operative assessment as the assessment made by the first instance judge in the particular case, the harder it will be to say that a previous appellate decision determines the outcome in the next case to come along.

Endicott identifies a second way in which judges make law through development of case law, by providing reasoning which commands general respect.\(^{57}\) He says, “Judicial reasoning binds future judges because judicial reasoning is treated as a good ground of law”; this too produces a coordinating effect, but this depends on whether judicial reasoning is generally sound, and that


\(^{57}\) Ibid, 132-133.
depends among other things “on the expertise and the wisdom of the judges”, with the presumption being that “judges in appellate courts are pre-eminently expert”.

But if the appellate court only conceives of its role as checking that the first instance court has directed itself in an appropriate way and then reached a result which cannot be said to be unreasonable, this second aspect of the way in which judges make law is also undermined. Society does not benefit from application of the wisdom of senior judges on some of the most difficult and sensitive cases which come before the courts. The reasoning of appellate judges is diverted from direct consideration of the merits of the cases before them; whereas part of the point of their promotion to appellate courts is to bring their wisdom to bear in a way which guides lower courts, not the other way round.

To tread between opposing risks to the rule of law, it seems we need to look for a more nuanced approach to appellate control in relation to proportionality review. How should the balance be struck?

_Analogies from other areas of domestic law_

The issues I have been discussing are not unique to the proportionality context. Courts have authority to rule upon the law, and in doing so frequently make evaluative judgments weighing up incommensurable factors. When is it appropriate for an appellate court to exercise its authority by making its own judgment and assessment? When is it appropriate to conduct a diminished checking review?
Points of statutory interpretation are usually regarded as a paradigm form of a question of law. The fact that it is a paradigm demonstrates an important truth: judicial review courts and appellate courts have a special constitutional role in producing authoritative general rulings on the meaning of law applicable across a wide range of cases.

This role extends to judgments regarding rationality in decision-making in public law. The rationality standard operates as a baseline of what is acceptable behaviour for a public authority. It is fundamental, so even though it may be a standard that varies somewhat according to context58 the statement of what is and what is not regarded as rational (and hence lawful) for a public authority to do is treated as a question of law and a proper issue for determination by an appellate court.

At the other end of the spectrum lie cases involving the exercise of discretions deliberately vested in first instance courts. Case management decisions and decisions whether to grant interlocutory relief pending trial are prime examples.59 Endicott argues that where the legislature deliberately enacts a very vague standard, it may mean that it intended to leave the decision to the courts and that the author of the rule did not intend to communicate any particular solution in the application of that rule; then, the role of the courts is more like invention than interpretation.60

At a mid-point on the spectrum between pure points of law and pure points of discretion lie cases in which the application of open-textured statutory terms or legal rules calls for a

59 See e.g. Hadmor Productions Ltd v Hamilton [1983] AC 191.
60 T. Endicott, Vagueness in Law, n. 46, 181-182.
multifactorial assessment. This is where Lord Clarke in Re B sought to locate appellate review of proportionality assessments. He did so by reference to wider principles of appellate review. As he said:

“… the correct approach of an appellate court in a particular case may depend upon all the circumstances of that case. So, for example, it has traditionally been held that, absent an error of principle, the Court of Appeal will not interfere with the exercise of a discretion unless the judge was plainly wrong. On the other hand, where the process involves a consideration of a number of different factors, all will depend on the circumstances. As Hoffmann LJ put it in In re Grayon Building Services Ltd [1995] Ch 241 at 254,

‘generally speaking, the vaguer the standard and the greater the number of factors which the court has to weigh up in deciding whether or not the standards have been met, the more reluctant an appellate court will be to interfere with the trial judge's decision.’”

61

A functional approach

Different factors point in different directions in terms of assigning responsibility for a proportionality evaluation of a measure between a first instance and an appellate court. The factors, and their relative weight, will vary from case to case. Accordingly, the arguments for assigning greater or lesser degrees of responsibility to a first instance as against an appellate court, and vice versa, will vary. And I suggest that this variation should be reflected in doctrine.

Sometimes a proportionality assessment may be closely bound up with findings of fact. In our system, finding the facts of a case is usually regarded as the responsibility of the first instance

61 In the Matter of B (a Child) [2013] UKSC 33, [2013] 1 WLR 4079, para 137.
court. Even when the evidence is in written form, a first instance court will often have had a better opportunity to understand it in depth.\(^{62}\) Where a proportionality assessment turns purely on the application of an uncontroversial general rule of law to the particular facts of a case, it may be appropriate to adopt the proposed standard appellate approach of review of that assessment by the first instance court.

The courts identify cases which involve questions of mixed fact and law. These may be analysed, like proportionality review, as involving the application of an open textured statutory term to the facts of a particular case. The cases often arise where there is an appeal on a point of law from a decision at first instance or in judicial review. Where is the boundary to be drawn between finding facts and questions of law?

*Edwards v Bairstow*\(^{63}\) concerned whether certain commercial activity constituted an adventure in the nature of trade which would be liable to tax. As Lord Radcliffe observed, in answering such a question there would be a wide area where the first instance tribunal could decide a case either way and it “could not be said to be wrong to arrive at a conclusion one way or the other”.\(^{64}\) Absent an express self-misdirection by the tribunal regarding the law, the appeal court would find an error of law only where the tribunal had arrived at a conclusion different from the “true and only reasonable conclusion” of law on the facts as found.\(^{65}\) That is to say, according to the reformulation by Lord Diplock in the *CCSU* case, an error of law will be found only if the conclusion drawn by the tribunal is irrational.\(^{66}\)

---


\(^{63}\) [1956] AC 14.

\(^{64}\) [1956] AC 14, 33.

\(^{65}\) [1956] AC 14, 36.

\(^{66}\) *Council of Civil Service Unions v Minister for the Civil Service* [1985] AC 374, 410-411 (Lord Diplock).
In drawing the boundary, John Laws argued that an appellate court should be willing to adopt a functional approach. The dividing line between law and fact should be driven by the functional responsibilities and capacities of the bodies in question. As he says, “the boundary between the very ideas of law and fact is not fixed; it largely depends on the answer to a question which transcends the internal rules of any legal system, namely what are the characteristics of a decision which ought to expose it to control by the courts?”

This suggests a way forward for the appellate approach to proportionality review. An appellate court should adopt a primary decision-making function where the characteristics of a decision by a first instance court are such that it ought to be exposed to full control (as opposed only to review control) by the appellate court. The appellate court should adopt a primary decision-making function when it is able to add value to the normative exercise in deciding whether a measure can be regarded as proportionate, where that potential for added value sufficiently reflects the additional costs and delay associated with an appeal.

Generally, that is unlikely to be the case in relation to reviewing facts found by the first instance court. But the appellate court has a constitutional function to articulate and police general legal norms. Thus, there may be a spectrum of potential engagement by an appellate court, depending on the precise nature of the issue which arises in relation to a proportionality assessment. On this approach, there will be differences of degree, regarding how far the appellate court should be drawn into acting as primary decision-maker to make the assessment afresh for itself. Depending on the circumstances in a particular case, it may be possible for the appellate court to

---

accept findings of fact made at first instance (subject only to rationality review) and then supply its own view of the values in contest in that factual position and of the normative outcome.68

Experience shows that this is a viable way of approaching proportionality review. The German Constitutional Court approaches proportionality analysis by allowing for a functional division of responsibilities: since it is a specialised constitutional court to which constitutional questions arising from proceedings in the ordinary courts are referred, it relies to a significant extent on the factual findings of those courts.69 But it will also exercise a primary fact-finding function where necessary.

In our law, perhaps a distinction might be drawn between finding facts which are of general importance for good law-making in a particular field, such as the way in which a measure advances a legitimate public purpose or the ways in which it affects large classes of persons,70 and the finding of particular facts specific to a claimant’s case. In US and Canadian law, such a distinction is drawn between so-called legislative facts and adjudicative facts.71

There is always an element of judgment in applying a legal standard to a set of facts. There is no hard and fast divide between a statement of the law and its application.72 The application of a rule of law is one way of showing its meaning. So, the issue is really an institutional one, of the

68 Cf Austin v UK (2012) 55 EHRR 14, para. 61: “..., since pursuant to arts 19 and 32 of the Convention it is the Court’s role definitively to interpret and apply the Convention, while it must have reference to the domestic court’s findings of fact, it is not constrained by their legal conclusions as to whether or not there has been a deprivation of liberty within the meaning of art.5(1)”.
71 I am grateful to Timothy Endicott for drawing this to my attention.
72 This is the way the common law operates. The Strasbourg and domestic caselaw on the application of Convention rights is a form of common law of human rights, which supplies determinate content for the open-textured Convention rights: see P. Sales, “Law and Democracy in a Human Rights Framework”, ch. 15 in D. Feldman (ed), Law in Politics, Politics in Law (2013), 219-220 and 231-233.
extent to which the appellate court should be exercising control over the meaning of law, including in its application as well as by way of statements of principle at a high level of generality.

The issue is similar to that which arises when an appellate court has to consider making a statement of a common law rule or principle which will cover cases beyond the case at hand as a matter of precedent. The more the court makes a statement of wide application, the more cases fall within its grasp and the more closely the action of the court has quasi-legislative effect - and the more important it is that the court’s own authority is brought to bear in determining the matter.

In similar fashion, a measure which is subject to proportionality assessment may be of a kind which has implications for the future for wide classes of case or, on the other hand, may be much narrower in its effects. The willingness of an appellate court to exercise direct control over the proportionality assessment of the measure ought to reflect such considerations.

The constitutional function of appellate courts implies that there is a level of generality at which it might be necessary for an appellate court to assess proportionality itself in order to articulate, or check the articulation, of a general rule for society. Also, the exercise of a primary decision-making function by an appellate court may be important under circumstances of value pluralism and incommensurability to demarcate where the boundaries of what will be taken to be proportionate are set, where such determination may have wide ramifications.

Litigation under the HRA, where a declaration of incompatibility is sought in relation to primary legislation, engages the responsibility of the courts in another way. Where the declaration is
sought in relation to the effect of the legislation on a wide class of case,\textsuperscript{73} the courts are asked to exercise a significant review function in relation to what Parliament has done and the argument for the appellate courts should act as primary decision-makers is a strong one.

First, cases involving assessment of the compatibility of legislation with Convention rights are in fact about the application of a legal test and the issues which arise are really just a sub-set of the issues which arise in relation to a decision involving mixed fact and law. If incompatibility with Convention rights is demonstrated, typically by reference to a proportionality analysis, then either section 3 of the HRA authorises the court to modify the interpretation of the legislation or, if that is not possible, section 4 authorises the court to issue a declaration of incompatibility.

Secondly, review of legislation for its compatibility with human rights is a major constitutional function for the courts, in checking that a proper balance is maintained between two constitutional values – rule of law and democracy.\textsuperscript{74} Where a general legislative scheme is in issue, respect for the democratic process implies that the body with authority to rule upon this should be the appellate court. In particular, ultimate authority to rule upon this should lie with the apex appeal court – the Supreme Court. Parliament, the Government and citizens are all entitled to expect that the judges of the highest appellate court, with the greatest decision-making authority within the court system, should apply their own minds and judgment in deciding whether a statute on its ordinary reading is incompatible with Convention rights and so liable to have a declaration of incompatibility made in relation to it. A declaration of incompatibility is a major public act with significant consequences for public respect for Parliament and its

\textsuperscript{73} As distinct from in relation to a one-off case, or small group of cases, where incompatibility might be found to arise because it is not “possible” to construe the legislation to accommodate the Convention rights in those few cases: see n. 25.

legislation. It may also have a significant effect in relation to the assessment by the Strasbourg Court regarding whether the United Kingdom has met its obligations under the ECHR.75 Conversely, when the domestic courts carry out their checking function and uphold the compatibility of legislation with Convention rights, it is the judgment of senior appellate courts which is likely to carry the most weight in providing reassurance about that to minorities, the general public and the Strasbourg Court.76

Parliament and the public are entitled to the direct application of what one might dare to call the wisdom of judges in superior courts, acquired through long practice in the law and as judges.77 In a sense, this is what the domain of legitimation associated with the rule of law owes to the domain of legitimation associated with the democratic process: that its most experienced and most authoritative judges should directly exercise their judgment in ruling on the compatibility with human rights of what Parliament has done.

Conclusion

A dialogue should take place between the two lines of authority which have developed regarding the approach of the appellate courts to proportionality review. The varied circumstances in which proportionality analysis may be called for mean that it is difficult to determine the proper approach for an appellate court by means of a definitive bright line rule. The approach to be adopted should be responsive to a range of factors which vary from case to case. It follows that the question of the approach to be adopted should be matter of legal argument, and counsel should be prepared to raise the issue in order to seek reasoned determinations from appellate

76 Sales, “Partnership and Challenge”, n. 74, 468.
courts. The courts should gradually seek to clarify the boundary, case by case. This represents the best hope for arriving at as much clarity as possible.