What would it take? The potential and limits of proportionality analysis in law

Ralf Poscher

Submit your article to this journal ↗

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

What would it take? The potential and limits of proportionality analysis in law

Ralf Poscher

Department of Public Law, Max Planck Institute for the Study of Crime, Security and Law Freiburg, Germany

ABSTRACT

Proportionality has become one of the primary standards for violations of fundamental and human rights in various constitutional and international jurisdictions. Proportionality in the strict sense requires a comparison of the infringement of a right with the achievement of a state interest pursued by a measure. The incommensurability of the two is an often-raised challenge to this comparison. The Article does not adopt a fully-fledged sceptical stance. Instead, it looks to proportionality as a means of overcoming the challenge of incommensurability. It then examines the structure of comparisons in general and of comparisons of incommensurables employing proportionality in particular. As for the latter it focuses on its performance by exploring the example of decathlon, where this type of comparison works exceptionally well. The example allows us to analyse the preconditions for comparing incommensurables. Unfortunately, most of these preconditions are lacking in the law. This explains why the proportionality principle in law only produces convincing results in cases where the disproportionality is very pronounced. In many other cases, however, the courts decide based on proportionality as well. This raises the question: What is really going on in these cases? The article concludes with a brief account of a benevolent explanation.

In many jurisdictions at the national and international level, proportionality has become a central – if not *the* central – standard of judicial review in cases involving constitutional and human rights.[1] It has friends and foes, and there is a lively debate on whether it strengthens or weakens protections for individual rights.[2] This article will explore a theoretical precondition of this normative debate, namely the potential of the

---

CONTACT Ralf Poscher ✉ public-law@csl.mpg.de 📧 Department of Public Law, Max Planck Institute for the Study of Crime, Security and Law, Günterstalstrasse, 73, Freiburg D- 79100, Germany

[1]Aharon Barak, *Proportionality: Constitutional Rights and their Limitations* (Cambridge studies in constitutional law, CUP 2012), with a diagram on 182.

[2]Just see the debate between Stavros Tsakyrakis, 'Proportionality: An Assault on Human Rights?' (2009) 7 Int J Const L 468; Stavros Tsakyrakis, 'Proportionality: An Assault on Human Rights?: A Rejoinder to Madhav Khosla' (2010) 8 Int J Const L 307, and Madhav Khosla, 'Proportionality: An Assault on Human Rights? A Reply' (2010) 8 Int J Const L 298, in the field of human rights. See also David M Beatty, *The Ultimate Rule of Law* (reprint OUP 2010).

proportionality standard in law. An analysis of the proportionality standard will reveal that whilst it can serve to outlaw instrumentally irrational and the most egregiously disproportionate state actions, it fails – for theoretical reasons – to explain many decisions which purport to rely on its proportionality in the strict sense. The latter involves the balancing of different interests, values, and rights in individual cases, and it is the incommensurability of these components to which critiques often point.

This article begins, in (I.), by examining this challenge to the rationality of judgements on proportionality in the strict sense. In so doing, it mainly pursues theoretical interests: it aims to examine why, on the one hand, it is not justified to view judgments based on proportionality in the strict sense as sceptically as many have done, while, on the other, it is justified to consider the performance of this procedure to be very low, given that it is only suitable for cases with strong discrepancies between the severity of rights infringements and the achievement of state interests. The aim is not to prove the obvious, but to understand in detail what kinds of elements come into play when we compare incommensurable entities, and which of these elements are not in fact served by proportionality in the strict sense in law. To this end, in laying the theoretical foundation, the analysis starts out by recapitulating the basic structure and types of comparisons in (II.). It first presents the elements and the structure of comparison to better describe the features of incomparability and incommensurability. (III.) initially describes the structure of incommensurable comparisons followed by an example demonstrating that comparisons of incommensurables can be extremely accurate and precise. The example is taken from a sport made up of different disciplines: decathlon. This serves to highlight what is needed to make comparisons of incommensurables accurate and precise, thus, laying the groundwork for assessing the performance of proportionality judgments in law in (IV.). Decathlon as an example is used to show in detail which preconditions are not met when it comes to the comparison of fundamental rights and state interests required by proportionality in the strict sense. Further, the detailed reconstruction of the challenges for proportionality judgments in law will explain why they enable judgments in some cases despite meeting the preconditions of comparing incommensurables only to a very limited extent, and why proportionality judgments in certain areas of the law perform above average. However, having established proportionality's usefulness in a limited number of cases, the article will then look at the fact that courts massively inflate the test for proportionality in the strict sense in many rulings which do not actually meet the relevant structural conditions. This raises the question of what actually goes on under the guise of proportionality in court decisions. Accordingly, in part (V.), the article will sketch out an answer to this question, which has been outlined in more detail in previous publications.

I. Elements of the proportionality standard

Proportionality is cut and sliced in different ways in different jurisdictions – sometimes even for the same reason: The German Federal Constitutional Court, for instance, focuses on the proportionality in the strict sense. It does so even in cases that could be decided by the necessity test[3] because it does not want to reproach the legislature for

---

[3]For example, BVerfGE 141, 220 (265–268) (english translation: <https://www.bverfg.de/e/rs20160420_1bvr096609en. html> accessed 19 September 2024, paras 92–102).

acting 'irrationally' by infringing fundamental rights more than necessary to achieve its purpose. The Canadian Supreme Court, on the other hand, tends to present balancing considerations in terms of the necessity test[4] as it does not want to appear to encroach on the legislature's prerogative of accommodating different political interests.[5] Both courts ultimately seek to avoid a confrontation with the legislature. They, however, choose different ways. Whereas the German court has a tendency to present necessity considerations in the context of proportionality in the strict sense, the Canadian Court has a tendency to present considerations of proportionality in the strict sense in the context of necessity.

This article will not address such manoeuvres, however interesting they might be from a comparative perspective. Instead, it will take a theoretical approach.

From a theoretical perspective, proportionality analysis has five elements:[6] legitimate aim, legitimate means, suitability of the means in relation to the aim, necessity of the means as the least intrusive measure to achieve the aim, and proportionality *stricto sensu*, i.e., the degree of the intrusion into a fundamental right compared to the degree of achievement of the legitimate aim, which is also sometimes addressed as the final or balancing stage of proportionality.

Suitability and necessity require a causal relation between means and ends. Does a certain government measure help to accomplish its intended aim? Are there less intrusive measures that cause at least as positive an effect as the measure under scrutiny? In substance, suitability and necessity amount to an instrumental rationality test: It is irrational to employ a means that is not suitable to further a self-imposed aim, and it is also irrational to cause more harm than necessary.[7] Suitability and necessity are of forensic importance. Both the administrative bureaucracy and the legislature can get things wrong or overlook less harmful means to pursue their aims. However, even though both elements of the proportionality standard can pose intricate empirical issues, they raise no deep theoretical questions.

Proportionality in the strict sense is quite distinct from suitability and necessity. First, it does not concern a causal but an evaluative relation. The degree of the infringement of fundamental rights must be compared to the degree to which the measure furthers a legitimate government objective. Comparing the degrees of infringement does not rely on an empirical causal relationship between the two but rather evaluates them according to some normative standard, namely whether they are proportionate – that is to say

---

[4]For example, *Corbiere v Canada (Minister of Indian and Northern Affairs)* [1999] 2 SCR 203; *R v McKinlay Transport Ltd* [1990] 1 SCR 627.

[5]Niels Petersen, *Proportionality and judicial activism: Fundamental rights adjudication in Canada, Germany and South Africa* (CUP 2017) 123 ff; Niels Petersen, 'Alexy and the "German" Model of Proportionality: Why the Theory of Constitutional Rights Does Not Provide a Representative Reconstruction of the Proportionality Test' (2020) 21(2) German LJ 163 <https://www.cambridge.org/core/journals/german-law-journal/article/alexy-and-the-german-model-of-proportionality-why-the-theory-of-constitutional-rights-does-not-provide-a-representative-reconstruction-of-the-proportionality-test/2EB57D7431F604A4FE37663F41F99206> accessed 19 September 2024.

[6]For different ways to subdivide proportionality Dieter Grimm, 'Proportionality in Canadian and German Constitutional Jurisprudence' (2007) 57 U of Toronto LJ 393; Giovanni Sartor, 'The Logic of Proportionality: Reasoning with Non-Numerical Magnitudes' (2013) 14(8) German LJ 1419, 1447 <https://www.cambridge.org/core/journals/german-law-journal/article/logic-of-proportionality-reasoning-with-nonnumerical-magnitudes/FB1F106A1A817D3073DAF6849A9E2DB6> accessed 19 September 2024; George Letsas, 'The Scope and Balancing of Rights: Diagnostic or Constitutive?' in Eva Brems and Janneke Gerards (eds), *Shaping Rights in the European Convention on Human Rights* (CUP 2014) 42; Francisco J. Urbina, *A Critique of Proportionality and Balancing* (CUP 2017) 5–6.

[7]A mean, however, does not violate necessity if the alternatives are more costly even if they are less intrusive.

whether the degree of infringement exceeds or corresponds to the degree to which the measure furthers the legitimate aim. An example of such normative judgments of proportionality in the strict sense are judgments about the culpability of an offender. Although there is a causal relationship between the offender's action and the harm she causes, the relationship between the harm and her guilt is not causal but normative. It is only causal in the Davidsonian sense that reasons can be causes.[8] Our normative reasons for condemning the offender's action can be a cause for attributing blame. Unlike suitability and necessity, the basic relationship between the offender's action and her guilt is not causal but normative.

Second, since proportionality in the strict sense is not based on a causal relation, it does not require one. It can thus also apply to judgments which do not rely on a causal relationship between means and ends. In German law, for example, administrative fees must be proportionate to the value of the service a government agency provides. However, the fee does not *cause* the service or vice versa. Rather, proportionality requires that the amount of the fee be normatively evaluated against the value of the service it is supposed to finance. Another example is the above one from criminal law, where the relation between the punishment and the offender's culpability must be proportionate. The guilt or severity of the crime, however, does not *cause* the punishment; instead, it is the standard by which the amount of punishment is evaluated.

In the context of criminal sentencing, this feature specific to proportionality in the strict sense is sometimes described as retrospective proportionality.[9] The retrospective nature of assessing criminal sentences accounts for and implies the non-causal nature of the underlying relation since there is no retroactive causality. This characterisation is nevertheless misleading because the special characteristic of proportionality in cases such as criminal sentencing is not a question of direction in time but of a lack of any causal relation between the two objects under comparison (i.e., guilt and sentence).[10] The causal elements of proportionality in the wider sense do not apply here. Thus, it would be nonsensical to ask whether an offender's guilt was 'suitable' to cause the penalty or vice versa, or whether there would be a penalty that infringes less upon the offender's culpability. The culpability of the defendant does not empirically cause the punishment or vice versa. It is only through the normative principle that a punishment has to be proportionate in the strict sense to the culpability of the defendant that the two are connected. Other normative proportionality standards not based on a causal end-means relationship can be found in tax law – for example, when the law requires that the tax burden be proportionate to the taxpayer's ability to pay,[11] or in German labour law, when it is required that the means chosen by unions and employer representatives in a labour dispute be proportionate to one another.[12]

Depending on the perspective, proportionality in the wider sense can be narrower in scope than proportionality in the strict sense, but also wider. Specifically, proportionality

---

[8]Donald Davidson, 'Actions, Reasons, and Causes' in Donald Davidson (ed), *Essays on Actions and Events* (Clarendon Press 2001) 3–19.
[9]Antony Duff, 'Proportionality and the Criminal Law: Proportionality of What to What?' in Emmanouil Billis, Nandor Knust and Jon P Rui (eds), *Proportionality in Crime Control and Criminal Justice* (Hart Publishing 2021) 30–34.
[10]On this important precondition of the suitability and necessity element of proportionality, Stefan Huster, *Rechte und Ziele: Zur Dogmatik des allgemeinen Gleichheitssatzes* (Duncker und Humblot 1993) 131–39.
[11]BVerfGE 82, 60 (90); BVerfG NJW 1990, 2869 (2872).
[12]BVerfG NZA 1991, 809 (811 f.); BAG NJW 1971, 1668 (1669).

in the wider sense is narrower since it requires an underlying causal relation between means and ends. It is wider only in comprehensiveness, in that it includes the suitability and necessity standards next to proportionality in the strict sense.

Third, unlike the other elements of the proportionality standard, proportionality in the strict sense poses questions regarding its theoretical viability when applied to the law, where it requires the comparison of seemingly incommensurable items: Is it even theoretically possible to apply proportionality in the strict sense under these circumstances? This shall be explored in the next section.

Incommensurability as the key issue of proportionality?

Scepticism towards proportionality in the strict sense in law is grounded in the incommensurability of the interests, values, and rights that must be balanced to arrive at an proportionality in the strict sense judgment.[13] Weighing security against liberty interests may seem like comparing apples and oranges, or as *Antonin Scalia* put it, like asking 'whether a particular line is longer than a particular rock is heavy'.[14] There seems to be no common scale to measure interests, values, and rights to be balanced as required by proportionality in the strict sense. Unlike empirical suitability and necessity judgments, the comparative judgment required for proportionality in the strict sense appears to lack an objective standard; it conveys the impression of being solely a question of subjective preferences.

Despite these theoretical challenges, the proportionality in the strict sense plays a significant role in the proportionality review of many national and international courts. Analysis of proportionality in the strict sense are also front and centre in the fundamental-rights jurisprudence of the German Federal Constitutional Court. Some of its more recent decisions focus exclusively on proportionality in the strict sense, and the court presents all of its – sometimes very detailed – constitutional specifications as being dictated by proportionality in the strict sense.[15] As already noted, other courts are less candid. At least sometimes, they conceal balancing considerations within other elements of their proportionality analysis. In *Nur* for example, the Canadian Supreme Court applied the minimal impairment test but therein held that the provision in question was 'grossly disproportionate'.[16] Even though it can at times appear unclear which elements of the proportionality test the European Court of Justice employs, in substance proportionality in the strict sense is an integral part of its doctrines.[17] Equally, the

---

[13]Bernhard Schlink, *Abwägung im Verfassungsrecht* (Schriften zum öffentlichen Recht vol 299, Duncker und Humblot 1976) 154–91, 210–14; Alexander T Aleinikoff, 'Constitutional Law in the Age of Balancing' (1987) 96(5) Yale LJ 943, 972–83; Jürgen Habermas, *Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats* (Suhrkamp 1992) 315 f; Tsakyrakis, 'Proportionality: An Assault on Human Rights?' (n 2) 471; Grégoire Webber, 'Proportionality, Balancing, and the Cult of Constitutional Rights Scholarship' (2010) 23 Can JL & Jurisprudence 179, 194.

[14]*Bendix Autolite v Midwesco Enterprises, Inc* 486 US 888, 897 (1988).

[15]For two extreme cases, see BVerfGE 141, 220 (267–286) (english translation: <https://www.bverfg.de/e/rs20160420_1bvr096609en.html> accessed 19 September 2024, para 98–144); BVerfG NJW 2022, 1583 (1584) (english translation: <https://www.bverfg.de/e/rs20220426_1bvr161917en.html> accessed 19 September 2024, para 147 f).

[16]*R v Nur* [2015] 1 S.C.R. 773; *Bennett v Canada (AG)* [2011] FC 1310; cf dissenting opinion in *Alberta v Hutterian Brethren of Wilson Colony* [2009] 2 SCR 567, para 134 (Abella J) for further references, see also Petersen, *Proportionality and judicial activism* (n 5) 105.

[17]See for example ECJ, Case C-293/12 and C-594/12 *Digital Rights Ireland* EU:C:2014:238, paras 66–72; Case C-203/15 and C-698/15 *Tele2 Sverige/Watson* EU:2016:970, paras 122 f, in which the court formulated a series of procedural safeguards

European Court of Human Rights even if not always in name, but in substance relies on balancing for its proportionality evaluations.[18]

It is thus not surprising that attempts have been made to overcome the incommensurability of interests, values, and rights. Unfortunately, all attempts to date to develop a common standard for items such as 'fundamental rights' and 'state interests' have failed.

They are either unconvincing in themselves or impracticable even when all methodological objections are put to one side. There are two basic strategies to arrive at a common standard. One strategy relies on creating an overarching social good that encompasses fundamental rights and state interests. *Aharon Barak*, for example, has proposed using 'social significance' as a uniform evaluation parameter.[19] Yet even if we were to become able to compare different kinds of legal positions after converting them into the currency of social significance, *Timothy Endicott*'s point remains cogent: This conversion requires us to eliminate the very incommensurabilities that the uniform 'standard' of social significance is intended to overcome.[20] The second strategy relies on the aggregation of individual empirical preferences. An econometric solution might be to assess incommensurability empirically by asking rights-bearers what they would be willing to pay for the goods that are to be compared. However, this would raise considerable methodological difficulties,[21] and even if they could be overcome in an experimental setting, the solution would not be workable in constitutional law terms. Moreover, it is questionable whether empirical preferences determined by such experiments would yield any results that could help define normative relationships between the state's policy aims and infringements of fundamental rights – just as empirical morality does not allow us to infer normative moral judgements. The incommensurability of the interests to be balanced and weighed apparently remains the central and irreconcilable problem of the proportionality principle.

Proportionality as the solution to incommensurability?

Yet changing perspective reverses the picture. According to this view, rather than incommensurability being the key issue of proportionality, the latter is in fact the solution to the incommensurability problem.[22] Precisely *because* we cannot compare two legal interests with the same standard, we can only compare the degrees of proportional fulfilment using different standards applicable to the respective legal interests at issue. By

---

for data retention regimes on the basis of 'strict necessity'. In Case C-511/18, C-512/18 and 520/18 *Quadrature du Net* EU:C:2020:791, paras 129–168, however, it demanded many of the same safeguards based on proportionality in the strict sense, para 130: 'the protection of personal data must apply only in so far as is strictly necessary. In addition, an objective of general interest may not be pursued without having regard to the fact that it must be reconciled with the fundamental rights affected by the measure, by properly balancing the objective of general interest against the rights at issue'.

[18]See for example *Guimon v France* App no 48798/14 (ECtHR, 11 April 2019), para 52; *Bărbulescu v Romania* App no 61496/08 (ECtHR, 5 September 2017), para 141; *Schüth v Germany* App no 620/03 (ECtHR, 23 September 2010), paras 74 f; see also Aharon Barak, *Proportionality: Constitutional rights and their limitations* (Doron Kalir tr Cambridge studies in constitutional law, reprint CUP 2012) 348 ff.

[19]ibid.

[20]Timothy Endicott, 'Proportionality and Incommensurability' in Grant Huscroft, Bradley W Miller and Grégoire Webber (eds), *Proportionality and the Rule of Law: Rights, Justification, Reasoning* (CUP 2016) 318.

[21]On these, see Petersen, *Proportionality and judicial activism* (n 5) 41 ff.

[22]Bruce Chapman, 'Law, Incommensurability, and Conceptually Sequenced Argument' (1998) 146 U of Pennsylvania L Rev 1487.

comparing to what extent legal interests have been realised under their own respective standards, we can weigh and balance that which would otherwise be incommensurable.[23] *Bruce Chapman* has illustrated the idea with the example of a dog show.[24] Poodles and German shepards cannot be compared to each other. However, poodles and German shepards can each be assessed based on how well they represent their respective breeds. The scores thus determined decide which dog is Best in Show. This method can be used even if some dogs (such as poodles) are judged only on their appearance, while others (such as German shepherds) are also judged based on their herding and protective skills.

In fact, seemingly 'incommensurable' comparisons are a routine part of everyday life. We compare a vacation in the mountains with one by the sea; an Italian restaurant with a Japanese one. If absolutely necessary, one could also form an opinion as to whether *Beethoven*, for example, was a greater artist than, say, *Jeff Koons*. It seems evident that any such judgment, if it is not just guided by hedonistic spontaneity, is based on similar comparisons to those made at dog shows. In everyday comparisons, we also judge to what extent seemingly incommensurable things or phenomena approach an – admittedly implicit – standard of excellence within their *respective genres*.

More abstractly, these comparisons can be described as comparisons of *proportional realisation* measured by particular reference values. We measure the degree to which the poodle has come close to the ideal of its breed and compare this degree to the degree the German shepherd has come close to the ideal of its breed; the degree to which any particular Italian restaurant corresponds to a Michelin-starred Italian restaurant and compare this degree to the degree any particular Japanese restaurant corresponds to a Michelin-starred Japanese restaurant, etc. Thinking in degrees of realisation, which rely on the proportionate fulfilment of some reference values, allows us to compare the otherwise incommensurable alternatives with which life routinely confronts us.

Sports shows us that this type of comparison can even yield precise and objective results. Performances in the individual disciplines of a decathlon, each of which is incommensurable with the others, are measured precisely.[25] Thus, in judging decathlons, we see something happening that Justice *Scalia* had deemed impossible – physically different categories can be compared: for example, the time of a run with the distance of a jump.[26] Which factors enable such numerically exact comparisons of otherwise incommensurable performances in the decathlon? How do the conditions for comparing degrees of realisation in law differ from those in sports, and what does this imply for judging the effectiveness of this approach? We are under no illusion that a precision

---

[23]A formalised attempt by Paul-Erik N Veel, 'Incommensurability, Proportionality, and Rational Legal Decision-Making' (2010) 4(2) LEHR 177, 195 ff; see Petersen, *Proportionality and judicial activism* (n 5) 47 ff; cf also Christoph Engel, 'Besonderes Verwaltungsrecht und ökonomische Theorie' (2011) Preprints of the Max Planck Institute for Research on Collective Goods 2/2011, 16–18 <https://hdl.handle.net/11858/00-001M-0000-0028-6DB0-7> accessed 20 September 2024, who proposes comparing the benefits and costs (in terms of freedom) of a state action to the maximum potential benefit and costs, and approving the action only if the benefits are preponderant.

[24]Chapman (n 22), 1492.

[25]For the original instance of this example, Alfred F MacKay, *Arrow's Theorem: The Paradox of Social Choice: A Case Study in the Philosophy of Economics* (Yale UP 1980) 71 ff; Marlies Ahlert and Hartmut Kliemt, 'Unverrechenbare Werte, verrechnende Prioritäten' in Björn Schmitz-Luhn and André Bohmeier (eds), *Priorisierung in der Medizin: Kriterien im Dialog* (Kölner Schriften zum Medizinrecht vol 11 Springer 2013) 240; cf also Weyma Lübbe, *Nonaggregationismus: Grundlagen der Allokationsethik* (ethica Band 29 mentis 2015) 28.

[26]Niels Petersen, 'How to Compare the Length of Lines to the Weight of Stones: Balancing and the Resolution of Value Conflicts in Constitutional Law' (2013) 14(8) German LJ 1387.

comparable to that in decathlon is achievable when it comes to law. But how exactly are degrees of realisation compared; what are the required elements and procedures; and where and how does the law fall behind? The in-depth analysis of comparisons of degrees of realisation in this article does not aim to state the obvious but to understand why exactly the law generally performs poorly when it requires comparing incommensurables. In addition, the following paragraphs will shine a spotlight on why comparisons can nevertheless deliver results in cases where the differences are more pronounced, and why some areas of the law lend themselves more to proportionality judgments than others. It will also discuss the seemingly innocuous attempts at formalising balancing in constitutional law, as suggested by *Robert Alexy* and his followers.

II. How do comparisons work?

Proportionality in the strict sense is based on comparing the degrees of realisation of the purpose of some state action, on the one hand, with the degree of realisation of fundamental rights, on the other. It follows that the comparison of the degrees of realisation should reflect the basic structure of any comparison.

Structure and types of comparisons

Comparisons always presuppose not merely two objects as *comparata*, but also a property which serves as the basis of comparison.[27] The earth and the moon can be compared both in terms of a property, such as their mass, and in terms of a relation, such as their distance from the sun. The *tertium comparationis* is not an additional object, but rather a property or relation. A comparison always consists of a three-valued relation. It presupposes at least two objects and a property or relation that forms the standard of comparison. Thus, comparisons between two objects can be distinctive not only with regard to the objects but also with regard to the property or relation against which the comparison is made.

The relevant property for the comparison is determined by the purpose of the comparison. Thus, sports that are amenable to the same standards of measurement may be compared by consulting different characteristics, depending on the purpose of the comparison. Archery and javelin tosses, for example, could theoretically be compared by measuring the distance the projectile travelled. However, archery is not judged by distance, but rather by proximity to a given target. The two disciplines exist to fulfil different purposes. One tests the athletes' strength and technique, the other tests their technique and precision. Like the property guiding the comparison, the purpose of the comparison also determines the polarity of the property to which the comparison refers. When camera tripods are compared by weight, the comparison can be based on either portability or stability. Depending on which criterion is chosen, heavier tripods may be preferred to lighter ones, or vice versa. The polarity can refer not only to

---

[27]Ralph Weber, 'Comparative Philosophy and the Tertium: Comparing What with What, and in What Respect?' (2014) 13 (2) Dao 151, 152, speaks somewhat vaguely of 'aspects' as the *tertium comparationis*. However, the aspects can be more precisely defined as properties or relations. This does not imply a position in the problem of universals; the structure of comparisons is not changed depending on whether one takes a realist or nominalist understanding of properties and relations.

higher or lower but also to equal values, or any other relation between the *comparata* with respect to the relevant property. Thus, someone might compare different saddlebags for horse riding to find one with a weight that matches the one he or she already owns to avoid imbalance.

One way to distinguish comparisons from mere categorisations might be to invoke the (respective) purposes, which allow for evaluating the results of a comparison. In contrast to mere categorisation, comparisons serve to evaluate a certain result based on a property that displays a certain polarity. They differ from mere categorisation in that comparisons rely on a *five-valued* relation between two objects, a property, and an evaluative purpose which provides a reason for the choice of property and determines its polarity. Even though some values remain implicit in many comparisons, it generally takes five elements to give a full account of a comparison: two objects, an evaluative purpose, a property, and the polarity of the comparison.

Different types of properties and relations enable different types of comparisons. This is also worth noting because it shows that the concept of a property allows for reflexivity. Properties can themselves have properties – a property of properties, which will prove to be important for our handling of incommensurability. Nominal or qualitative properties and relations allow only nominal comparisons. Two lemons, for example, can be compared to see if they belong to the *Primofiori* variety. Comparison on the basis of a nominal property permits an evaluation of objects when a polarity is assigned to the property of comparison, e.g., when only *Primofiori* lemons are assigned high value. However, nominal comparisons do not provide a basis for ordinal or cardinal, i.e., metric rankings. Instead, ordinal or metric properties or relations are required, which then enable an ordinal or cardinal ranking of the objects. Lemons can be compared using the ordinal relation 'heavier than' and ranked accordingly. However, they can also be compared using the metric property 'weight in grams', which allows them to be placed in a cardinal ranking. Such a cardinal order not only permits the objects to be ranked, but also indicates the degree of difference between them. One lemon can then be described as not only heavier, but about 20 grams heavier than the other.

Comparing *degrees* of realisation as a potential solution to the incommensurability problem is based on metric properties. In the case of ordinal properties, an object's classification provides no information about the spectrum of the evaluation index in which it occurs. If item B is ordinally placed between A and C, it is not possible to tell where the three items are on an overall index, nor by how many grading increments they differ. All three could, for instance, be in close proximity at the lower or upper end of the scale; or A might be located at the upper end of the spectrum and C at the lower end, with B occupying any point in between. Comparing the degrees of realisation based on ordinal properties or relations is thus impossible.[28] Accordingly, degrees of realisation are a metric property. No matter how the comparison of degrees of realisation is structured, this must rely on a more complex property or relation that allows for scaling. Mere nominal or ordinal properties are not enough. This is one of the differences between proportionality in the strict sense and necessity. Necessity evaluations just require an ordinal ranking of options. A

---

[28]See also that the weight formula, Robert Alexy, 'On Balancing and Subsumption: A Structural Comparison' (2003) 16 Ratio Juris 433, 443–48, contrary to its ordinal three-level appearance, is based on a metric understanding of the level values, Petersen, 'Alexy and Proportionality' (n 5) 165 f: As mere ordinal values, they could not be calculated by means of a formula.

measure is necessary if it there is no other measure that is ordinally less intrusive and achieves the purpose ordinally as well or better. However, this is not the main difference between necessity and proportionality in the strict sense. Necessity requires a comparison of the different effects of different means on fundamental rights, on the one hand, and on state objectives, on the other. It does not require an assessment of the trade-offs between the effects on fundamental rights and state objectives.

Incomparability

In evaluative ethics, incomparability of objects is sometimes inferred from their incommensurability.[29] Yet as the previous discussion shows, this principle cannot apply absolutely. For all objects, at least some nominal properties can be found to permit comparison. This is also true for comparisons between abstract and physical objects, which are often cited as examples of incomparability.[30] Even the number 5 is nominally comparable to an actual lemon regarding whether the objects are abstract. In an absolute sense, thus, there are no incomparable objects.

However, objects may be incomparable with respect to a particular property or relation if the property or relation which guides the comparison cannot be meaningfully applied to both or to either of the objects to be compared. Two numbers cannot be compared as to which is yellow, and a number and a lemon cannot be compared as to which has the higher numerical value.

Similarly, two objects can be incomparable with respect to a metric property. The number 5 and a lemon are not only relatively incommensurable with regard to the property 'weight in grams', but also relatively incomparable. From relative incommensurability, however, only a relative, not an absolute incomparability follows. Incommensurability is a special case of incomparability, namely incomparability with respect to a certain metric property.

This also applies to objects which are put in relation to each other in the context of the constitutional doctrine of proportionality. Even if no conventional property can be found to guide comparisons between (for example) freedom of expression and national security, this does not entail that they are not in any way comparable. They can be nominally compared not only in terms of whether they are abstract objects but – to give a more normatively meaningful example – in terms of whether they are objects that the Basic Law recognises as legitimate purposes of state action.

Some authors distinguish between the incomparability of values as such and that of *instantiations* of values and infer the incomparability of the latter from the former.[31] If the *values* of freedom and equality are incomparable, so the argument goes, a provision that protects liberties and one that guarantees equality cannot be compared as *instantiations of values*. Yet, as a caveat to the seemingly obvious notion that a property – incomparability – that applies to values must also apply to their instantiations, two things must be noted:

[29]This view seems to prevail, see for instance Joseph Raz, *The Morality of Freedom* (Clarendon Press 1986) § 13; for further references, Ruth Chang, 'The Possibility of Parity' (2002) 112(4) Ethics 659, 660.

[30]For example, Ruth Chang, *Making Comparisons Count* (Studies in Ethics, Routledge; Taylor and Francis 2002) 85.

[31]Nien-hê Hsieh, 'Incommensurable Values' in Edward N Zalta (ed), *The Stanford Encyclopedia of Philosophy*, Fall 2020 Edition, <https://plato.stanford.edu/entries/measurement-science/> accessed 19 September 2024, sec 3.1.

First, the above discussion has shown that it is scarcely possible to make general statements about the incomparability of values. As for other items, the comparability of values is also relative to the property or relation on the basis of which they are to be compared. For example, freedom and equality can be compared *nominally* in terms of whether they contribute to justice. Whether they can also be compared ordinally in terms of which value is more important for justice is certainly debatable, and the cardinal comparison of which value reaches which result on a 'scale of justice' seems speculative at best. However, as with all other objects, the incomparability of values must be specified in terms of a certain property. Absolute incomparability does not hold for values either.

Second, what is usually meant by such incomparability claims is that at least some values cannot be compared on the basis of *another value*. This is an obvious position for anti-reductionist value pluralists. Thus, an axiological pluralist, who does not hold that justice and beauty are merely different manifestations of the good, might also hold that there is no value by which justice and beauty can be compared. Does it also follow that *instantiations* of justice and beauty cannot be compared? Would it, therefore, be impossible to compare a court decision and a lyrical poem in terms of the values of justice and beauty assigned to each?

Yet, why should the realm of values differ from that of dog shows? In both cases, the comparison could reference the degree of proportionate fulfilment of a respective ideal. Thus, one could ask how close the court decision comes to the ideal of a just judgment, and how far the poem approaches a poetic ideal. It is certainly true that the details of such a comparison would be difficult to elaborate, but cases are conceivable in which a poorly reasoned, glaringly incorrect court decision would come off worse in this comparison than a well-written poem with rich layers of deep meaning. Thus, even a value pluralist could judge which of the two embodiments of values is the better specimen of its kind – just as in the dog show, where guard dogs are judged according to quite different characteristics than lapdogs. Yet even dog shows do not involve a direct comparison of the different species as such. If we judge by degrees of realisation, then the incomparability of values does not – contrary to what might seem the obvious conclusion – necessarily imply the incomparability of instantiations of values.

Commensurability and incommensurability

Commensurability was originally a mathematical concept. Ancient mathematics had not yet discovered irrational numbers. This made it impossible, for example, to express how much longer the diagonal of a square was compared to its sides, even if the diagonal could be ordinally classified as longer. For ancient mathematics, diagonals and lateral lines were ordinally comparable, but not commensurable. The discovery of mathematical incommensurability is attributed to *Hippasus of Metapontum* (c. 6th century BC). Legend has it that the gods drowned him out of anger about the theoretical hardship his discovery caused.[32] This example from antiquity again shows that incommensurability is a special case of incomparability. In the following discussion,

---

[32]On the ancient mathematical history of the term, ibid 1.1.

the terms will therefore be used in this sense, even if they are sometimes used differently in the literature.[33]

Two objects are commensurable with respect to a property if the property is a metric property[34] and the objects are comparable with respect to this property. Two monetary assets are commensurable in terms of the amount in one currency. Both can be measured by numerical amount, and both numerical amounts can be compared metrically. Their difference can also be measured using the metric property. Two objects are incomparable with respect to a metric property if the property cannot be meaningfully related to one or both of the objects of comparison.[35] The property of numerical monetary value cannot, for instance, be meaningfully applied to the colour green and the number 5. They are incomparable with respect to this metric property.

As with any other kind of incomparability, incomparability with respect to metric properties is not absolute. In light of the infinite number of properties and relations of objects,[36] it can hardly be ruled out that some metric unit might also be found by means of which any two objects could be compared. Any two objects will be commensurable in some respect. This also applies to matters which are considered incommensurable in constitutional law. Even for values such as freedom of expression and national security, metric properties can be found which permit comparison. In corpus linguistics, for example, researchers examine language corpora to see how frequently certain expressions are used in certain contexts, thus allowing for conclusions about their social relevance and possibly changes in their significance. Here, comparisons are based on the numerical frequency of expressions used to describe a certain object. Similar comparisons could be made on the basis of a legal language corpus. The frequency of use of certain legal terms in legal documents could signify, for example, shifts in their relevance to legal practice. A higher frequency of court decisions employing the term 'national security' could indicate a growing challenge for fundamental rights. Even if such comparisons have no significance for constitutional doctrine, they again underline that it is hardly possible to find genuine incommensurability *per se* without specifying the property or relation to which the judgment refers.

This also means that there can be quite different commensurabilities for two objects, even if one metric property seems an obvious basis for comparison. Thus, for two sums of money, their respective cash value appears to be the self-evident candidate for a commensurable comparison. Two tax debts, for example, can be compared on the basis of their monetary value. If A has to pay €100 in taxes and B has to pay €200, then B pays twice as much tax as A. However, the two tax burdens could also be compared using other metric

---

[33]In the discussion about evaluative comparisons Raz (n 29) ch. 13 passim, does not distinguish between the two terms; Chang, *Comparisons* (n 30) 1, uses the term incomparability for a special case of evaluative comparison, in which a positive ratio determination on the basis of an at least ordinal property in the sense of 'better, equal, worse' is not possible; Hsieh (n 31) 1.2, makes a distinction according to whether the lack of comparability refers to values (incommensurability) or instantiations of values (incomparability).

[34]Some authors also extend the notion of incommensurability to ordinal properties, Hsieh (n 31) 1.1, but this is irrelevant for comparing degrees of realisation, as it relies on metric properties due to the gradual character of the comparison.

[35]It would hardly make sense to speak of the incommensurability of objects in a comparison based on a merely nominal property. At least in conceptual terms, this does not make sense even for ordinal properties. Whether one should expand the meaning of this concept need not be addressed in the present case, since the comparison of the *degrees* of realisation presupposes a cardinal comparison in any case.

[36]See Geert Keil, 'Über die deskriptive Unerschöpflichkeit der Einzeldinge' in Geert Keil and Udo Tietz (eds), *Phänomenologie und Sprachanalyse* (mentis 2006), 83.

properties and relations, for instance 'percentage of income'. If A earns €1,000 € and B €4,000, A pays twice as much tax as B as a percentage of income even though only half the numeric amount. Comparing taxes as the ratio of tax to income produces a completely different result than comparing raw amounts.

This example shows that the number of possible comparisons is limited only by the number of possible relations which can be used to develop comparative values. The complexity of the relation forming the basis of comparison can be increased at will. Tax law, for example, generally accounts for the burden of the taxation on people's lives. Taxes on income that just barely covers the existential necessities of life are much more burdensome; furthermore, once past the subsistence level, income levels exhibit decreasing marginal utility. This observation forms the justification for progressive taxation. In a progressive tax system, different percentage tax amounts can ultimately result in equal tax burdens. The comparison is thus driven by a relationship that incorporates progressivity as a functional characteristic. When this characteristic is applied, progressive tax burdens can be equal.

Whether two objects are the same or different – and how different they are – is a function of the metric relation which drives the comparison. Metric relations can be determined using any sort of mathematical function. Their number is thus potentially infinite. Even one metric property by which two objects can be compared – such as monetary value in a currency – is sufficient to allow the derivation of a potentially infinite number of relations whose values can then be compared.

Commensurability is achieved through units of measurement. In our context, in which we try to explore the potential and limits of comparisons relying on degrees of realisation of reference values, it is worth noting that all our conventional units of measurement are themselves based on degrees of realisation of a reference value. For example, the length of an object in metres was determined over centuries by the relation of the object's length to the length of the prototype meter in Paris. With the help of stipulated reference values, conventional metric properties can be created and also changed. Since 1983, for example, the metric property 'length in metres' is no longer determined by the relation to the original prototype meter in Paris, but based on the relation to a natural constant, namely the distance that light travels in a vacuum in 1/299 792 458 of a second.[37] Measurements require three elements: two reference values to create a scale and a scaling function. Most of our everyday measurements of length, time, and weight set one reference value to be a natural zero – extensionlessness, timelessness, and weightlessness – and stipulate the other, historically, via paradigm objects and, currently, using natural constants. This is only different for temperature. Unlike the Kelvin scale, absolute zero is neither the zero value for commonly used Celsius nor Fahrenheit scales.[38] Furthermore, our everyday scales are usually linear. However, linearity is not a constitutive feature of measurements. For example, the Richter scale, which enables the commensurable comparison of the strength of earthquakes, is exponential.

Both the cardinal comparison of objects according to a metric property and all our conventional standards for length, volume, weight, time etc. are based on the comparison
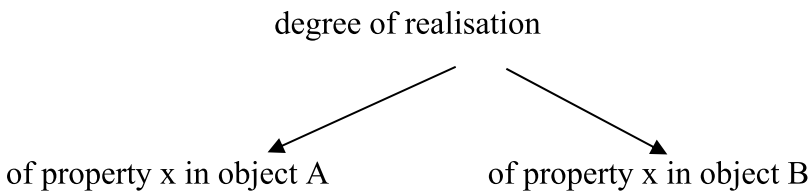
---

[37]Definitions of SI Base Units (National Institute of Standards and Technology) <https://physics.nist.gov/cuu/Units/current.html> accessed 19 September 2024.

[38]Cf Hasok Chang, 'Spirit, air and quicksilver: The search for the "real" scale of temperature' (2001) 31(2) Historical Studies in the Physical and Biological Sciences 249.
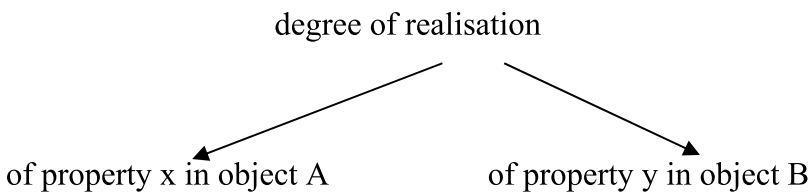
of the degree of realisation of reference values and are thus, in this sense, judgments of proportionality. Until 1983, when two jumps were compared in a long jump competition, this was done based on the degree of realisation that the distance between the jump-off mark and the first contact with the ground had in relation to the prototype meter in Paris. A jump of 'five metres' meant it realised the length of the prototype meter five times. The measurement of distance is the proportionate relation – one to five – of the prototype meter and the jump. Commensurable comparisons are thus comparisons of degrees of realisation, even with our conventional units of measurement. The fact that degrees of realisation must be compared in dog shows, decathlons, and constitutional law is not a peculiarity of dog shows, the decathlon, or constitutional law, but a particular feature of all commensurable comparisons. What, then, is special about dog shows, decathlons, and the law?

Commensurable comparisons consist in comparing the degrees of realisation of a shared property. Incommensurable comparisons consist in comparing the *comparata* regarding different properties.

Commensurable Comparisons

degree of realisation

of property x in object A   of property x in object B

Comparisons of Incommensurables

degree of realisation

of property x in object A   of property y in object B

In the dog show example, poodles and German shepherds are compared regarding the degree of realisation of different properties, namely the ideal poodle and the ideal German shepherd. But how is that possible given the general structure of comparisons, which require a shared property of the *comparata*? This obstacle can be overcome thanks to the reflexivity of properties. We can base the comparison on a property that the different properties share. If the properties are metric properties, we now know that they all share the property of realising them to a certain degree which allows for assigning metric values to them. Thus, it becomes possible to compare this second-order property, which the different properties of the *comparata*, and consequently the *comparata* themselves, share. Both *comparata* share the property of having metric values assigned to them corresponding to the degree of realisation of a metric property – even if the metric properties are not shared. The comparison of *comparata* with different properties can only be

achieved by basing the comparison on the numeric values of the degrees of realisation of the different properties.

Such comparisons do not always make sense. This made it easy for Justice *Scalia* to ridicule proportionality judgments. It does not make sense to compare the numerical values of a line's length with the numerical value of a rock's weight only because both line lengths and rock weights share the property of having achieved a certain numerical value according to their measuring unit. That such comparisons of second-order properties are possible does not necessarily make them useful. However, that is not so different for first-order comparisons either. To compare the weight in grams of a stamp and an aircraft is also possible, but equally senseless. Actually, a more revisionist take on the distinction between commensurability and incommensurability might insist that the only difference between them is that comparisons of degrees of realisation relying only on one metric property are more often sensible or useful. It could also be pointed out that, in commensurable comparisons, what is actually compared is not the property as such but the degrees of realisation of that property. That said, the fact that this is more often useful than the comparison of degrees of realisation of different metric properties makes no difference to commensurability. In both cases, only the numerical values of the degrees of realisation are commensurable. Whether such a comparison makes sense is circumstantial in any event. Suppose a certain material has very low adhesive qualities which make it as difficult to form one metric unit of a line than to aggregate one metric unit of weight. In that case, it might be useful to compare the numeric value of each metric property to assess the skill or ingenuity of an engineer working with this material. Regardless of whether such revisionist takes on the distinction seem advisable, the fact remains that there is a structural difference between the commensurable and incommensurable comparisons regarding the number of properties involved, and that this structural difference does not preclude comparisons of incommensurable properties. This is the case because a comparison can always be based on the degrees of realisation – a secondary property that both *comparata* share, even if only for different properties.

III. Comparing the incommensurable

In having detailed all the relevant elements, the preceding sections have set the stage for a complete description of the structure of incommensurable comparisons. To get a better understanding of the limitations of incommensurable comparisons in law, this structure will first be discussed in abstract terms, followed by the example of the decathlon with its highly performative comparison of incommensurables, and only then for the law.

The abstract structure

The abstract structure of incommensurable comparisons relies on the same elements as comparisons in general. However, to arrive at these elements, incommensurable comparisons require more work.

Objects and purpose

The abstract structure of incommensurable comparisons has to comprise the elements that every comparison entails. It has to relate at least two objects to a purpose of

comparison. In the dog show example, the purpose of the comparison of dogs of different breeds is to find the most suitable dog to be crowned Best in Show.[39]

Properties

As just noted, comparisons of incommensurable properties have the structural characteristic of pertaining to two different properties. These properties are by and large not conventional properties, even when they relate to these conventional properties. If it were a quality of poodles to maintain a certain weight and for German shepherds to attain a certain height, it would still make no sense to compare the numerical values of conventional weight and height. Rather, some new metrics would have to be defined that relate to the ideal weight of poodles and the ideal height of German shepherds. The degrees of realisation of these new metric properties could then be meaningfully compared.

Since both properties have to be metric, they require the creation of scales,[40] as scales are a structural element of metric properties. It follows that unconventional scales for each of the *comparata* would have to be stipulated. Scaling in turn requires the definition of an upper and a lower reference value and the determination of a scaling function for the interval. In the case of dog shows, scales might be based on an average dog of a breed as the lower and ideal types as the upper reference value for each breed. The scales could be linear, but also progressive, or regressive, due to exponential marginal efforts to approach the ideal.

The construction of new incommensurable metric properties is often complicated by multidimensionality. Whereas most of our everyday commensurable comparisons are based on unidimensional properties such as length, weight, or time, those used in incommensurable comparisons are often multidimensional. This makes it necessary to develop a scale for each dimension, to make them commensurable regarding their degree of realisation, and to assign a weight to each dimension to deal with trade-offs between the dimensions. Multidimensionality leads to 'nested' incommensurability. The process of creating meaningful commensurability of degrees of realisation of incommensurable properties has to be repeated for each dimension. The commensurable degrees of realisation of each dimension can then be used to calculate the degree of realisation of the multidimensional property. However, multidimensionality is not an issue specific to comparisons of incommensurable properties. If two lemons are to be compared with respect to their yellowness, the comparison relates to a shared property; but hue, saturation, and luminance make for at least three dimensions of the intensity of the colour. This would make it necessary to develop scales and aggregation formulas for at least three dimensions.

Weight, polarity, and relevance

Although it might seem natural to assign equal weight to the incommensurable properties under comparison, they could also be weighted in other ways that might make more

---

[39]For example: Andrew Das, 'Westminster Dog Show 2021: Wasabi the Pekingese Wins Best in Show' *The New York Times* (New York, June 14 2021) <https://www.nytimes.com/live/2021/06/13/sports/westminster-dog-show#who-won-on-saturday> accessed 19 September 2024.

[40]Eran Tal, 'Measurement in Science' in Edward N Zalta (ed), *The Stanford Encyclopedia of Philosophy*, Fall 2021 Edition, <https://plato.stanford.edu/entries/value-incommensurable/> accessed 19 September 2024, sec 3.4.

sense or even be called for.[41] In the example of the dog show, it might be that breeding and training a poodle that comes close to the ideal of its breed is much easier than breeding and training a German shepherd accordingly. If that were the case, it might be reasonable to assign more weight to the ideal German shepherd.

In addition, the polarity of the properties has to be determined. In the case of the dog show, the dog which comes closest to the ideal of its breed wins. However, as mentioned above, depending on the respective purpose it might make more sense to favour a point on the scale that indicates a lower degree of realisation, equilibrium, or anything in between.

Establishing new incommensurable metric properties as objects of comparison extends the range of possible comparisons. However, the effort that goes into the construction of these properties is only worth it if each property actually serves the purpose of the comparison at hand. While comparisons based on incommensurable properties are flexible when it comes to which properties are used, these do need to be relevant and speak to the same purpose. For example, the metric property 'ideal poodle' does not serve the purpose of judging both poodles and German shepherds, but it does serve the purpose of judging poodles – and vice versa for the metric property 'ideal German shepherd'. That said, both properties are relevant when judging the quality of each dog.

Measurement

In a final step, the degrees of realisation have to be measured. For commensurate comparisons of conventional properties, such as weight in kilograms or length in centimetres, measurement is generally not an issue. The situation can look much more complex for the scales stipulated/constructed for incommensurable comparisons. Multidimensional metric properties in particular might require multiple measurements; what is more, we lack objective instruments corresponding to those used for conventional metric properties.

The showcase example: decathlon

In measurement theory, two dimensions of measurement are distinguished: their accuracy and their precision.[42] Accuracy pertains to how well a measurement tracks differences in reality; precision tracks how reliable the measurement is. In archery for example, accuracy tracks how close to the bullseye an arrow hits; precision tracks the spread of all arrows. While measuring differences in incommensurable comparisons evidently poses a number of challenges, there are nevertheless examples in which they are extremely accurate and precise. One such example is decathlon. Comparisons in decathlon of performances in incommensurable disciplines such as sprints and jumps are both accurate and precise. They can track differences in performances down to hundredth of seconds and centimetres, and the comparisons of these minute differences in performances is highly reliable. Why is it that a sport that consists of comparing the performance

---

[41]On the necessity of justifying the substitution rates between the objects of comparison, see Veel (n 22), 192.
[42]Tal (n 40) 8.3.

of athletes in ten very different disciplines works so well? This deserves a closer look to see how the different elements of comparisons are handled in this context.

Objects and purpose

The *comparata* of the decathlon are the performances of each athlete in the ten constitutive disciplines. The purpose of the comparison is equally straightforward: to rank the competitors in order of their overall athletic performance.

Properties

As such, the comparison of the overall performance is not an incommensurable comparison since all athletes are compared based on the same overall performance metric. However, consisting of ten disciplines, decathlon is an example of a multidimensional scenario (i.e., ten dimensions), making for tenfold nested incommensurability. Hence, in order to establish an overall performance metric, we need to find a way to compare performances across disciplines. These comparisons are based on incommensurable measurements in each discipline, which have to be converted into point scores that enable the calculation of an overall score. Based on ten newly established metric properties designed to further the purpose of the comparison, these point scores enable us to compare and add up the performances in different disciplines down to hundredth of seconds and centimetres in accuracy, for example in a 100-meter sprint and the length of a jump track.

One factor which contributes to the accuracy and precision of comparisons in decathlon is that the properties compared across disciplines are not in themselves multi- but unidimensional. Thus, cross-discipline comparisons in decathlon do not also entail the weighing up of different aspects of performance. Conversely, other sports such as figure skating or gymnastics rely on multidimensional performance properties such as athletic prowess and aesthetic execution. It would complicate the comparisons in decathlon considerably if it were to include such disciplines. In fact, sports that are affected by such multidimensionality when it comes to their objectivity acknowledge this issue and aim to address it by involving several referees that try to create a certain objectivity by averaging a plurality of assessments.

However, even though the decathlon properties are unidimensional they still have to be constructed. A comparison across disciplines cannot rely on the numeric values that are measured in the different disciplines according to different conventional metrics such as seconds and metres. To compare the numeric values of a 10.65-second 100-meter sprint with a 6.85-meter jump neither makes sense nor serves the purpose of the comparison. This means that new metric properties have to be constructed for which it is possible to determine degrees of realisation that do serve the comparison despite the disciplines being compared based on different properties.

Combined sports events can look back on a long tradition of establishing these kinds of properties. Ancient pentathlon, dating back to 700 B.C.,[43] and early modern combined sports competitions were based on ordinal rankings. They just added up the ordinal

---

[43]IAAF Council, *IAAF Scoring Tables for Combined Events* (2001) <https://worldathletics.org/download/download?filename=53f7d332-be0c-434c-8467-1d9078966147.pdf&urlslug=IAAF> accessed 19 September 2024, 7.

numbers of each athlete's placements in the different competitions and thus arrived at a point score that allowed a crude overall cardinal ranking.[44] The elaboration of metric scales based on reference values that enable more nuanced cardinal rankings has pre-occupied the combined disciplines since the late nineteenth century. The first scoring tables were linear and set the upper reference value at the current world or national record, and the lower reference value often at the average performance of junior athletes.[45] The resulting interval between the two values was generally subdivided to form a 1,000-point scale. Linear scales were used up until 1934.[46] However, such linear scoring tables could not take into account that it becomes progressively more difficult to further enhance top performances. Progressive formulas were first developed in 1912, but not adopted by the International Association of Athletics Federations (IAAF), the predecessor organisation to today's World Athletics, until 1934. The progressive formula currently in use was developed by the Czech Viktor Trkal in 1984 and reflects nine principles set by the IAAF.[47] The two most important principles require that the scoring tables be progressive and that comparable performances in each discipline be assigned the same point score. If the formula were based on world records alone, it would require that the world record in each discipline correspond to the same point score. However, setting an upper reference value under the current formula is more complex. It is calculated based on the world record and the average of the performances of the 30 best specialists and the 100 best decathletes as of 1984.[48] Trkal's efforts to comply with the nine principles of the IAAF resulted in the formula:

$$S = A \times (P - B)^C \text{ for disciplines in which greater distance or height, i.e. a higher}$$
$$\textit{metric, indicates a better performance}$$

$$S = A \times (B - P)^C \text{ for disciplines in which faster time, i.e. a lower metric, indicates}$$
$$\textit{a better performance}$$

S is the point score; A is a calibrated parameter based on the point score of 1,000 being assigned to the upper reference value; B is a performance resulting in 0 points; P is the performance in the metric of the discipline, i.e., the result recorded; and C, i.e., >1, serves as the progression factor.[49] For example, for men's 100-meter sprint, A is set at 25.4347, B at 18.00, and C at 1.81; for men's pole vaulting, A is set at 0.2797, B at 100.00, and C at 1.35.[50] On the basis of Trakl's formula, scoring tables can be calculated. Below, you can find an excerpt from such a table for men's 100-meter sprint:[51]

---

[44]ibid 7 f.
[45]ibid 11.
[46]ibid 12 f.
[47]ibid 17 f.
[48]Michael Fröhlich and others, '30 Jahre Bewertungstabelle im Zehnkampf: Ist eine Revision nötig?' (2016) 57(2) LSB (Berlin) 81, 82.
[49]Guillaume Chèze, 'Decathlon Rules: An Axiomatic Approach' (2021) 2(1) Mathematics and Sports 1, 3.
[50]IAAF Council (n 43) 22.
[51]ibid 48.

Perf.	Points	Perf.	Points	Perf.	Points	Perf.	Points	Perf.	Points
9.50	1223	10.00	1096	10.50	975	11.00	861	11.50	753
9.51	1221	10.01	1094	10.51	973	11.01	858	11.51	750
9.52	1218	10.02	1091	10.52	970	11.02	856	11.52	748
9.53	1215	10.03	1089	10.53	968	11.03	854	11.53	746
9.54	1213	10.04	1086	10.54	966	11.04	852	11.54	744
9.55	1210	10.05	1084	10.55	963	11.05	850	11.55	742
9.56	1208	10.06	1081	10.56	961	11.06	847	11.56	740
9.57	1205	10.07	1079	10.57	959	11.07	845	11.57	738
9.58	1202	10.08	1076	10.58	956	11.08	843	11.58	736
9.59	1200	10.09	1074	10.59	954	11.09	841	11.59	734
9.60	1197	10.10	1071	10.60	952	11.10	838	11.60	732
9.61	1195	10.11	1069	10.61	949	11.11	836	11.61	730
9.62	1192	10.12	1066	10.62	947	11.12	834	11.62	728
9.63	1190	10.13	1064	10.63	945	11.13	832	11.63	725
9.64	1187	10.14	1062	10.64	942	11.14	830	11.64	723
9.65	1184	10.15	1059	10.65	940	11.15	827	11.65	721
9.66	1182	10.16	1057	10.66	938	11.16	825	11.66	719
9.67	1179	10.17	1054	10.67	935	11.17	823	11.67	717
9.68	1177	10.18	1052	10.68	933	11.18	821	11.68	715
9.69	1174	10.19	1049	10.69	931	11.19	819	11.69	713
9.70	1172	10.20	1047	10.70	929	11.20	817	11.70	711
9.71	1169	10.21	1044	10.71	926	11.21	814	11.71	709
9.72	1166	10.22	1042	10.72	924	11.22	812	11.72	707
9.73	1164	10.23	1040	10.73	922	11.23	810	11.73	705
9.74	1161	10.24	1037	10.74	919	11.24	808	11.74	703
9.75	1159	10.25	1035	10.75	917	11.25	806	11.75	701
9.76	1156	10.26	1032	10.76	915	11.26	804	11.76	699
9.77	1154	10.27	1030	10.77	912	11.27	801	11.77	697
9.78	1151	10.28	1028	10.78	910	11.28	799	11.78	695
9.79	1149	10.29	1025	10.79	908	11.29	797	11.79	693
9.80	1146	10.30	1023	10.80	906	11.30	795	11.80	691
9.81	1144	10.31	1020	10.81	903	11.31	793	11.81	689
9.82	1141	10.32	1018	10.82	901	11.32	791	11.82	687
9.83	1139	10.33	1016	10.83	899	11.33	789	11.83	685
9.84	1136	10.34	1013	10.84	897	11.34	786	11.84	683
9.85	1134	10.35	1011	10.85	894	11.35	784	11.85	681
9.86	1131	10.36	1008	10.86	892	11.36	782	11.86	679
9.87	1128	10.37	1006	10.87	890	11.37	780	11.87	677
9.88	1126	10.38	1004	10.88	888	11.38	778	11.88	675
9.89	1123	10.39	1001	10.89	885	11.39	776	11.89	673
9.90	1121	10.40	999	10.90	883	11.40	774	11.90	671
9.91	1118	10.41	996	10.91	881	11.41	771	11.91	669
9.92	1116	10.42	994	10.92	878	11.42	769	11.92	667
9.93	1113	10.43	992	10.93	876	11.43	767	11.93	665
9.94	1111	10.44	989	10.94	874	11.44	765	11.94	663
9.95	1108	10.45	987	10.95	872	11.45	763	11.95	661
9.96	1106	10.46	985	10.96	870	11.46	761	11.96	659
9.97	1103	10.47	982	10.97	867	11.47	759	11.97	657
9.98	1101	10.48	980	10.98	865	11.48	757	11.98	655
9.99	1099	10.49	977	10.99	863	11.49	755	11.99	653

IAAF Scoring Tables for Combined Events / Tables de Cotation pour les Epreuves Combinées

As the table shows, 100-meter sprint performances can be compared with all other disciplines down to one hundredth of a second in terms of accuracy.

Weight, polarity, and relevance

In multidimensional comparisons, there are two possible dimensions of weight. One relates to the weighing of the different dimensions – here the different disciplines. The

IAAF makes its ambition in this regard very clear, with the second of its nine principles for the development of the scoring table requiring the following: 'Results in various events should, as far as possible, yield about the same number of points if the results are comparable as to quality and difficulty'.[52] In decathlon, there is no 'supreme' discipline whose results carry more weight. The second dimension relates to the weight of the object of comparison. Here, too, decathlon assigns equal weight to the overall performance of each athlete. An alternative would be to assign unequal weight factors to the performance of different athletes, for example in the event of women and men competing against each other.

Since the purpose of athletic competitions is to rank athletes in order of their performance, the polarity of the properties is obvious. Sprint events reward faster times, i.e., lower values; field events reward greater distance or height, i.e., higher values. This polarity is captured by the two variants of Trkal's formula. As for the corresponding point scores, higher numbers designate a better performance.

What is more, point scores and their polarity obviously relate to athletic performance and thus to the purpose of the comparison.

Measurement

With scoring tables in place, this leaves only measurements. In principle, this seems straightforward for the disciplines that make up decathlon, since all disciplines can rely on unidimensional conventional measurements. However, even here things are not as simple as they might appear. The reason for this is that some of the disciplines use different measurement methods. Sprint events, for example, can be measured automatically or manually. In turn, an adjustment factor (0.24 for events below 400 m, 0.14 sec for 400 m) applies when manually timing an event; this factor is subtracted from the B parameter of the formula.[53]

The history of scoring tables in decathlon and the current formula –including the values developed for the different disciplines – show that establishing relevant scales by way of reference values that allow for a precise determination of degrees of realisation for comparisons of incommensurable properties is anything but trivial. The fact that establishing such scales means to stipulate values leaves room for alternative constructions and contestation. Critics of the current decathlon formula accuse it of falling short of the aspired equal weighting of all disciplines.[54] Some believe the current reference values to be outdated and biased, claiming that they distort the evaluation of individual performances and foster selection effects among athletes – in turn disadvantaging some disciplines and contradicting the spirit of the decathlon: upholding the equal value of all ten disciplines.[55]

IV. Proportionality in the strict sense in law

The example of decathlon has shown that accurate und precise comparisons of incommensurables are indeed possible. However, it has also demonstrated what it takes to

---

[52]ibid 17.
[53]ibid 25.
[54]Chèze (n 49), 4 f.
[55]Fröhlich and others (n 48), 94 f.

deliver them, and this is the main point of this article. It does not so much focus on whether precise comparisons between rights and state interests are possible. It rather focusses on what kind of efforts would have to be expended to create the preconditions to compare incommensurables with some degree of precision. It would take the setting of reference values and scaling functions determined for all the dimensions of such often multidimensional *comparata*; not to mention a determination on how the *comparata* are measured and weighted with accuracy and precision. For the law, the phenomenology looks far from promising – something that will become apparent when each of the necessary preconditions for comparisons of incommensurables are examined in more detail.

Comparing fundamental rights and state interests

Judgments on proportionality in the strict sense require the comparison of incommensurables such as freedom of speech and national security. How can the challenges inherent in such comparisons, which have been so effectively overcome in decathlon (albeit not uncontested), be addressed in constitutional and human rights law – two fields in which the proportionality standard is regularly invoked? The answer can be broken down by revisiting the individual elements outlined above.

Objects and purpose

In fundamental rights law, the objects of comparison are most often a specific fundamental right and a countervailing state interest. The state interest can be a policy aim, the protection of fundamental rights of third parties, or some value.

The purpose of comparing fundamental rights with state interest is quite different from the comparisons in decathlon or dog shows. The purpose of comparing the latter is to select the best athlete or dog. Selecting for the best is an almost natural purpose of comparisons. It is also how we use comparisons in everyday life situations, for example when we hunt for the best restaurant, movie, car, etc. In constitutional law, however, the purpose of comparison is not to select the best interest, when we compare fundamental rights and state interests. Rather the purpose of comparison has been established in legal doctrine. We compare the intensity of a fundamental rights infringement with the achievement of a state interest in order to ensure some kind of balance between the two for the purpose of protecting fundamental rights, in particular to safeguard the interests of minorities against those of the majority, which controls which state interests are pursued.

Properties

The properties of fundamental rights and state interests that are compared as part of a judgment on proportionality in the strict sense are the severity of the fundamental rights infringement versus the achievement of the state interest. Unlike the properties compared across the different disciplines of decathlon, these properties are often highly multidimensional. State measures can infringe upon different aspects of fundamental rights, the relevance of which might even be contested. For example, with some fundamental rights it is unclear whether their relevance should be determined solely with regard to individual freedom, or also in terms of how the freedom of

society as a whole is affected. Some infringements are severe but affect only a few rights-bearers; others – such as dragnet searches or mass data retention – are relatively moderate but affect many rights-bearers. Moreover, the individual and the societal dimension of a right can themselves be multidimensional. Thus, someone who understands freedom of speech in terms of its functional value to democracy will accord more weight to political speech than private expressions of opinion, while someone who views freedom of expression as a good in itself will make smaller distinctions, if any. Multidimensionality also affects state interests. A state measure often simultaneously pursues several regulatory aims, and these might be achieved to different degrees.

What would it take to enable a reliable comparison of these complex multidimensional properties? Theoretically, the first requirement would be consensus as to the relevant dimensions. Secondly, it would take the establishment of a metric property for each of the dimensions of fundamental rights and the state interest by stipulating an upper and lower reference value and a scaling function for the interval in between. One would also need to stipulate an aggregation function for the values of each dimension. All of this would need to happen – thirdly – to establish the higher-order metric properties 'fundamental right infringement' and 'achievement of state interest' as a precondition to comparisons of degrees of realisation of both incommensurable properties. The challenge for the law is much more challenging than for decathlon, in which all events that have to be compared are unidimensional. Each event is measured in just one dimension – either time or distance. The comparison of 'fundamental right infringement' and 'achievement of state interest' resembles a combined sports event in which two disciplines must be compared that are measured in multiple dimensions such as figure skating and dancing, which are judged by athletic performance and aesthetics. This is not to say that metric properties for higher-order incommensurability comparisons could not be stipulated, it, however, takes the stipulation requirements literally to another level.

Nothing that comes remotely close to meeting such requirements can be observed for the typical judgments on proportionality in the strict sense in fundamental rights law. None of the preconditions of reliably comparing degrees of realisation of incommensurable properties are met. On the surface, this makes judgments on proportionality in the strict sense appear impossible. So, does it follow that the critics who regard incommensurability as an insurmountable problem for proportionality have had it right all along?

In fact, we need only turn to everyday comparisons between incommensurable options such as restaurants, holidays, jobs, etc. to see that such a radical conclusion is not warranted. We largely seem to be able to compare incommensurable options just fine without explicitly meeting all preconditions of such judgments. However, the fact remains that – for theoretical reasons – such comparisons do require us to meet these preconditions somehow. What seems to happen is that we implicitly – largely intuitively and subconsciously – stipulate some reference values and some kind of scaling and weighting function when we evaluate incommensurable alternatives. From a phenomenological perspective, the process appears to be analogous to other intuitive judgments encountered in everyday life. When crossing a busy intersection, we intuitively estimate the speed and trajectory of vehicles approaching from different directions, and navigate our own path to avoid collisions. We do not consciously calculate the exact speeds or geometrical trajectories of other vehicles, yet we usually successfully avoid collisions.

However, in normative contexts, this implicit, intuitive, and largely subconscious process of stipulation has specific costs due to the idiosyncratic nature of normative evaluations in comparison to guessing physical outcomes. Firstly, it should be noted that the intuitive normative scales will be of a very crude nature. This is evident in our everyday judgments in that we will be able to make up our minds between incommensurable alternatives of restaurants, holidays, jobs, etc. if the difference in quality is obvious; however, we will struggle when the differences are less pronounced. Second, the implicit and largely subconscious process will result in different scales for different people. The implicit constructions will be idiosyncratic.[56] Anyone who has ever discussed which restaurant to go to or which movie to see with a larger group will probably vividly remember how these differences can play out. The same holds true for judgments on proportionality in the strict sense in law. There are very different opinions on how serious a security threat the traffic disturbances caused by climate activists who glue themselves to the streets are. Some see them on par with organised crime, if not terrorism; for others, they are merely a nuisance like other traffic jams caused by construction work or mass individual commuting. These differences are at least in part caused by the implicit differences in the construction of reference values for security and the scaling functions. The idiosyncrasy is also at full display in the famous Hutterian case of the Canadian Court. The Canadian legislator had eliminated a religious exemption from the requirement of driver's licence having to be equipped with a photo of the licence holder. The majority opinion held that the cost of not being able to hold a driver's licence for members of the Hutterian Brotherhood, who rejected to be photographed on grounds of their religious beliefs, 'do not rise to the level of seriously affecting the claimants' right to pursue their religion'.[57] The dissenting opinion of Judge Abella comes to the almost diametrically opposite evaluation: 'The harm to the constitutional rights of the Hutterites, in the absence of an exemption, is dramatic'.[58]

Conditions for establishing metric properties are somewhat less dire in legal domains in which the law itself provides some type of scale. One example would be criminal law, where punishment must be proportionate to the offender's culpability. Punishment and guilt are incommensurable; they, too, are only comparable based on the degrees of realisation of the severity of the punishment and the extent to which an offender is culpable. Although there are no decathlon-like scales here either, the legislature has established penalties and sentences for a wide range of offences. Murder stands at the far end as a reference point for the most severe punishment, and there are many point scores on the descending scale of severity from (for example) rape down to armed robbery down to petty theft. The legislator has thus provided a ranked catalogue of offences and punishments which provides orientation for classifying comparable offences within this legally predefined spectrum. It is therefore not surprising that assessments of so-called 'retrospective' proportionality (supra at footnote 6) in criminal law are sometimes viewed less sceptically than the assessment of 'prospective' proportionality.

---

[56]Cf Sartor (n 6), 1443, on the 'arbitrariness' of assigning weights to values.
[57]*Alberta v Hutterian Brethren of Wilson Colony* [2009] 2 SCR 567, para 99.
[58]*Alberta v Hutterian Brethren of Wilson Colony* [2009] 2 SCR 567, para 114 (Abella J).

Weight, polarity, and relevance

Since fundamental rights infringements and state interests are often multidimensional, two types of weights have to be assigned. For the aggregation of realisation values in each dimension, weights have to be assigned to each dimension of each of the two incommensurable properties. In addition, the properties as such need to be weighted.

In constitutional law, courts often emphasise the equal value of all fundamental rights.[59] However, most constitutional courts consider different state aims to be differently important. The Supreme Court of the United States, for example, distinguishes between legitimate, important, and compelling state interests.[60] The Canadian Supreme Court distinguishes administrative convenience from pressing and substantial objectives.[61] The European Court of Human Rights distinguishes permissible grounds from other legitimate aims.[62] Even in the decision which introduced the proportionality standard into German constitutional law, the Federal Constitutional Court differentiated between paramount common interests, important common interests, and mere general interests of the community.[63] These differences in assigned weights must be taken into account when comparing degrees of realisation. The degree of realisation of a less important state objective should not outweigh a loss of freedom of the same degree. In contrast to the decathlon, however, such weight assignments are rarely made explicit, or even determined anywhere near as precisely as in combined sports. Nor is there likely to be a consistent intersubjective agreement on the appropriate weighting in the law.[64]

It goes without saying that the assignment of weights at both levels – the dimensions of the properties and the properties as such – also heavily relies on intuitive decisions that stay largely implicit in the overall proportionality judgment.

The polarity of fundamental rights infringement and achievement of state interest is more complicated. Unlike in decathlon, the polarity of comparisons in fundamental rights law is not the *superiority* of one degree of realisation over the other, but rather some sort of *balance* between the degrees of realisation of rights and state interests. This balance that is the aim of proportionality in the strict sense should ensure that individual rights are defended to a certain extent against the interests of the majority and the state – that they should not be absolutely subordinated to collective goals. Comparing degrees of realisation should ideally guarantee that this balance has been properly struck.

There are, however, different notions of how an adequate balance should be conceptualised. 'Optimisation' models involve the most demanding standard. In German

---

[59]Matthias Cornils, '§ 168 Allgemeine Handlungsfreiheit' in Josef Isensee and Paul Kirchhof (eds), *Handbuch des Staatsrechts der Bundesrepublik Deutschland* (vol 7, 3rd edn C. F. Müller 2009), para 102; Niels Petersen, *Verhältnismäßigkeit als Rationalitätskontrolle: Eine rechtsempirische Studie verfassungsgerichtlicher Rechtsprechung zu den Freiheitsgrundrechten* (Jus publicum vol 238, Mohr Siebeck 2015) 58; in favor of a differentiation between at least the general freedom guarantee from more specific rights under the German constitution Jakob Hohnerlein, 'Grundrechtlicher Schutz der Willkürfreiheit: Eine Materielle Grenze der politischen Gestaltung?' (2022) 61 Der Staat 637, 651.

[60]*United States v Carolene Products Company* 304 US 144 (1938) fn 4.

[61]*Frank v Canada (Attorney General)* [2019] 1 SCR 3; Charles-Maxime Panaccio, 'The Justification of Rights Violations: Section 1 of the Charter' in Peter Oliver, Patrick Macklem and Nathalie Des Rosiers (eds), *The Oxford Handbook of the Canadian Constitution* (OUP 2017) 666.

[62]*Witold Litwa v Poland* App no 26629/95 (ECtHR, 4 April 2000), para 49; *Kohlhofer u Minarik v The Czech Republic* App no 32921/03, 28464/04, 5344/05 (ECtHR, 15 October 2009), para 96.

[63]BVerfGE 7, 377 (408) (english translation: <https://law.utexas.edu/transnational/foreign-law-translations/german/case.php?id=657> accessed 19 September 2024).

[64]Among other things, the proposal by Veel (n 23), 211 f., to determine proportionality by adapting the negotiation solution of *Nash* also fails because of the different weights of the goods to be weighed.

constitutional law, these models are associated with the formula of 'practical concordance' coined by *Konrad Hesse*, which has had a decisive influence on the case law of the Federal Constitutional Court: 'Constitutionally protected legal interests must be balanced with one another in the solution of the problem in such a way that each of them gains effectiveness. … Limits must be set to both interests so that both can achieve *optimal* effectiveness'.[65] The last half-sentence in particular can be understood to mean that optimal effectiveness aims to achieve the same degree of realisation and the highest possible realisation for both fundamental rights and state interests. The theory of fundamental rights of *Alexy* and his school, who understand fundamental rights as optimisation requirements, also aims at optimisation in the sense of equal degrees of realisation.[66] A narrower reading of the optimisation objective requires only that the degree of realisation of fundamental rights never lag behind the degree of realisation of state interests.

Less demanding normative approaches, by contrast, require that there be no *significant disproportion* in the degrees of realisation resulting in an inadequate realisation of fundamental rights.[67] If the state chooses a suitable means for a legitimate end, and this means is simultaneously the most protective of fundamental rights, the proportionality in the strict sense only prohibits the state from pursuing the legitimate end if this would sacrifice a disproportionately large degree of fundamental rights in return for a small gain in achieving the state's aim.

Both approaches allow degrees of realisation to be compared in two ways. On the one hand, we might compare how the overall degree of realisation of an affected fundamental right compares to the degree of realisation of the state aim as pursued by the state action. Does the degree of realisation of the fundamental right outweigh that of the aim? For illustrative purposes, let us counterfactually assume that the degrees of realisation and achievement of a state interest could be quantified with percentage values. Then a state measure would be unconstitutional, under optimising approaches, if it permits a fundamental right to be realised only to 50% overall, while the state aim is realised to 70%. Less demanding approaches which forbid only clear imbalances would – again simply as a matter of illustration – by contrast, reach a verdict of unconstitutionality at a ratio of, say, 10% to 90%.

While the nature of the comparison is rarely made explicit in constitutional law, it is probably very seldomly made based on a comparison of the overall degree of realisation. As a rule, courts are more likely to compare the *change* that a government measure causes to the respective degrees of realisation. The comparison aims at the difference between the *status quo ante* and the degree of realisation after the state action. The incremental loss of freedom is contrasted with the gain in achieving the state's aims. Under[68] optimising approaches, if the change is in balance or at least not to the disadvantage of the affected fundamental right, it is considered proportionate. In illustrative percentage values: a 20% increase in the achievement of a state aim that is bought with a 30% loss

---

[65]Konrad Hesse, *Grundzüge des Verfassungsrechts der Bundesrepublik Deutschland* (1st edn, Müller 1967) 279 – translation by the author.

[66]Robert Alexy, 'On the Structure of Legal Principles' (2000) 13(3) Ratio Juris 294, 300; Robert Alexy, 'Formal principles: Some replies to critics' (2014) 12(3) Int J Const L 511, 512.

[67]Ralf Poscher, *Grundrechte als Abwehrrechte: Reflexive Regelung rechtlich geordneter Freiheit* (Mohr Siebeck 2003) 396.

[68]Sartor, (n 6) 1441.

of a fundamental right would be disproportionate according to optimising approaches. Under the less demanding approaches, which focus on severe disproportionalities, only state aims realised by, say, 5% more, tied to a 40% loss in freedom, would trigger a finding of disproportionality.

Optimisation and gross disproportionality approaches differ in their normative content and are the subject of dispute within constitutional law. Interpreting proportionality in the strict sense as imposing an optimisation requirement, in particular, has been criticised as too demanding.[69] As a substantive matter, critics complain that this approach leads to constitutional overdetermination of the political process. If proportionality in the strict sense requires optimum balance in each case, it will ultimately mandate one single solution for each conflict between fundamental rights and state interests. Legislative outcomes would thus be pre-determined by the constitutional balancing, which requires an optimal balance. Optimisation models would, consequently, neutralise any political leeway, and constitutional balancing would take the place of politics. Further, as an institutional matter, constitutional optimisation would lead to an imbalance in the relationship between (constitutional) courts and the legislature. The final decision on the proportionality in the strict sense would then lie with the courts. If, however, proportionality in the strict sense is held to always require an optimal balancing of fundamental rights and state aims, the legislature would be placed under the all-encompassing curatorship of the courts. The principles theory of fundamental rights, as coined by *Alexy*, attempts to compensate for these institutional distortions, in particular by introducing 'formal principles' intended to guarantee the legislature's decision-making leeway despite the optimisation requirement.[70]

The point of this article, however, is not to take sides in these normative debates on constitutional theory, but rather to make a point about the theoretical feasibility of the positions in this debate. Against this backdrop, it suffices to take note of the fact that the optimisation approach is not only doctrinally more demanding but also requires more precision in estimating relative degrees of realisation. Optimisation presupposes much more precise measures than approaches focusing on gross disproportionality, since even relatively crude assessments can reveal a gross disproportion in terms of degrees of realisation.

Measurement

It is also obvious that constitutional and human rights law have no instruments of measurement at their disposal that would allow for any kind of measurement that comes remotely close to those in decathlon. Again, most of the measurements based on the idiosyncratic, implicit, and intuitively – largely subconsciously – created scales to assess the severity of fundamental rights infringements and the achievement of a state interest are themselves implicit, and largely intuitively and subconsciously applied. This also comes at a price for the measurements. Accuracy and precision of measurements of proportionality in the strict sense will be very weak. They will at best

---

[69]Matthias Jestaedt, *Grundrechtsentfaltung im Gesetz: Studien zur Interdependenz von Grundrechtsdogmatik und Rechtsgewinnungstheorie* (Mohr Siebeck 1999) 239 ff; Poscher, *Abwehrrechte* (n 67) 82 f; Franz Reimer, *Verfassungsprinzipien: Ein Normtyp im Grundgesetz* (Duncker & Humblot 2001) 333; Poscher (n 67) 82–83.

[70]On formal principles, see Martin Borowski, 'Formelle Prinzipien und Gewichtsformel' in Matthias Klatt (ed), *Prinzipientheorie und Theorie der Abwägung* (Mohr Siebeck 2013), 151.

be crude and display a fairly widespread, even though it would be difficult to differentiate whether the spread is due to measurement or scaling issues.

The upshot

So, what are we left with for judgments on proportionality in the strict sense in law? While they are not impossible, they do mostly rely on judgments that are idiosyncratic, crude, intuitive, and largely implicit for all elements that underlie and enable comparisons of incommensurable properties. Hence, they will provide only very limited accuracy and precision. Does this mean that the model of degrees of realisation is therefore completely unsuitable for gauging the relationship between the severity of an infringement of fundamental rights and the degree of achievement of government aims? This conclusion seems too extreme.

Let us imagine a decathlon event for which the reference values have not been set by the IAAF and where the measurements would be of very coarse granularity – for example, the long jump would be measured only in whole metres and the 100-meter race in 5-second increments. Even then, assuming sufficiently large performance differences in different disciplines, it would be possible to say which performances were better than others. This is due to the fact that even in the absence of established standards, reference values and scales would not be arbitrary. Even if some similar performances in different disciplines might be deemed equal or unequal according to different plausible reference values, extreme differences in performance will be judged as unequal (and one performance better than another) according to all of the reference values under consideration. If there are sufficiently divergent results in the individual disciplines, various reference values will overlap insofar as one result appears better and the other worse in relation to every plausible reference value. If the world record in the 100-meter race is 9.58 s and that in the long jump is 9.02 metres, then a performance of 10.2 s in the 100-meter race will, according to all reasonable determinations of the reference value and even with only very crude measurements, be better in relative terms than a long jump of 4 metres.

The same can be expected of the idiosyncratic judgments on degrees of realisation when comparing the severity of rights infringements and the degree to which the state achieves its aims. Even if both the construction of the scales and the measurements are implicit, crude, and intuitive, they will not be completely arbitrary. Thus, some overlap of the idiosyncratic judgments is to be expected if the differences between the degree of infringement and achievement of a state interest are sufficiently extreme. Despite the inherent idiosyncrasies, there is always some intersection of different comparisons of fundamental rights and state aims for which all reasonable constructions and measurements will arrive at the same result. If the severity of the restriction of a fundamental right diverges very clearly from the degree of achievement and importance of a state's aim, this discrepancy will be apparent no matter what implicit idiosyncratic standard is used.

Optimisation models could try to capture this effect by significantly increasing the granularity of measurements of realisation.[71] *Alexy*'s Weight Formula, for example,

---

[71]In metaethics, this corresponds to approaches that attempt to reconstruct our everyday world comparisons using imprecise cardinality, Derek Parfit, *Reasons and Persons* (OUP 1986) 431; Chang, *Comparisons* (n 30) 32 f., 145. Ruth

seems to rely on such granularity in that it only has three grading levels. However, the latest version of the Weight Formula provides for four factors, each divided into three grading levels, for each item of consideration. The possible combinations quickly add up. 'If one takes the refined Weight Formula with four variables on each side, that is, with eight variables altogether, one acquires, on the basis of the two triadic scales, 3^8, that is, 6561 constellations. … It is an important point concerning the rationality of balancing that this diversity has a well-ordered, rather than a chaotic, character'.[72] It is almost comical that *Alexy* is proud to claim such levels of well-ordered diversity for his formula, given the dire preconditions for such fine-grained comparisons in fundamental rights law. This does not even take into account *Alexy*'s further postulate of a 'double triadic' Weight Formula based on nine gradations.[73] The double triadic formula assumes 6,561 (9^4) gradings for each object of consideration and can represent a whopping 9^8 – i.e., 43,046,721 – different outcomes. This is more than 40 times as fine-grained as the comparisons between the performances in two disciplines in decathlon, which are compared via two 1,000-point scales ($1,000^2 = 1,000,000$ constellations). Even beyond the already astronomical numbers of the double triadic weight formula is the potential reach of the 'extended formula', which allows several principles to be weighed against each other and contrasted on each side of the equation, thus expanding the formula additively.[74] It is exactly the extreme levels of differentiation that *Alexy*'s formulas presuppose that can be held against them as a theoretically feasible model for comparisons of incommensurables in law.

It should have become clear why even the refined Weight Formula presupposes a degree of accuracy in determining degrees of realisation which cannot be expected given comparisons of incommensurable properties that rely to a large extent on implicit idiosyncratic intuitive constructions of scales and measurements. Overlap will generally occur only in cases involving quite extreme differences. The idiosyncratic judgments on proportionality in the strict sense are likely to overlap only when a fundamental right is severely infringed *and* when this drastic infringement is only offset by a small advancement of the state aim. In cases of stark disproportionalities, the decision can be based on a proportionality judgement for which different idiosyncratic constructions and measurements will overlap.[75] Cases of the German Federal Constitutional Court that relied on such stark disproportionalities concerned a forced lumbar puncture, which carried significant long-term health risks, to determine culpability in a misdemeanour,[76] or evictions that were likely to cause the death of tenants,[77] and similarly extreme cases.[78]

Chang, 'Parity, Imprecise Comparability and the Repugnant Conclusion' (2016) 82(2) Theoria 182, on the other hand, argues that this does not capture the equality of alternatives, since equality has no degree to which imprecise cardinality can be tolerant. Even if this initially seems plausible as a conceptual argument, it is questionable whether it takes sufficient account of granularities. Two sprinters who both run 9.6 s in a 100-meter sprint are equally fast when their performance is measured in tenths of a second, although their performance measured in hundredths may show a considerable difference. Equality in tenths is tolerant of differences in hundredths.

[72]Robert Alexy, 'Proportionality and Rationality' in Vicki C Jackson and Mark Tushnet (eds), *Proportionality: New Frontiers, New Challenges* (CUP 2017) 18.

[73]Alexy, 'On Balancing' (n 29) 445.

[74]Robert Alexy, *Law's Ideal Dimension* (OUP 2021) 173 f.

[75]Cf Duff (n 9) 31, on 'retrospective proportionality', for which the preconditions are much more favourable: 'The most we can hope to do is to mobilise a rough principle of disproportionality: that sentences should not be (manifestly) disproportionate to the seriousness of the crimes for which they are imposed.'

[76]BVerfGE 16, 194 (202); see also BVerfGE 17, 108 (117–120).

[77]BVerfGE 52, 214 (221 f); BVerfG NJW 2019, 2995.

[78]For example, BVerfG NJW 1994, 1719 (1719 f.); BVerfG NZM 2005, 657 (658 f.)

Judgments in law are thus no different from our everyday comparisons of incommensurable alternatives. Deciding between an Italian and a French restaurant in the same price range will be difficult and, if groups are involved, be affected by various idiosyncrasies. But if the choice is between a bad Italian restaurant and a good French one, the decision is easy. A disanalogy between proportionality judgments in everyday life and in law might be seen in the need to support the latter with arguments, whereas the former are often made purely intuitively. However, since weights, scaling functions and measurements are also idiosyncratic in law, argumentation does not change the structure of proportionality judgments in either domain. As the efforts in decathlon or the Canadian court in the Hutterian case have shown, different outcomes can be argued on the basis different idiosyncratic assignment of weights, scaling functions or measurements. Argumentation cannot therefore save proportionality judgments in law from their inherent idiosyncrasies.

For *intra*subjective comparisons, *Ruth Chang* has thus argued against the trivalence of comparative outcomes (better – equal – worse) in evaluative comparisons of incommensurable objects, and for an additional category of 'parity'.[79] This proposal can also be applied to comparing degrees of realisation. In terms of degrees of realisation, otherwise incommensurable alternatives we face in everyday life can be compared via idiosyncratically constructed and measured metric properties. Beyond the intersection of plausible and permissible construction and measurements, the framework does not generate any intra- and intersubjectively compelling information about whether the alternatives are better, worse, or equivalent. The spectrum of plausible and permissible constructions and measurements of degrees of realisation can, as it were, explain the mechanics which lead to the phenomenon which *Chang* associates with the concept of 'parity'. Similarly, decision theorist *Wlodek Rabinowicz* reconstructs the idea of parity of preference judgments as the intersection of different preference orders, and derives from this a complex taxonomy of different preference relations.[80]

As for *Chang*'s concept of parity,[81] supervaluationist logic developed for semantic vagueness could provide the logical structure of our proportionality judgments in law. Just as judgments about vague predicates receive the truth value 'super-true' if they are true according to all justifiable precisifications of a concept, statements about disproportionality in the strict sense would be 'super-true' if they were true under every plausible intuitive construction and measurement of the incommensurable properties.

---

[79]Chang, 'Parity' (n 29); Chang, *Comparisons* (n 30) S. 123–41.

[80]Wlodek Rabinowicz, 'Value Relations' 2008 (74) Theoria 18, 37–45.; Both Rabinowicz' and Chang's ideas relate to the reconstruction of evaluative judgments. Chang even seems to see parity as a peculiarity of evaluative as opposed to non-evaluative judgments, Ruth Chang, *Making Comparisons Count* (Studies in Ethics, Routledge, Taylor and Francis 2002) 143–44 f. However, the example of the decathlon already shows that the measurement of athletic performance can also have a similar structure. The same is true even for geometric evaluations. The question of which figure is 'larger', for example, may be ambiguous with respect to the area being measured and its outer edges. Edge lengths and surface area are incommensurable. A rectangle of 3 cm length and 3 cm width has an area of 9 cm² and edge length of 6 cm; a rectangle of 10 cm length and 0.5 cm width has an area measurement of 5 cm² and edge length of 10.5 cm. Which of the two is 'larger'? We can use not only the area or the edge length, but also an infinite number of conceivable functions involving the edge and area measure. However, this indeterminacy does not preclude a rectangle 1 cm long and 1 cm wide from being clearly identified as smaller than the two examples above. Such geometric indeterminacies are even relevant in the law. For example, in calculating street-cleaning fees, municipalities have used a wide variety of combinations of street frontage and property area measurements. In German law, many have been declared unconstitutional because they led to disproportionate fees, for example HessVGH, DVBl 1986, 778 (778); OVG Greifswald, LKV 1996, 379 (379 f.).

[81]ibid 147–49.

Another upshot of the analysis of proportionality in the strict sense pertains to necessity. It becomes clear why we can be more confident in our necessity than in our judgments on proportionality in the strict sense. Even though there are also evaluative judgments involved there are three key differences that make necessity judgments much more intersubjectively reliable. Frist, they do only require ordinal comparisons, they do not require the creation of scales. Necessity only requires determining whether there is an ordinally less intrusive measures that achieves a purpose ordinally as well or better than the one in question. Second, necessity judgments do not require to compare incommensurables. They do not require to compare the intensity of intrusions with the achievement of a state purpose, but just the ordinal comparison of different intrusions regarding there intrusiveness and the comparison of different achievements of the same purpose. Third, they do not require to assign a comparative weight to intrusions and purposes, since they do not compare the two with each other at all. This does not imply that necessity judgments will always be uncontested. However, the analysis of explains why necessity judgments are relieved of many of the burdens that plague judgements on proportionality in the strict sense.

V. What actually goes on

Proportionality in the strict sense only goes so far in law. For theoretical reasons, the standard may be considered effective only in cases of very pronounced disproportionalities. Thus, for theoretical reasons, the purpose of comparisons of incommensurable fundamental rights and state interests is limited to uncovering gross disproportionalities between fundamental rights infringements and the effects of a legitimate state interest. Courts nevertheless decide many cases that do not involve such gross disproportionalities based on balancing. This raises the question of what courts are actually basing these cases on.

One answer would surely be that courts simply superimpose their ad hoc idiosyncratic intuitive proportionality judgments onto the positions of the parties involved or even the legislature whose laws they review. No doubt, there might be cases in which this is the best explanation. In many cases, however, something more benign, and more systematic, seems to be at play.[82]

In fundamental rights and human rights cases, courts face a special challenge. Under most fundamental or human rights regimes, the same law which guarantees a right also allows the state to limit the very same right.[83] This is true not only for the German Basic Law but also for international human rights guarantees such as the Charter of Fundamental Rights of the European Union (CFR) or the European Convention on Human Rights. Art. 52 (1) of the CFR confers on states a general power to limit the rights it spells out, provided the limits are imposed 'by law and respect the essence of those rights and freedoms'. Similarly, the European Convention of Human Rights permits

---

[82]For a more elaborate account, see Ralf Poscher, 'Proportionality and the Bindingness of Fundamental Rights' in Emmanouil Billis, Nandor Knust and Jon P Rui (eds), *Proportionality in Crime Control and Criminal Justice* (Hart Publishing 2021) 49 ff; Ralf Poscher, '§ 4 The Basic Law as a Constitution of Proportionate Balance' in Matthias Herdegen and others (eds), *Handbook of Constitutional Law: A Handbook in Transnational Perspective* (C.H. Beck 2024) paras 23 ff.

[83]On the legal-cultural differences between European and US-American fundamental rights guarantees in this regard Paul de Hert, Proportionality in Modern Regulatory States Confused about Priorities, in Jan Czarnocki and Przemyslaw Palka (eds), *Proportionality in EU Digital Law* (Hart 2024) (forthcoming).

infringements of many of its rights if this is considered 'necessary in a democratic society in the interests of public safety, for the protection of public order, health or morals, or for the protection of the rights and freedoms of others'. This arrangement of guarantees and limitations poses a structural problem for the protection of rights since it raises the question of how the state can be bound by fundamental rights if it has a more or less unrestricted general power to limit them.[84]

The German Federal Constitutional Court's initial response to this question was to introduce proportionality review into constitutional law.[85] This approach has since been taken up in other countries.[86] The proportionality requirement imposed some general limits on the powers of the state to infringe fundamental rights, giving those rights at least some bite. Proportionality has now been explicitly introduced into the Charter of Fundamental Rights of the European Union. Art. 52 (1) CFR requires all limitations of rights to be 'subject to the principle of proportionality'.

However, the 'suitability' and 'necessity' elements of proportionality review place few limits on state action, since they require only that state action be instrumentally rational.[87] And, as we have seen, proportionality in the strict sense can, for theoretical reasons, only come into play when there are gross discrepancies between the degree of infringement and degree of realisation of the state's interest. Even assuming these limits to restrictions of fundamental rights are actually implemented, they rule out only irrational and grossly disproportionate state actions. Accordingly, they still provide only a weak protection for fundamental rights. The level of protection especially against infringements by the legislator for many fundamental rights guarantees[88] still lacks behind their normative aspirations.

Moreover, the protection that proportionality can provide for fundamental rights is not specific to individual fundamental rights. Proportionality applies across the board to all fundamental rights without consideration of the specific nuances of the various freedoms to which it is applied. However, it is for a reason that fundamental rights codifications do not simply provide one single right to 'freedom'. Contemporary codifications guarantee a comprehensive catalogue of rights specific to different protected interests, drawing on distinct historical experiences of violations and substantial differences of context. If proportionality were the sole limit to infringements, these differences could not be fully captured by ruling out irrational and grossly disproportionate measures.

Although the introduction of proportionality into constitutional law constituted a preliminary general measure to render them more effective, given the residual structural

---

[84]In German constitutional law the resulting problem of bindingness of fundamental rights was already central in the discussion under the Constitution of the Weimar Republic from 1919, on the historical development of the issue Poscher, '§ 4 The Basic Law as a Constitution of Proportionate Balance' (n 82) para 23 ff; on the structural nature of the issue Bernhard Schlink, 'Proportionality (1)' in Michel Rosenfeld and András Sajó (eds), *The Oxford Handbook of Comparative Constitutional Law* (1st edn OUP 2012) 727–29.

[85]BVerfGE 7, 377 (403–413) (english translation: <https://law.utexas.edu/transnational/foreign-law-translations/german/case.php?id=657> accessed 19 September 2024).

[86]Barak (n 1) 179 f.

[87]In US discussions, scholars therefore questioned whether the Supreme Court even conducts proportionality reviews; according to this reading, the court's arguments with regard to suitability and necessity rather serve only to uncover hidden, inadmissible legislative motives, Iddo Porat, 'Mapping the American debate over balancing' in Grant Huscroft, Bradley W Miller and Grégoire Webber (eds), *Proportionality and the Rule of Law: Rights, Justification, Reasoning* (CUP 2016) 410 f.

[88]This does not hold for each and every fundamental rights guarantee. Some have more restrictive limitation clauses also regarding the legislator; some – especially formal procedural guarantees – are not open to limitation at all.

weaknesses of this innovation, the German Federal Constitutional Court has further reinforced the binding nature of fundamental rights by developing doctrines specific to certain individual fundamental rights and presenting these doctrines as derived from proportionality analysis in the strict sense. However, these doctrines are neither the result of a balancing of degrees of proportionate realisation nor do they lend themselves to such a balancing act. Rather, they are doctrines specific to certain fundamental rights, the purpose of which is to enhance their binding effect, particularly in relation to the legislature. These doctrinal innovations are specific responses to the same issue as the introduction of the general proportionality test, namely the challenge of conferring meaningful legal force upon fundamental rights with open-ended limitation clauses. However, they are not the result of balancing. The smoking gun can be seen by the fact that once these doctrinal innovations are established, they apply to all relevant cases no matter what weighting factors might be at play.

Some court-imposed procedural safeguards, for instance, apply to every infringement of a right irrespective of its intensity or the applicable state interest. One example is the Federal Constitutional Court's creation of a right to data protection in the early 1980s. The court fortified the right with a slew of procedural safeguards such as the right to be informed about the data stored, to have false data corrected or deleted, etc. The court presented these safeguards as mandated by proportionality in the strict sense.[89] However, these safeguards apply irrespective of the type of personal data stored – whether a mere work address or a sensitive medical file – and irrespective of the state aim – whether for internal administrative purposes or to prevent a terrorist attack.

At least when it comes to German law, we can also see court efforts to establish some 'scales' specific to individual fundamental rights or certain infringement contexts. As far back as in its Pharmacy Decision from 1958, the Federal Constitutional Court developed a three-tiered theory of occupational freedom within its balancing considerations. Hermeneutically the theory relies on the distinction between the choice and practice of an occupation alluded to in Art 12 GG, which guarantees the freedom of occupation on the one hand and a 'scale' ranging from public interests of outstanding importance, to important, to simple public interests, on the other. There is finer granularity when it comes to the 'scales' for infringements on the right to data protection and national security, distinguishing between up to six different degrees of probability of security risks on the national security side, and up to four levels of endangered public or private goods.[90] Such rudimentary 'scales', however, are not the product of measurements on pre-existing fundamental rights or public interest scales. They are attempts to stipulate such a scale. As for any such stipulation it holds what Wittgenstein has pointed out for the standard meter: 'There is one thing of which one can state neither that it is 1 m long, nor that it is not 1 m long, and that is the standard metre in Paris'.[91] Like the standard meter the scaling attempts in constitutional law rely on a stipulation prior to which no

---

[89]BVerfGE 65, 1 (44–52) (english translation: <https://www.bverfg.de/e/rs20080227_1bvr037007en.html> accessed 19 September 2024, para 147–164).

[90]For a latest elaboration on the different levels of risks and public goods with a tendency towards sliding scales BVerfG, 18.2.2023, 1 BvR 1547/19 and 1 BvR 2634/20, paras 103–108 (english translation: <https://www.bverfg.de/e/rs20230216_1bvr154719en.html> accessed 19 September 2024, paras 103–108).

[91]Ludwig Wittgenstein, *Philosophical Investigations* (3rd edn, Macmillian Publishing Co. Inc. 1958) § 50; on the discussion of this point and Saul Kripke's critique: Saul Kripke, *Naming and Necessity* (Harvard UP 1972), 54.

measurement is to be had. The scaling attempts are not derived from proportionality but make proportionality judgements possible.

In many judgments relying on proportionality in the strict sense, we thus see a re-entry[92] of the original hermeneutical challenge of how to give fundamental rights with their aspiration to effectiveness some bite despite the state's sweeping powers to limit them. The proportionality standard was intended to tackle this hermeneutical issue but has only been partially successful. As a general standard, it is still too weak. As a result, proportionality in the strict sense has become a launching pad for more specific standards intended to invigorate specific fundamental rights by curtailing the state's power to limit them. These doctrines are better understood as a hermeneutic doctrinal elaboration of these rights under the guise of proportionality in the strict sense than as attempts to balance degrees of realisation, which (for theoretical reasons) only succeeds when courts are confronted with gross disproportionalities. Their hermeneutic character is sometimes even quite obvious when the results of the purported weighing exercise, which is not hermeneutic, [93] can be explained much better by hermeneutic aspects. For example, the German Federal Constitutional Court required a warrant for online infiltration of personal computer systems by security agencies as a result of an alleged balancing of the right to privacy against national security.[94] However, the hermeneutic argument of drawing an analogy to house searches, which explicitly require a warrant under the German constitution, is the apparent real reason for this specific legal construction. Even though doctrinal innovations like these are stipulative and their content sometimes contestable, they can in principle be hermeneutically justified – just as the introduction of the proportionality into constitutional law itself – as legal constructions to accommodate the tension between the aspiration to effectiveness of fundamental rights and their sometimes almost unrestricted limitation clauses.

Disclosure statement

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

---

[92]Re-entry as the re-emergence of a form within a form in the sense of Niklas Luhmann, *Introduction to Systems Theory* (Polity Press 2013), 54 with reference to George Spencer-Brown, *Laws of Form* (E. P. Dutton 1979).

[93]On the difference between proportionality judgements and hermeneutics Mattias Kumm, 'The Idea of Socratic Contestation and the Right to Justification: The Point of Rights-Based Proportionality Review' (2010) 4(2) LEHR 141, 142 f.

[94]BVerfGE 120, 274 (331 f.) [2008] (english translation: <https://www.bverfg.de/e/rs20230216_1bvr154719en.html> accessed 19 September 2024, paras 258 ff).