



Lord Nicholls on the likelihood of crimes

Andreas Kapsner¹

Received: 11 October 2021 / Accepted: 30 September 2022 / Published online: 25 October 2022
© The Author(s) 2022

Abstract

In the decision to the British appeals case *Re H (Minors)*, Judge Lord Nicholls, talking about criminal behavior, stated that “that the more serious the allegation the less likely it is that the event occurred”. There is actually quite a bit of discussion about the conclusions that should be drawn from this observation in the literature, but I have not found much discussion of the question whether the observation is *right*. I find this surprising, and in this essay I want to inspect this question.

Keywords Likelihood of crimes · Crime statistics · Severity of crimes · Criminal and civil cases · Lord Nicholls · Standard of proof

1 Introduction

This essay deals with a purported correlation between the severity of crimes and their frequency that has been claimed to hold in an important British case. Roughly, the claimed correlation is that more serious crimes occur less frequently. As Britain is a common law country and, consequently, much weight is put on precedent and earlier judgments, it is important to be sure to fully understand such claims. However, given the broad nature of the claim, its interest goes far beyond common law jurisdictions.

It is in an appeals decision by the House of Lords, *Re H (Minors)*¹, that we find this remarkable opinion by Judge Lord Nicholls (p. 150):

When assessing the probabilities the court will have in mind as a factor, to whatever extent is appropriate in the particular case, that the more serious the allegation the less likely it is that the event occurred and, hence, the stronger should be the evidence before the court concludes that the allegation is established on the balance of probability. Fraud is usually less likely than negligence.

¹ [1996] 1 All ER 1.

✉ Andreas Kapsner
Andreas.Kapsner@lrz.uni-muenchen.de

¹ Munich Center for Mathematical Philosophy, Ludwig-Maximilians-University, Geschwister-Scholl-Platz 1, 80539 Munich, Germany

Deliberate physical injury is usually less likely than accidental physical injury. A stepfather is usually less likely to have repeatedly raped and had non-consensual oral sex with his under age stepdaughter than on some occasion to have lost his temper and slapped her. Built into the preponderance of probability standard is a generous degree of flexibility in respect of the seriousness of the allegation.

Though couched in several hedging terms, there is a clear statement here of an interesting relation between the serious nature of some types of behavior with their frequencies, which I will call **Nicholls' Thesis** and abbreviate as **NH**.

There is actually quite a bit of discussion about the conclusions that should be drawn from **NH** in the literature, in particular about certain questions concerning standards of proof in certain civil cases.² (Note that I will discuss these questions in the next section to give the appropriate context, but I will not be able to say much more substantive about them in the confines of this essay.)

In contrast to the amount of considerations of its impact, I have not found much discussion of the question whether Nicholls' observation is *right*, and even less discussion of *why* it might be thought to be true. One prominent place in which **NH** is rejected is a later ruling by the same court, written by Baroness Hale, to be discussed in Sect. 3. Beyond this and sources that cite Lady Hale's argument, I have only found two academic discussions where these questions are addressed (if briefly): (Redmayne, 1999) and (McBride, 2009). Interestingly, Redmayne and McBride come to diametrically opposed conclusions, which I will discuss in several of the sections below.

I will start, though, by outlining the legal background of Nicholl's decision, which has to do with questions of the appropriate standard of proof in certain difficult cases. I will not draw any conclusions about this larger issue, but it seems worthwhile to situate the discussion in this way, as it gives context not only to Nicholl's decision, but also to Lady Hale's critique of it.

2 Legal background: standards of proof

Nicholls' Hypothesis occurs in an opinion about *civil* cases that contain, as integral part of the case, allegations of *criminal* behavior. This might sound confusing, but the case from which the Nicholls Hypothesis stems, *Re H (Minors)*, is a good example that illustrates this general phenomenon.

A young girl had made allegations that her step-father had raped her. The ensuing criminal case, which was not directly the matter of the appeal, had been decided in his favor. In the ensuing civil case, which *was* the matter of the appeal, the local authorities had tried to take custody of the other children in the household, on the grounds that they were in danger of future harm from the father, given the older girl's testimony in the criminal trial.

Cases in which civil and criminal matters mingle in this way can become quite complicated on a theoretical level. In particular, it has been unclear to some judges

² See, e.g., Cobley (2006), Keating (1996), Keenan (1997), Bendall (2009), Roebuck (1996) and Hjalmarsson (2013).

and legal theorist which standard of proof should be applied to the criminal charges that are part of the larger civil case.

Usually, the only standard of proof in civil cases is the balance of probabilities, in simple terms: More probable than not. In criminal cases, the standard is higher: The evidence must convince *beyond reasonable doubt*. That standard had been judged not to have been met in the criminal case against the step-father. Now, the problem is this: Should the judge apply the criminal or the civil standard of proof to the question whether the step-father had raped the child, which was the question on which the care order hinged?

In earlier British cases, there is a certain strand of thinking that the appropriate standard in such cases might be a third kind of standard, intermediate between the criminal and the civil standard. The judge cited those opinions, and, arguably, took them to heart when deciding that he had no other choice than to leave the children in the father's care, even though he did not seem pleased by that outcome.

Part of the question the appellate court had to decide was whether the judge had applied the correct standard of proof, and it is in this context that we find the expression of Nicholls' Hypothesis quoted above.

The dialectics of the case are actually rather difficult to untangle (a task that would go beyond the scope of this paper), but in a nutshell, Nicholls holds that the balance of probabilities is the only standard that applies in civil cases. **NH** is meant to underwrite that claim. Nicholls is not presenting his hypothesis as some kind of uninterpreted statistical observation, but rather as an obvious platitude (if not obvious to anyone, then, at least, to any judge). That there is, as I mentioned, little discussion about the truth of the hypothesis suggests that many, indeed, felt it ringing true when reading it. Nicholls claims that the intuitive grasp of **NH** is what leads judges to mistakenly think that there might be a higher standard of proof that attaches to more serious crimes. Whether this strategy can succeed under the assumption that **NH** turns out to be true is an intriguing question, but in this paper I will be only focused on whether we should accept **NH** or reject it.

This is interesting quite irrespective of the original question about the standards of proof. For example, if true, it could guide friends of Bayesian approaches to legal proceedings to find better priors for their calculations [see Urbaniak and Di Bello (2021) for a general introduction to the topic, and Sect. 7.3 therein for the problem of finding appropriate priors].

About the further questions about the standards of proof, it suffices to say that Lord Nicholls' ruling has led to no shortage of confusion of a type that his fellow judge Lord Lloyd had already warned of in his ruling of the same case.

3 Lady Hale on NH

In a later ruling by the same court,³ Lady Hale was careful to be respectful of Lord Nicholls' arguments, deploring the confusion they gave rise to but crediting it to unso-

³ *Re B (Children)* [2008] UKHL35.

phisticated readings of these arguments rather than any inherent incoherence. However, in a short passage she explicitly addressed **NH** and seems to reject⁴ it:

As to the seriousness of the allegation, there is no logical or necessary connection between seriousness and probability. Some seriously harmful behaviour, such as murder, is sufficiently rare to be inherently improbable in most circumstances. (...) Other seriously harmful behaviour, such as alcohol or drug abuse, is regrettably all too common and not at all improbable.

This would be a splendid counter-example, if only it were one. But it isn't. Murder is more rare than alcohol abuse, and on all sensible measures of gravity of offense, it is also more serious, just like **NH** would have predicted.

The part I have elided does nothing to make this argument any better, and in fact, it thematically belongs to the next paragraph, which is why I cut it out. Here it is:

Even then there are circumstances, such as a body with its throat cut and no weapon to hand, where it [i.e., murder] is not at all improbable.

The next paragraph goes on:

Nor are serious allegations made in a vacuum. Consider the famous example of the animal seen in Regent's Park. If it is seen outside the zoo on a stretch of greensward regularly used for walking dogs, then of course it is more likely to be a dog than a lion. If it is seen in the zoo next to the lions' enclosure when the door is open, then it may well be more likely to be a lion than a dog.

What these thoughts seem to establish is that a preliminary assessment of the likelihood of a crime based on **NH** would have to be adjusted by further evidence. But surely, nothing Lord Nicholls' wrote disputes this.

Though Lady Hale's ruling has much to commend itself, this particular part of it seems unconvincing to me. **NH** withstands this attack and deserves a closer look.

4 Restating the hypothesis

Let me start by making an attempt at succinctly capturing Nicholls' hypothesis:

Nicholls' Hypothesis (NH) The more serious a crime is, the less frequently it occurs.

Some observations about my choice of words⁵ are in order: I write about crimes, where Nicholls only wrote "serious allegations". Allegations can, of course, also pertain to objectionable behavior that is not criminal. However, the larger context of the discussion (see Sect. 2) makes clear that criminal behavior is the main focus here. Should we in what follows find out that the hypothesis holds for criminal behavior,

⁴ What she rejects is actually a stronger version of **NH** that claims a necessary correlation, not one that is usually observed. However, her words seem to be directly directed at Lord Nicholls' statement.

⁵ They are not meant to be substantially different from McBride's summary, "(...) serious crimes are less common than minor crimes." (McBride, 2009, p. 340). However, I found it easier to spell out some upcoming augmentations with help of the phrase I chose.

then we can come back to an investigation into the more wide-ranging version of the hypothesis for a later date.

Even focusing on crimes, I need to concede that “crimes” and “allegations of crimes” are not the same, and their frequencies could differ considerably.⁶ It is true that Nicholls wrote about allegations; however, again, the rest of the quote makes clear that it is the frequency of the actual crimes that Nicholls is thinking about. The detour via “allegations” introduces further problematic questions that it is best to also set aside for another time.⁷

Further, note that I do away with all the hedging phrases (mainly “usually”). McBride writes about those that

... the qualification is of no account here, as it simply acknowledges the notorious ‘reference class’ problem: of course, deliberate injury is not ‘less likely’ than accidental injury in the boxing ring and so on. This just shows that it is always possible to restrict the reference class of a probability statement until the statement no longer holds true but this does not affect the truth of the statement as a general statement. (McBride, 2009, p. 327.)

And further:

Plainly, one can conceive of situations—e.g. a concentration camp, the Fall of Berlin—where murder or rape was more common than assault or indecent assault (...) (McBride, 2009, p. 340.)

I agree with him that Nicholls meant his hypothesis in a general sense, not restricted to too narrow scenarios, and applying to countries with a working jurisdiction. We can, as I suppose Nicholls did, think of modern day Britain as an example of what is intended. These things made explicit, we can erase the “usually”s from the hypothesis, as I did.

5 Two views on what a ‘serious’ crime might be

An important question⁸ we have to consider when we want to further scrutinize **NH** is what it means for a crime to be more serious than another, and how that might be measured.

There are two ways to understand the notion of seriousness in this context: As a moral judgment, or, more positivistically, as a measure of the response of a legal code and the legal practices based on it:

⁶ Indeed, one can think of situations where the two have little to do with each other at all. Imagine a legal system of a regime in which making allegations carries great personal risk of some sort, and that risk is getting more and more severe as the seriousness of the allegations increases. In such a situation, one would expect few serious allegations, no matter what the frequency of the corresponding crime might be.

⁷ Even if we grant that “a stepfather is usually less likely to have repeatedly raped and had non-consensual oral sex with his under age stepdaughter than on some occasion to have lost his temper and slapped her”, there is still room to doubt that an allegation of the former is less likely to be true than an allegation of the latter.

⁸ Important enough to warrant a dedicated section, as I am grateful to a reviewer for pointing out.

Moral Seriousness A crime is more serious than another if it is morally more reproachable.

Legal Seriousness A crime is more serious than another if a legal code punishes it more severely.

If we want to go beyond a common-sense notion in our analysis, then the first option seems the more problematic one: Whose moral judgments are we talking about? Experts or laypeople? Are we considering judgments of people of a particular jurisdiction (and at a particular point in time), or should we be more general? Going deeper into philosophical territory: What do these judgment track, if anything? If we give a moral realist answer here: What if most people are actually wrong about the seriousness of a given crime? And so on. I will not attempt to answer these questions, for the purposes of this essay I will go with examples that, I think, most readers will have a clear moral intuition about (such as in Sect. 3 above, where I assumed general agreement about murder being more serious than alcohol abuse).

The second kind of seriousness, the one that is reflected by law and legal practice, seems easier to get a conceptual handle on. But actually, to get to a useable comparison of seriousness is far from trivial. This problem will come up throughout what follows, and they will take specific forms that are related but not identical.

But which of the two notions should we focus on? Let me make a remark about their interaction to answer this:

Observe that when these two ideas of seriousness are in harmony with each other, we will be apt to applaud the given legal system for its sense of justice. Where they come apart, we will be critical of it. When innocuous behavior is punished, something is amiss. If reprehensible behavior routinely goes unpunished, then something went equally wrong. In especially dire circumstances, the two notions might even be diametrically opposed regarding some crimes, such as circumstances in which refusal to committ moral crimes is punished. These are the kinds of examples we have, at the end of the last section, agreed to set aside as outside of the scope of **NH**. But, as will become clearer below, one need to suppose a much greater harmony between the two notions of seriousness than just one that rules out evil regimes if one wants to argue for **NH**. Indeed, I think that so much harmony is needed for that argument that, should **NH** turn out to hold for one idea of seriousness, it is apt to hold for the other as well. I will explain my reasons for this in Sect. 6.2.

6 Why might **NH** be true?

However, before we get into the details of testing it out, let us ask why **NH** might strike someone as a reasonable hypothesis in the first place.

6.1 Empirical extrapolation

I noted above that Nicholls does not present his hypothesis as something that simply fell out of an examination of large data sets about criminal behavior. However, the one

defense of the truth of the hypothesis that I found in the literature holds that this is exactly how one should go about arguing for it.

[C]learly the concept of seriousness is different from that of improbability but all [Nicholls' Hypothesis] says is that it is a matter of judicial notice that serious crimes are less common than minor crimes. This is not logically but empirically true: to get some idea of the figures, we can note that the government's crime survey, combining both police and British Crime Survey (BCS) figures, states that in 2007–2008 the police recorded 17,000 'including 800 homicides)serious assault' (including 800 homicides) as against a total of 1,000,000 'all assaults' (Home Office, 2008, pp. 64–65). As far as sex offences are concerned, based on 2006–2007 BCS data, it appears that less than 1% of all adults reported having experienced a serious sexual assault as against 2% reporting less serious sexual assaults and all these offences constitute just under 1% of all recorded crime in 2007–2008. (McBride, 2009, p. 340)

Of course, if the hypothesis were true, then we should hope for some kind of empirical support for it. However, if the only reason to entertain the hypothesis were the bare crime statistics, then we have to address several problems here.

First, the few figures McBride cites are hardly enough to support any general claims. We would have to examine vastly greater data sets, and it is unclear that we have enough reliable information for that.

The literature that deals with the many problems of crime rate statistics is vast.⁹ These problems directly impact any attempt to supply an empirical base to **NH** and in that sense, I agree with Redmayne, who writes:

It is often suggested¹⁰ that serious allegations are harder to prove than less serious ones because serious wrongs are less likely to have occurred. Such a claim is difficult to defend; we simply do not know very much about the general prevalence of fraud, negligence, arson or child abuse, and this makes the claim that serious forms of behaviour are rarer than less serious ones problematic. (184)

To expand on that sentiment: It is not only the amount of data that is the problem, it is also the kind of data we have access to. As I discussed above, Nicholls is best understood to be talking about the frequencies of crimes rather than the frequency of crimes being reported to the police. It's plausible to assume that these numbers correlate roughly, but also that they are never exactly the same. What is more, the gap between occurrences and reports might be consistently greater for some forms of crimes than for others. Maybe some crimes are unlikely to be reported, while others might invite a greater number of false allegations.

We can, of course, try to make educated guesses about the respective sizes of such deviations, but (a) we might be quite wrong about those (cf. the #MeToo movement,

⁹ See, for a small sample of critiques from a variety of perspectives: Brownstein (2000), Hope (2005), Hickman and Rice (2010), Skogan (1974), Center and Smith (1972).

¹⁰ Note that he is directly discussing Lord Nicholls's *In Re H* ruling in this quote, notwithstanding the talk about this being "often suggested". I have not seen this suggested often before Lord Nicholls.

which made plausible claims of certain crimes being much more wide-spread than previously expected). And (b), by making such guesses, we would leave the realm of the purely empirical extrapolation, and thus of the idea we are examining in this subsection.

What else might we turn to, if we find the number of crime reports to the police wanting? We might look at conviction rates instead. That might, at least, help to screen off some false allegations. However, that the number of convictions tracks actual crimes more closely than allegations is surely contentious (It might, indeed, even be contentious that is *should*. The higher standard of proof in criminal cases is meant to make conviction of the innocent highly unlikely, at the expense of a higher likelihood of letting the guilty go free).¹¹

Again, we could try our luck with conviction rates modulated by speculations about the number of times the law has gone awry in either direction, but that also does not seem like a promising path, for much the same reasons as the ones given above.

If anything, then, conviction rates seem even less suited than frequencies of crime reports, but it seems unclear what other empirical data we might rely on. What we would be better off with would be some intuitively plausible explanation of the suggested correlation between crime severity and frequency.

6.2 Intuitively plausible explanations

Let us start here, then: Is there any hard-wired connection between the rate of occurrence of some behavior and its being judged more or less serious? Well, fortunately, we don't live in a society that is so perversely conformist that it punishes eccentric and uncommon behavior for the sole reason of its infrequency and, moreover, finds such behavior more seriously offensive the more seldom it is observed. The possible explanation we are looking for should, therefore, be slightly more indirect.

Here is another line of thought: It might simply be harder to commit a more serious crime, and failed attempts may land on a lower level of severity. Given extreme circumstances, I might intend to kill a grown man with my bare hands, but I would probably fail. It is conceivable, though, that I might hurt him in the process, provided he is in equally bad shape as I am.

However, across different crime types, this is surely not a general pattern. Anyone who knows how to operate a car is physically capable of murder, while successful art forgery is a feat that is beyond most.

To spare you more non-starters of that nature: In thinking about the matter, I have only found two candidate explanations that seem plausible and have the required generality to potentially support Nicholls' wide-ranging claim:

Moral Compass People's inner moral compass tends to prevent them from doing wrong, and kicks in with more and more strength as crimes get worse and worse.

¹¹ There is another problem further downstream: In so far as Nicholls' Hypothesis is seen as something that guides judicial decision making, then if it was supported by conviction rate data alone, we would be trying to calibrate the epistemic engine of our legal system using its own output. The problematic nature of such a circular process are obvious, though maybe such problems will be found to be inevitable in some form or another, as they can be seen as a quintessential pattern of epistemological predicaments.

Deterrence The more serious a crime, the harder the punishment. This punishment may be of a legal or a social nature, but in the present context the legal punishment is more salient. So, if criminal justice has some deterrent effect, one should expect it to deter more serious crimes than minor ones.

These two mechanisms operate on the two different ideas of seriousness discussed in Sect. 5. And as I believe the effects of both are real (to what exact extent is another question, see below), I think that in circumstances where they work against each other, **NH** has little chance to be found to hold. This requires a fair amount of harmony and mutual reinforcements between **Moral Compass** and **Deterrence**, and this is the reason that above I have speculated that **NH**, should it hold for one sense of seriousness of crime, will hold for the other, as well.

The two mechanisms have different strengths and weaknesses as potential explanations. **Deterrence** is somewhat more objective than **Moral Compass**, at least given a single legal code. What punishment is to be expected will be consistent across individuals, how morally wrong it strikes them not. Moreover, in fully explaining what a moral compass is, what it point towards and how it might be disturbed will involve giving answers to the difficult questions I mentioned in Sect. 5.

However, the idea of deterrence has its own problems. There is an extensive literature that deals with the question whether criminal laws and punishments indeed have the deterrent effect lawmakers hope.¹² There are situations in which criminalizing certain behavior (say, drug use) might lead to more crime rather than less. That said, on the whole, I think I agree with the following quote:

The expectation of official action clearly does not deter all offending, but it equally clearly deters a lot of offending, and when it is altered or removed, people's behavior changes. (Kennedy, 2012, p. 9)

This acknowledges that deterrence works, but it surely it doesn't work perfectly. This is, in itself, not problematic for our purposes. However, there is reason to think that deterrence doesn't work uniformly well or badly across crime types and situations. For example, it as a more pronounced effect on premeditated crime than on impulsive acts. The problem for deterrence as a (partial) explanation of **NH** is that these kinds of differences might be systematic enough to deter people more from certain crimes of lesser severity than from more serious ones.

Moreover, there are crimes that are harder to prove than others and therefore more frequent. Physical sexual abuse is more serious than exhibitionism in well-frequented public places, but it is probably also more frequent, in part because it is harder to be found guilty of it.

These kinds of problems might lead to disturbances great enough to undermine **NH**. I will get back to this point towards the end of this paper.

¹² The essays collected in Nagin et al. (2018) give a good sense of the current state of this discussion

7 Is the hypothesis true?

Now that we have some notion for why the hypothesis *might* be thought to be true, let us ask the crucial question: *Is it true?*

It will turn out that the answer, as far as I can see, depends on which of several more detailed versions of the hypothesis we are dealing with.

That these versions exist is due to a somewhat hidden ambiguity in my formulation of NH, to repeat:

Nicholls' Hypothesis The more serious a crime is, the less frequent it occurs.

The ambiguity lies in the word “crime”, which can either be read as “a particular *kind* of crime” or “an individual criminal *act*”.

That ambiguity, of course, needs to be resolved, and I will do this in the following subsections. I should, however, first give my reasons for phrasing **NH** in the way I did, for of course I have built this ambiguity into it deliberately. My reason is simply that I find the same ambiguity in Nicholls' opinion and in its interpretation by others. Phrasing NH as I did gives me now the chance to investigate the different readings in order to see whether one of them is correct.

7.1 NH crime types

Let us start with the one that Redmayne attacks:

NH Crime types The more serious a *kind* of offense it is, the less frequently crimes of that type occur.

First, some remarks:

This way of phrasing the hypothesis presupposes that we have a suitable way of individuating crimes. If we are allowed to play fast and loose with this, then it will be easy enough to come up with counter-examples to the hypothesis. Being threatened with a gun is more serious than being threatened with a stuffed peacock, but it is also more common. If the two are considered different crimes, then we have an all too cheap counter example already.

Individuating crime types, of course, is one of the jobs of a penal code. Though there may be difficult cases, I will proceed by assuming that a satisfactory partition is available that way. What is worth noting, though, is that the hypothesis then becomes relative to a given legal code. Inasmuch as such codes differ from one country to the next, we might get further complexities to sort through, but I don't expect the problems here to be unsurmountable.

What is more problematic than the mere categorization of crimes is the problem mentioned in Sect. 5, the needed ranking of crime types by seriousness or severity. How to set up such indices is far from clear. Attempts have been made to base them on maximum penalties set by parliaments, actual sentencing practice in the courts, and public and expert opinion (MacKinnell et al., 2010, p. 2). Giving these parameters

different weights will lead to different orderings. MacKinnell et al. (2010) compares some options, and applies them to Australian crime statistics.¹³

But let us assume that we have somehow agreed on which order of seriousness to use, and that it at least roughly corresponds to our intuitive sense of the matter. Should we then expect Nicholl's Hypothesis to be true? Redmayne thinks not:

Even if there is a general correlation between gravity and frequency of wrongdoing, it beggars belief to presume that the correlation always holds. As Robertson has observed, there may be some minor regulatory crimes which are in fact very rare. (Redmayne, 1999, p. 184)

In other words, Redmayne thinks that is not hard to give counter examples of the kind I discussed in Sect. 3 given by Lady Hale. And, even though I argued that her counter example fails, I actually agree with Redmayne. I have not been able to access the paper by Robertson that is cited by Redmayne, but it is not too hard to think of examples that better illustrate the problem. For example, it is illegal to enter the British parliament in a suit of armour.¹⁴ This is surely a very uncommon crime. At one point in time, it would probably have been seen as a very serious offense, but nowadays I have trouble believing that it would be considered more than an act of extreme mischief (unless the would-be knight were to make use of his medieval weapons). At the very least, I cannot imagine that it would be punished more severely than murder, which is much more frequent.

One might object that it is unfair to pick such an unlikely crime as example. But of course, that would come close to asking to only to test the truth of the hypothesis in cases where it will come out true. Infrequent but not-so-serious crimes are the domain in which we will find many counter examples. To give one that is a little more up to date:

In the UK, criminal defamation (libel) was abolished by the Coroners and Justice Act 2009. Before that, very few criminal defamation cases were brought before the courts [see McMahon (2001)]. But does that mean that it was among the most heinous of crimes?

A reviewer of this essay wonders whether we should take such counter-examples all too seriously:

Is the **NH** Crime Types to be understood as admitting no, or few exceptions, or should it be understood as making a broad generalization that may admit many counter-examples, but still be sufficiently true to carry important theoretical and moral weight?

This might be plausible (depending on what exactly the theoretical and moral weight is taken to be, i.e. what one wishes to argue for using **NH**) if the many counter-examples of infrequent but also not-too-serious crimes add up to only a insignificant fraction of

¹³ For the United Kingdom, there is a "Crime Severity Score" that is filed under "Experimental Statistics". Without going into details, it strikes me as at a pretty experimental stage, indeed. (<https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/datasets/crimeseverityscoreexperimentalstatistics>).

¹⁴ see https://www.lawcom.gov.uk/app/uploads/2015/03/Legal_Oddities.pdf.

the overall crime number. Maybe one can find a cut-off point of infrequency above which the hypothesis holds. This, as far as I can see, is however not the case.

At least in the above mentioned Australian study¹⁵ that compares different ways of setting up a severity order of crimes, on none of them a strong correlation between severity and frequency is observable, and I don't find this surprising. Moreover, the missing pattern does not emerge if we consider only crimes above a certain frequency-threshold.

7.2 NH individual crimes

Let us move on to reading the hypothesis in a way that might work better. Instead of thinking of kinds of crimes, we might read "crimes" as single criminal acts, and correspondingly rephrase **NH** thus:

NH Individual Crimes The more serious a single criminal act is, the less likely it is to occur.

To arrive at an uncontentious measure of the severity of single acts is, of course, an even taller order than to do the same for crime kinds if we look at the notion as specified by laws and legal practice. For no legal code specifies exactly how serious a particular criminal act is. The only way this could be the case was if all instances of a given type crime would receive exactly the same punishment. But this is not the case in any actual legal systems I am aware of. Instead, legal systems acknowledge the fact that there are more and less serious instances of the same kind of crime, and usually specify a range of possible punishments. Here, the role of the judges, and their own particular sense of seriousness, comes in. Inasmuch as this is a moral judgement itself, it isn't even possible to avoid reference to such judgments here. On the upside, it might actually be somewhat easier to agree on a sense of moral seriousness of single criminal acts and to compare pairs of them, as opposed to whole crime types.

In any case, suppose a measurement and comparison of seriousness can be achieved, what can be said about instances of minor, but infrequent crimes?

First, we need to assure that the ordering of serious is not so fine grained that these are assigned a measure of seriousness that is distinct from all other criminal acts. For if we did that, then only few instances of crimes with that measure of severity would be observed, which goes against **NH** and the assumption that we are dealing with minor crimes.

So, suppose that we have a coarse enough measure of severity so that exemplars of infrequent minor crimes are "swallowed up" by roughly equally serious more common criminal activities. The uncommon occurrences of criminal libel, for example, might end up at the same level of severity as some acts of theft (though other acts of theft might be more or less serious).

Given all that, would **NH** be true?

Well, I don't want to claim that it is impossible to adjust the parameters that are only loosely specified in this thought experiment in a way that **NH** comes out true. But I have the feeling that, even if such a specification is possible, there will be many others

¹⁵ MacKinnell et al. (2010).

that are equally plausible for which **NH** does not hold. Consider the first specification that is coarse grained enough to make the hypothesis true, and compare it with its slightly more precise predecessor. What makes the former more plausible than the latter (other than that it makes the hypothesis true)?

That the hypothesis in this form can support much theoretical weight seems very implausible to me.

7.3 NH individual crimes of a given type

Lastly, let me propose a kind of synthesis of the two readings above:

NH Individual Crimes of a Given Type Given a particular kind of crime, the more severe an instance of it, the less likely it is.

Now we are getting to something that strikes me as at least *prima facie* plausible. Given the crime of theft, I would expect more petty instances of it to be more common than more serious ones. Moreover, I think **deterrence** and **Moral Compass** supply plausible explanations for this correlation. To give an example, in some states, speeding beyond a certain threshold is considered a crime. Assuming that threshold is more often surpassed only narrowly than by a lot, we could explain this by the driver's sense of putting others in danger or (maybe more likely) by worries about legal repercussions.

This is a version of the hypothesis that seems worth entertaining and testing out, though there may of course be problems that will lead us to conclude that it does not hold in all cases. Amongst others, the problems I pointed out in Sect. 6.2 might distort the picture so much that the hypothesis fails. Suppose there are certain types of white collar crimes that are serious forms of theft but that are not usually punished. In that case I am not sure that they might not be more prevalent than minor, but often punished, instances of theft.

Or consider that of many types of crime there will be premeditated and impulsive instances, and we've discussed above that this might skew the likelihood order so much that **NH** will be seen to fail.

And of course, the problems of comparing the seriousness of criminal acts remain. However, it seems more plausible that we can actually get somewhere here, where we only have to compare similar kinds of behavior.

All that said, if there is any hope for **NH**, then I think this version is its best bet.¹⁶ To judge whether that bet is lost or won requires decidedly more research.

8 Conclusion

We have seen three versions of Nicholls' Hypothesis: The first relates to crime types and is surely false. The second relates to individual criminal acts, and I have voiced doubts as to whether it can be spelled out in a way that is very convincing. Lastly, I considered the hypothesis to be about individual criminal acts within a given crime

¹⁶ That's a hunch. A reviewer writes: "The author seems to acknowledge that some of the same objections against earlier formulations of the NH still apply to this version of the hypothesis, but does not explain why he thinks those objections would be less damning as to the third formulation." Maybe they aren't.

type. This is the version that has at least a chance of being true, and I have presented two mechanisms that might explain its truth, **Deterrence** and **Moral Compass**.

However, this version is also theoretically not the most exciting one, and I leave it as an open question whether many interesting arguments can be based on it. In particular, I leave an examination of the question whether what is left of NH after restricting it to individual acts in a particular crime type is enough for the dialectical purposes of either Lord Nicholls or others who have quoted him for future research, though I allow myself the remark that I have serious doubts that it is.

As a general summary, I think it clear that anyone who would like to base any kind of substantial argument on NH will have a lot of work to do for this argument to turn out convincing.

Acknowledgements I would like to thank the two anonymous reviewers for their careful reading and their critiques and comments, which certainly helped to improve the essay. My research has been supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project 436508789.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bendall, C. (2009). The demise of the enhanced standard of proof in child protection cases: re b [2008] ukhl 35: Cases. *Journal of Social Welfare and Family Law*, 31(2), 185–191.
- Brownstein, H. H. (2000). The social production of crime statistics. *Justice Research and Policy*, 2(2), 73–89.
- Center, L. J., & Smith, T. G. (1972). Crime-statistics-can they be trusted. *American Criminal Law Review*, 11, 1045.
- Cobley, C. (2006). The quest for truth: Substantiating allegations of physical abuse in criminal prosecutions and care proceedings. *International Journal of Law, Policy and the Family*, 20(3), 317–343.
- Hickman, M. J., & Rice, S. K. (2010). Digital analysis of crime statistics: Does crime conform to Benford's law? *Journal of Quantitative Criminology*, 26(3), 333–349.
- Hjalmarsson, J. (2013). The standard of proof in civil cases: An insurance fraud perspective. *The International Journal of Evidence & Proof*, 17(1), 47–73.
- Hope, T. (2005). *What do crime statistics tell us*. Oxford University Press.
- Keating, H. (1996). Shifting standards in the house of lords re h and others (minors)(sexual abuse: Standard of proof). *Child & Fam. LQ*, 8, 157.
- Keenan, C. (1997). Finding that a child is at risk from sexual abuse: Re h (minors)(sexual abuse: Standard of proof). *The Modern Law Review*, 60, 857.
- Kennedy, D. M. (2012). *Deterrence and crime prevention: Reconsidering the prospect of sanction*. Routledge.
- MacKinnell, I., Poletti, P., & Holmes, M. (2010). Measuring offence seriousness. *Crime and Justice Bulletin*, 142, 1–12.
- McBride, E. (2009). Is the civil 'higher standard of proof' a coherent concept? *Law, Probability and Risk*, 8(4), 323–351.

- McMahon, M. (2001). Defamation claims in Europe: A survey of the legal armory. *Communications Lawyer*, 19, 24.
- Nagin, D. S., Cullen, F. T., & Jonson, C. L. (2018). *Deterrence, choice, and crime: Contemporary perspectives* (Vol. 23). Routledge.
- Redmayne, M. (1999). Standards of proof in civil litigation. *The Modern Law Review*, 62, 167.
- Roebuck, J. (1996). Allegations of child sexual abuse: Should 'unacceptable risk' be the only criterion for refusal of contact?: An evaluation of reh and ors (minors)(sexual abuse: Standard of proof)[1996] 1 all er 1. *James Cook University Law Review*, 3, 116–136.
- Skogan, W. G. (1974). The validity of official crime statistics: An empirical investigation. *Social Science Quarterly*, 25–38.
- Urbaniak, R., & Di Bello, M. (2021). Legal Probabilism. In Zalta, E. N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2021 edition.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.