Individualisation at Sentencing: The Effects of Guidelines and "Preferred" Numbers

Julian V. Roberts
University of Oxford

Jose Pina-Sanchez
University of Leeds

Ian Marder[*]
University of Leeds

Psychology and law; Sentencing; Sentencing guidelines

Over a decade after the introduction of sentencing guidelines in England and Wales, little is known about their effects on consistency and individualisation. Limited research has addressed the issue of consistency, and no research has explored another key concept, namely individualisation. This is regrettable since one criticism of guidelines is that they undermine the principle of individualisation at sentencing, and this critique is examined here. The article explores two potential threats to individualisation, using sentence length data from the Crown Court Sentencing Survey. One threat may arise if a guideline constrains judges to sentence within a restricted range, leading to a less individualised approach to sentencing. The second is more fundamental, and consists of the tendency to favour some sentence lengths over others—a preference for certain "round" numbers. The article reports an analysis of custodial sentences for assault offences. Results indicate that sentence lengths for assault offences are affected by a preference for certain numbers—a tendency first observed by Francis Galton in the 19th century. On a more positive note, we find no evidence that the sentencing guidelines for assault and burglary introduced in 2011 have diminished the degree of individualisation in sentencing. We also find that courts report taking a larger number of sentencing factors into account under the new guidelines, further evidence that the guidelines have not undermined individualisation.

Sentencing guidelines promote more consistent and principled sentencing. Promoting greater consistency may come at a cost, however. Guidelines which are excessively restrictive will achieve consistency, but at the expense of individualisation. To date, little research has explored the effects of the Sentencing

[*] Our thanks to Andrew Ashworth, Rory Kelly, Keir Irwin Rogers, Oren Gazal-Ayal, and the journal's reviewers for comments on an earlier draft. Views expressed herein are solely those of the authors.

Council's guidelines on consistency or individualisation. The Council routinely publishes projections of the impact of its guidelines on the prison estate.[1] To date it has yet to conduct research on the question of whether the guidelines have enhanced consistency. It is also important to address critics'[2] concerns that the new guidelines format may impair sentencers' ability to distinguish between cases—to *individualise* the sentence.[3]

Academic studies have begun to fill the research void, although a systematic examination of the effects of the guidelines on consistency remains to be conducted. Three studies have examined the effects of the Council's first guidelines (covering assault and burglary). Pina-Sánchez analysed trends before and after the introduction of the assault guideline and found that "consistency improved in all the offences studied after the new guideline came into force."[4] Other studies have explored consistency in the application of the guidelines, and demonstrated that for several high-volume offences, the guideline factors were being applied in a consistent manner across courts.[5] While limited in scope, these studies suggest that the English guidelines have had a positive effect on promoting a more consistent approach to sentencing.

Consistent outcomes alone are insufficient to achieve justice in sentencing. A mandatory sentencing law which imposed a five-year sentence on all convictions, regardless of the relative seriousness of individual cases, would ensure consistency but would lack a second essential element: *individualisation*. In light of the importance of both consistency and individualisation, it is surprising that almost all the empirical research on sentencing has explored consistency or sentencing disparity while neglecting individualisation in sentencing. The present research explored individualisation in the Crown Court. More specifically, we addressed two potential threats to individualisation, one arising from the implementation of guidelines, and another—of a more fundamental nature—which relates to the psychology of decision-making.

[1] See *https://www.sentencingcouncil.org.uk/* [Accessed 29 November 2017].

[2] In a recent work, Graeme Browne asserts that "The current system of English sentencing guidelines … seeks consistency at the expense of individualised justice" (p.150). G. Browne, *Criminal Sentencing as Practical Wisdom* (Oxford: Hart Publications, 2017). Elsewhere, John Cooper argues that the guidelines undermine personal mitigation, a key component of individualised sentencing: J. Cooper, "Nothing Personal" in A. Ashworth and J.V. Roberts (eds), *Sentencing Guidelines: Exploring the English Model* (Oxford: Oxford University Press, 2013), p.158. See also testimony at House of Commons Justice Committee, Tuesday 14 December at Q.48.

[3] At the conceptual level, consistency and individualisation are not incompatible. Consistency may be defined as applying a common set of criteria fairly across cases, and individualisation requires a court to reflect legally-relevant factors associated with individual offenders. In practice, however, the pursuit of consistency of outcome may impair a court's ability to take account of individual cases characteristics.

[4] J. Pina-Sánchez, "Defining and Measuring Consistency in Sentencing" in J.V. Roberts (ed.) *Exploring Sentencing Practice in England and Wales* (London: Palgrave Macmillan, 2015), p.87.

[5] J. Pina-Sánchez and R. Linacre, "Sentence Consistency in England and Wales: Evidence from the Crown court Sentencing Survey" (2013) 53 *British Journal of Criminology* 1118; J. Pina-Sánchez and R. Linacre, "Enhancing Consistency in Sentencing: Exploring the Effects of Guidelines in England and Wales" (2014) 30 *Journal of Quantitative Criminology* 731. Irwin-Rogers and Perry explored sentencing for burglary and their analyses "provided a strong indication that the courts were sentencing in a manner that was consistent with the domestic burglary guideline". K. Irwin-Rogers and T. Perry, "Exploring the Impact of Sentencing Factors on Sentencing Domestic Burglary" in Roberts (ed.) *Exploring Sentencing Practice in England and Wales* (2015).

Threats to individualisation: (a) Sentencing Guidelines and (b) "Penal Clustering"

Tensions may arise when a system attempts to maximise both consistency and individualisation. For example, the grid-based sentencing guidelines found in many US jurisdictions (including Minnesota and Pennsylvania), employ a two-dimensional structure. Crime seriousness constitutes one dimension and criminal history the other; the grid provides a range of sentence for each cell of the grid. Thus, under the Minnesota standard grid, a conviction for a seriousness level six offence for an offender in criminal history category three carries a sentence range of 34–46 months.[6] Courts must impose a sentence within this range, or find "substantial and compelling" reasons to justify a departure from the guidelines. Since a single grid is used, all offences[7] must be assigned to one of only eleven offence seriousness levels. Given that there are many hundreds of criminal offenses in Minnesota, crimes of variable seriousness will be assigned the same level of gravity in the grid. This arrangement enhances the consistency of outcomes, yet a loss of individualisation may result.

(a) The English sentencing guidelines

Sentencing guidelines in England and Wales take a different approach. The guidelines require courts to follow a clearly-articulated methodology which allows sentencers considerable discretion within that structure, as well as the power to depart from the guidelines when it would be in the interests of justice to do so. Definitive sentencing guidelines were first introduced in England and Wales in 2004.[8] Since then, offence-specific guidelines have been issued for the most common crimes. As with the US schemes, the goal of these guidelines is to promote a more consistent approach to sentencing. The guidelines provide more restrictive sentence ranges than those established by statutory maxima.[9] For example, although the statutory maximum for assault occasioning actual bodily harm is five years' imprisonment,[10] the ceiling of the guideline range is only three years. Assuming that courts sentence within the guideline range, the distribution of sentence lengths is restricted—unless a court elects to depart from the guideline by invoking the interests of justice provision.[11] Under such a guideline, judges may be less likely to use the full range of their discretion.

The English guidelines do not assign crimes to a small number of seriousness levels (as is the case in Minnesota); each offence has its own separate guideline. In 2011, the Sentencing Council issued a new-format guideline.[12] The previous

[6] Minnesota Sentencing Guidelines Commission, *Minnesota Sentencing Guidelines and Commentary* (Minneapolis: Minnesota Sentencing Guidelines Commission, 2015).

[7] There is a separate sentencing grid for sexual offences.

[8] In December 2004, the Sentencing Guidelines Council issued three guidelines, covering Seriousness, Reductions in Sentence for a Guilty Plea and New Sentences. See M. Wasik, "Sentencing — The Last Ten Years" [2014] Crim. L.R. 477; and A. Ashworth and J.V. Roberts, "The Origins and Evolution of Sentencing Guidelines in England and Wales" in Ashworth and Roberts (eds.) *Sentencing Guidelines: Exploring the English Model* (2013).

[9] Sentencing Council of England and Wales, *Assault Offences Guideline: Professional Consultation* (London: Sentencing Council of England and Wales, 2010) p.4, *https://www.sentencingcouncil.org.uk/wp-content/uploads /ASSAULT_Professional_web.pdf* [Accessed 29 November 2017].

[10] Offences Against the Person Act 1861 s.47.

[11] Coroners and Justice Act 2009 s.125(1)(b).

[12] See J.V. Roberts, J.V. and A. Rafferty, "Sentencing Guidelines in England and Wales: Exploring the new Format" [2011] Crim. L.R. 680.

assault guideline focused on forms of specific conduct and provided few sentencing factors. The Council's new format guideline was designed to provide greater guidance, principally be moving away from specific assaultive acts and by providing a clear, step-by-step methodology. Key to this methodology is a two-step approach to considering factors relevant to harm, culpability and more generally, aggravating and mitigating factors. By specifying the factors to be considered, and by restricting the range of sentence, the guidelines aim to enhance consistency across courts, but might the English guidelines have resulted in less individualisation? For several assault offences[13] the new format guideline provides fewer categories (three rather than four), each with its own starting point and ranges. Under the new format, individualisation may have suffered as courts assigned the same number of cases to fewer categories.

Individualisation at sentencing requires a court to differentiate among cases to reflect important individual characteristics relating to the offence and the offender. The introduction of the new assault offences guideline[14] may have had one of two effects. By imposing tighter sentence ranges and by requiring cases to be assigned to one of three categories, the guideline may have resulted in *less* individualised outcomes—through the mechanism noted earlier (cases of variable seriousness being "squeezed" into the same guideline category and hence the same range of sentence). On the other hand, by enumerating and highlighting important factors relating to the case (e.g. harm, offender culpability, personal mitigation) the guideline may have promoted greater differentiation among cases, thereby increasing individualisation. More or less individualisation? That is the question, and the answer is fundamental to the success and future of the guidelines.

The Council's new guideline format has been adapted for subsequent guidelines,[15] so findings from the first two guidelines will carry lessons for other offence-specific guidelines which are subject to a similar structure. In addition, since the US grid-based approach to guidelines has been rejected in all other countries, the English guidelines represent the principal model for countries seeking to introduce greater structure to their sentencing.[16] These findings will therefore have relevance for other jurisdictions which have indicated an intention to adopt sentencing guidelines.[17]

(b) Preferred sentences, or penal clustering

The second potential threat to individualised sentencing concerns the phenomenology of decision-making. If sentencing were determined by a computer

[13] For example, the previous (Sentencing Guidelines Council) guideline for s.20 assaults had four guideline categories while the newer Sentencing Council guideline contains only three categories.

[14] Unfortunately, comparable data are not available for the sentencing of assault before the introduction of the first (SGC) assault offences guideline. Had such data been available it would be possible to draw conclusions about the effects of guidelines more generally, rather than the effects of a change in the guideline format.

[15] In its more recent guidelines the Council has modified this structure, although the essential architecture which requires courts to proceed through a sequence of steps is common to all the guidelines.

[16] The New Zealand Law Commission created a set of guidelines but these are not publicly available as the government has yet to approve the necessary implementation legislation: see W. Young and C. Browning, "The Origins and Evolution of Sentencing Guidelines: A Comparison of England and Wales and New Zealand" in Ashworth and Roberts (eds.) *Sentencing Guidelines: Exploring the English Model* (2013). South Korea has introduced sentencing guidelines which follow the English model.

[17] The Scottish Sentencing Council is beginning to issue guidelines (see *https://www.scottishsentencingcouncil.org.uk/* [Accessed 29 November 2017]) and the other jurisdictions actively contemplating the introduction of guidelines include Canada, some Australian states, the Gulf states and Israel.

capable of detecting and accounting for all legally-relevant (and only legally-relevant) factors, we would expect clear separation of cases along the dimension of severity (as measured by sentence length). This could be represented graphically by a smooth relationship between two dimensions, namely, sentence length and crime seriousness. The computer would have no a priori preference for any specific sentence lengths, and would be indifferent to the appeal of "round numbers". Human decision-makers, however, are likely to be influenced by the intuitive appeal of, say, one year over a term of 13 or 11 months, or four years rather than 50 months. A sentencer contemplating imposing a term of custody in the 10–14 month range may be more likely to settle on a year, and this means that some cases receiving this sentence merited less or more time in prison than exactly one year. The consequence is what we call *Penal Clustering*[18]—the tendency of courts to favour certain, specific sentence lengths. Penal clustering is the equivalent of a university marking distribution with, say, 24 possible marks (running from F up to A+) but where three marks—A, B, or C—account for almost all marks assigned.

The 19th century statistician, Sir Francis Galton, noted this phenomenon in a prescient article published in *Nature* over a century ago. Galton drew attention, for the first time, to the fact that certain custodial sentence lengths were used by judges[19] very frequently, while others were rarely imposed. Beginning with the assumption that "terms of imprisonment awarded by judges should fall into a continuous series",[20] Galton found that some sentence lengths were very popular, while others were never imposed. For example, a sentence of five years was 20 times more likely than a sentence of six years, and this "extreme irregularity" could not be explained by any legally-relevant circumstance. Galton argued that the human preference for particular numbers (then converted into fixed sentences of imprisonment) "interferes with the orderly distribution of punishment in conformity with penal deserts".[21] The appeal of certain sentence lengths undermined a court's ability to make finer distinctions based on differences in crime seriousness. In short, this preference for certain numbers impaired individualisation.

The only subsequent empirical analyses of preferred sentence patterns in England were conducted and reported by Pease and colleagues over 30 years ago.[22] Since then, however, no further research has been conducted on this issue, and much has changed, including the inception of sentencing guidelines. Ashworth[23] has suggested that the tendency to cluster exists today, but there has been no empirical test of this assertion. This article explores the extent to which the phenomenon that Galton identified all those years ago is still present in English sentencing.[24] We addressed two principal questions relating to individualisation in sentencing:

[18] Earlier scholars used the term *Preferred Sentences*: see K. Pease and M. Sampson, "Doing Time and Marking Time" (1977) 16 The *Howard Journal of Criminal Justice* 59.

[19] Although this article focuses on clustering in judicially-imposed punishments, the phenomenon of numerical clustering also affects statutory maxima, where certain sentences are more common than others.

[20] F. Galton, "Terms of Imprisonment" (1895) 1338(2) *Nature* 174, 174.

[21] Galton, "Terms of Imprisonment" (1895) 1338(2) *Nature* 174, 175.

[22] C. Fitzmaurice and K. Pease, *The Psychology of Judicial Sentencing* (Manchester: Manchester University Press, 1986), Ch.7 and K. Bottomley and K. Pease, *Crime and Punishment: Interpreting the Data* (Milton Keynes: Open University Press, 1986), p.105.

[23] A. Ashworth, *Sentencing and Criminal Justice*, 6th edn (Cambridge: Cambridge University Press, 2015), p.122.

[24] There is surprisingly little research on the phenomenon in other jurisdictions. One study in Canada examined parole eligibility dates for murder and found evidence of sentence length clustering: see C. Jones and M.B. Rankin,

- Do courts use a wide range of custodial sentence lengths or is there evidence of "penal clustering"[25] around a limited number of specific lengths?
- Did the introduction of the new assault offences guideline (in 2011) or the burglary offence guideline (2012), both of which aimed to promote greater consistency, adversely affect individualisation?

Method

Measuring penal clustering and individualisation in sentencing

Our measure of *penal clustering* is the degree to which sentences of imprisonment spread smoothly across the range of available sentence lengths—the measure adopted by Galton in his 19th century research. *Individualisation* is a more challenging concept, both to define and measure. There has been no scholarship on this dimension of sentencing, and for this reason we adopt a relatively straightforward measure. We employ two measures of individualisation: (i) the *number* of unique sentence lengths arising from the distribution of all custodial sentences imposed; and (ii) the *proportion* of all cases falling into the most frequently-imposed unique sentence lengths.[26] We assume that a distribution lacking individualisation would have fewer unique sentences than one in which individualisation is present.

It is important to note that these measures are not simply the opposite of variability. Two distributions of numbers may have the same "spread" or variability but very different degrees of individualisation. For example, consider these two illustrative distributions of sentence lengths: x: = .5,.5,1,1,1,1,1,1,1.5,1.5 and y = .5,.6,.7,.8,.9,1,1.1,1.2,1.3,1.4,1.5. These distributions have the same spread of numbers but the latter has more unique lengths, and therefore greater individualisation. A distribution of sentence lengths *lacking* individualisation would require only a small number of specific sentence lengths to capture all cases. If the new guideline has constrained individualisation, more cases will fall into a smaller number of unique sentences. The distribution of sentences will be captured by fewer sentences.

Two further measures related to individualisation at sentencing addressed other concerns about the guidelines. It may be argued that applying a guideline with three levels leads courts to overlook the diversity of sentencing factors present in a case. For this reason we compared the number of aggravating and mitigating factors taken into account under the old and new guideline. In addition, we tested the hypothesis that mitigation plays a diminished role under the Council's new

"Justice as a Rounding Error: Evidence of Subconscious Bias in Second Degree Murder Sentences in Canada" (2014) 52(1) *Osgoode Hall Law Journal* 109.

[25] The phenomenon of penal clustering may also be described as *unwarranted uniformity*, the opposite of unwarranted disparity. Thus, if two crimes, one of which merits three years and another five years in custody, are both assigned the same sentence (say, four years), this may be described as a loss of individualisation or excessive uniformity, since the two cases should be distinguished. See J. Pina-Sánchez and R. Linacre, "Refining the Measurement of Consistency in Sentencing: A Methodological Review" (2016) 44 *International Journal of Law, Crime and Justice* 68.

[26] The two measures are correlated, as a higher number of individual sentences will mean that a larger number will be needed to accommodate all sentences imposed.

format guidelines.[27] We compared the number of mitigating factors cited by courts before and after introduction of the guideline to see if the guideline had reduced the number of factors being employed by courts.

Data Source

The data are drawn primarily from the *Crown Court Sentencing Survey* (CCSS) administered by the Sentencing Council. Between 2011 and 2015, sentencers in the Crown Court were required to complete a data return every time they sentenced an offender.[28] The analyses draw upon data from 2011 and 2012 (the years relevant to determining the impact of the new format guidelines). The analysis is restricted to cases receiving an immediate custodial sentence since sentence length (in days) constitutes our principal variable of analysis in the exploration of individualisation and penal clustering. We compared the degree of clustering generally, and then determined whether there was any change following the introduction of the new guidelines. For burglary we drew upon the Court Proceedings database.

Findings

(a) Penal clustering

Before examining the evidence for penal clustering, we note an obvious manifestation of the appeal of one category of numbers. Of all custodial terms imposed for assault, fully three-quarters involved an even rather than an odd number of months.[29] This anomalous statistical pattern reveals the influence of cognitive biases on the determination of sentence and marks the distribution as deriving from a human decision-maker. The preference for even numbers has been well-documented in a range of decision-making contexts,[30] and these data show that the proclivity for one category of numbers influences sentencing decisions as well.

We now present the distribution of sentence lengths imposed in the 6,743 cases of assault sentenced to immediate custody in 2011, in order to evaluate the degree to which sentences cluster around certain, specific lengths. This distribution is represented graphically in Figure 1 which reveals that a high volume of cases cluster within a small number of unique sentence lengths. Although the cases are distributed across 194 different sentences, most were concentrated within a few specific lengths. The ten most frequently-imposed sentences capture over half (56%) of the total sample. Indeed, five sentences (one year; one and a half years;

[27] Cooper, "Nothing Personal" in Ashworth and Roberts (eds), *Sentencing Guidelines: Exploring the English Model* (2013), makes this claim, and other commentators have criticised the new guidelines for containing an imbalance of factors. For example, the guideline for GBH contains 50 factors, two-thirds of which are aggravating.

[28] There is a degree of missing data; some judges failed to complete the form when sentencing but the problem is neither large nor systematic. For further discussion of the limitations of the Crown Court Sentencing Survey, see M. Dhami and I. Belton, "Using Court Records for Sentencing Research: Pitfalls and Possibilities" in J.V. Roberts (ed.), *Exploring Sentencing Practice in England and Wales* (London: Palgrave Macmillan, 2015).

[29] This dominance of even numbers is more striking than that found in the only other study which has explored judicial preferences for even numbers. Jones and Rankin, "Justice as a Rounding Error: Evidence of Subconscious Bias in Second Degree Murder Sentences in Canada" (2014) 52(1) *Osgoode Hall Law Journal* 109, 126 found that Canadian judges imposed an even number of years in 62% of cases when determining periods of parole ineligibility.

[30] See, for example, Y. Nishiyama, "A Study of Odd- and Even-Number Cultures" (2006) 26(6) *Bulletin of Science, Technology & Society* 479.

two years; six months; eight months) account for almost 40% of all custodial sentences. The single most frequent outcome (one year) accounted for 13% of cases, as can be seen in Table 1.

Table 1: Most frequent custodial sentence lengths (in days), all assault cases, 2011

Sentence Length (in days)	360	540	720	180	240	270	480	450	300	1,080
% of cases in each sentence length category	13%	7%	6%	6%	5%	4%	4%	4%	4%	3%

Crown Court Sentencing Survey (CCSS); percentages rounded

Focusing on a more restricted band of sentence lengths highlights the clustering effect. Figure 2 shows the percentage of specific sentences within the range from 310 to 410 days. Although sentencers may choose from 100 unique sentences within this range, over 90% of the sentences imposed were exactly 360 days. A further 4% were 330 days and 2.2% were 312 days. Thus, only three unique sentences were necessary to account for 97% of all cases sentenced within this broad range of sentence lengths.

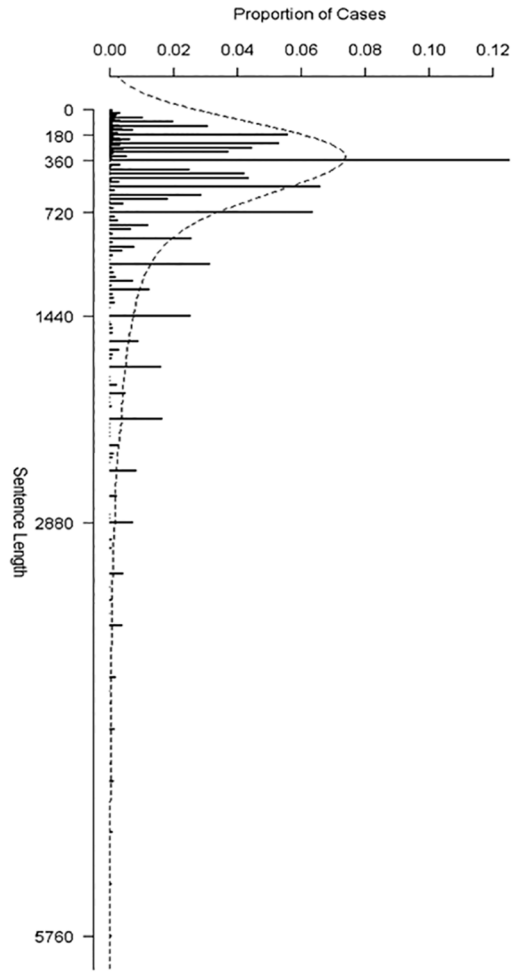

Figure 1: Distribution of custodial sentence lengths for cases of assault, 2011

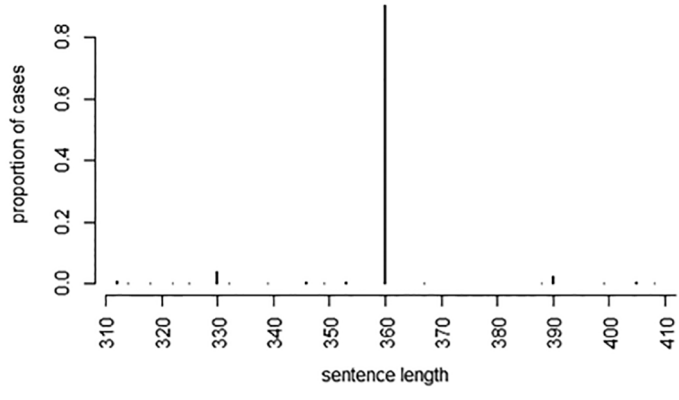

Figure 2: Proportion of sentences falling within the (310–410) day range

Offence-specific sentences

It is possible that clustering only emerges when the court has a serious offence which may result in a sentence within a very broad range. To explore this possibility, we compared a more serious assault offence (Causing grievous bodily harm with intent to do grievous harm (GBH))[31] and a less serious offence (Assault occasioning actual bodily harm, hereafter ABH).[32] Assault offences share certain commonalities and are grouped under the same sentencing guideline, but there are great differences in terms of harm and culpability (reflected in the guideline ranges). The more serious offence of GBH carries a maximum penalty of life imprisonment and a guideline maximum of 16 years; significantly higher than the five-year maximum and three-year guideline ceiling for ABH.[33]

First, we document the degree of clustering in sentences imposed for GBH. The sample now contains 1,526 sentences of immediate custody. Once again, clustering was evident, and to a greater degree than in the total assault sample. As can be seen in Table 2, the ten most frequent sentences encompass almost three-quarters (71%) of all sentenced cases. Furthermore, expanding this list to include the 19 most frequent sentences captures more than nine out of ten cases (91%). Only three specific sentences (one year; two years and 18 months) are necessary to accommodate almost four cases in ten. We then repeated the analysis for the less serious offence of ABH, and found a comparable degree of clustering. There were 2,291 sentences imposed for this offence, with 111 unique outcomes. The ten most frequent outcomes accounted for fully 70% of all sentences. Clustering thus occurs regardless of the seriousness of the crime.

Table 2: Most frequent custodial sentence lengths (in days), GBH and ABH Offences, 2011

GBH										
Sentence length (in days)	360	720	540	480	600	450	900	1,080	630	420
% of cases in each sentence length	14%	12%	10%	7%	6%	6%	5%	5%	3%	3%
ABH										
Sentence length (in days)	360	240	180	540	270	720	300	480	450	600
% of cases in each sentence length	16%	8%	8%	7%	7%	7%	5%	5%	5%	5%

Crown Court Sentencing Survey (CCSS); percentages rounded

Clustering/spreading by court

To address the question of whether there is significant variation in the degree of clustering in different courts across the country, we use as an index of clustering

[31] Offences Against the Person Act 1861 s.18.
[32] Offences Against the Person Act 1861 s.47.
[33] The range of recommended outcomes remained the same when the new guideline of assault was introduced. The maximum penalty is five years, or seven years if there is a racial or religious aggravating factor.

the number of specific sentences imposed divided by the number of cases sentenced. Thus, a court which sentenced 10 cases and used 10 specific sentences represents zero clustering and a perfect degree of "spread", with a spread factor of 1.00. A court that sentenced 40 cases and used only eight unique sentences would represent a low spread, high cluster score (.20). We ranked the Crown Court centres across the country regarding the degree to which they "cluster" their prison sentences—as represented by the "spread" factor. There was considerable range in the degree to which courts "cluster". Not surprisingly, courts sentencing more cases employed more unique sentences. However, variability in the degree of clustering did emerge, even after controlling for the number of cases sentenced. For example, in Swansea, 30 cases were assigned to one of 19 unique sentences, while in Leeds, the same number of cases were assigned to fewer unique sentences (13) suggesting greater clustering in the latter court.[34]

These analyses show that sentencers draw upon a small number of "popular" sentences. The threat to individualisation in sentencing observed by Galton in the courts of Victorian England persists today. The pattern of clustered terms of imprisonment is unexpected since the sentencing guidelines in England and Wales do not prescribe specific outcomes (but rather sentence *ranges*, leaving courts free to choose the actual sentence). Clustering would be expected in a jurisdiction with many mandatory sentences since these, by definition, require a court to cluster cases at a specific sentence. There are no such mandatory sentences for assault offences, however, and so this cannot explain the pattern of results which must be attributed to judicial rather than legislative decision-making. Whatever the explanation for the constrained distribution arising from courts' preferences for certain sentences, the consequence is a restriction on individualisation.

At this point, we examine the effect of the Council's new format guidelines on clustering and the use of sentencing factors.

(b) Assessing the impact of the new format Guidelines on individualisation

Since data are available for the period before the new guideline came into force, it is possible to evaluate the effect of the introduction of the new assault offences guideline in June 2011, and the burglary guideline of January 2012.[35] As noted, the new guidelines sought to improve the structure and coherence of the old guideline, with a view to promoting greater consistency.[36]

Examining all assault sentences imposed after the introduction of the new guideline, we find that matters improved in terms of our measure of individualisation in sentencing. After the implementation of the guideline we observe *more*, not fewer, specific sentence lengths being imposed. Equating the two samples in terms of the number of cases reveals that, in the post-guideline phase, the proportion of sentences encompassed by the 10 and 20 most common outcomes *decreased*, from 58.3% to 53.7% and from 79.3% to 76.1%; these

[34] The full table of centres is available from the first author (julian.roberts@worc.ox.ac.uk).

[35] We were not able to analyse guidelines introduced after 2012 due to a change of approach used to record custodial sentence outcomes in the CCSS. Whereas the first release of the CCSS provides custodial sentences in days, subsequent releases only provided only intervals of months or years, thereby preventing analysis of sentencing individualisation.

[36] Sentencing Council of England and Wales, *Assault Offences. Definitive Guideline*, *http://www.sentencingcouncil.org.uk/wp-content/uploads/Assault_definitive_guideline_-_Crown_Court.pdf* [Accessed 29 November 2017].

pre-post implementation differences are statistically significant; there is less than one chance in 100 that they could have arisen as a result of random variation.[37]

In other words, greater individualisation was observed following implementation of the new guideline. This finding is striking because it includes data from only the first year following introduction of the new guideline. It is reasonable to expect a new guideline to take at least a year to "bed down"—and to achieve its anticipated effect on trial court sentencing practices. Yet these data reveal statistically significant effects on individualisation within 12 months. We also tested the hypothesis that the new guideline for domestic burglary (introduced in 2012) reduced the degree of individualisation relative to the year previously. Consistent with the findings for assault, the results indicated no loss of individualisation following the introduction of this guideline.[38]

(c) Effects of the Guidelines on the use of sentencing factors

With respect to sentencing factors we also found no evidence of loss of judicial sensitivity to individual sentencing factors. In fact, the new format assault guideline had the opposite effect. Under the old assault guideline, an average of 3.65 sentencing factors were recorded across all cases. This average rose to 4.67 under the new assault guideline, a statistically significant increase. The effect presumably reflects the fact that the new format assault guidelines enumerate up to 50 factors. In contrast, the previous guidelines contained only a few common factors besides the type of activity comprising the assault, directing courts instead to consult the generic guideline for determining offence seriousness. Thus the previous guideline for section 18 assaults contained only a single additional sentencing factor (provocation) in addition to the seven factors common to all assaults. In contrast, the new guideline for this offence contains 50 sentencing factors relating to harm, culpability, seriousness or personal mitigation.[39]

Analysis of responses to the Crown Court survey revealed no reduction in the number of mitigating factors cited by courts following introduction of the guideline; the average number rose slightly from 1.24 to 1.33. Further analyses revealed that the number of aggravating factors increased to a greater extent than the number of mitigating factors, reflecting possibly the greater number of aggravating factors listed in the guidelines.[40]

Discussion and conclusions

This empirical research explored two potential threats to individualisation at sentencing: a tendency to impose certain sentence lengths; and the introduction of the new format guidelines—both of which may have further compressed different cases into a restricted number of sentences. Regarding the first issue, we demonstrated that despite the discretion that sentencers exercise in England and

[37] Two-sample tests for equality of proportions generate p-values of < .001 and < .007, respectively.

[38] This analysis drew upon CPD data from the Ministry of Justice. A t-test for the difference in proportions between the pre and post scenarios for the percentage of cases within the top 10 outcomes shows that the difference is not statistically significant, p-value = 0.2186.

[39] See fn.35, at pp.4–5.

[40] For example, the guideline for section 18 GBH contains 33 aggravating and 17 mitigating factors. This asymmetry reflects the reality that there are many ways to aggravate the commission of a crime, fewer ways to commit a crime under circumstances of lesser harm.

Wales, many custodial sentences imposed for assault are assigned to one of a restricted number of sentence lengths. The consequence is that the principle of individualisation is constrained. In his 1895 article, Galton concluded with, in his words, a little "moralising".[41] He decried the fact that "irrelevant influences" such as the "unconscious favour or disfavour felt for particular numbers" affects the choice of sentences of imprisonment. The phenomenon of penal clustering represents a threat to individualisation to the extent that a one-year sentence draws in cases which merit a shorter (or longer) sentence.[42] The consequence is a degree of unfairness in sentencing outcomes. If a court unconsciously[43] "rounds up" a sentence of imprisonment, the offender is effectively serving time in prison for which no justification exists.

One way of counteracting this procrustean tendency by courts would be to sensitise sentencers to the phenomenon, and to encourage them to sentence without regard to "round" or appealing numbers (such as one or two years). This might involve providing them with feedback on the distribution of sentence lengths imposed. Such remedial initiatives could target the courts which display the greatest degree of penal clustering. A second remedy would involve re-writing the guidelines so that sentence lengths are expressed in months, at least up to some limit (such as three or five years). This may also reduce the tendency to cluster around specific round numbers such as 12 months. Finally, guidance to sentencers to the effect that there is a presumption to "round down" when determining sentence lengths may mitigate any "rounding up" tendency.

Our conclusion regarding the other potential threat to individualised sentencing is more positive. Previous research has demonstrated an increase in consistency following the introduction of the Council's first guidelines[44] and, in this paper, we have demonstrated that this has not come at the price of diminished individualisation—at least for the first two guidelines issued by the Council, covering assault and burglary offences. This suggests that sentencing reform is not a zero-sum game; guidelines can enhance consistency without impairing individualisation. Assuming that these findings apply to other offences for which guidelines apply, we conclude that sentencing in England and Wales has not become less individualised as a result of the introduction of the new format guidelines.

Jurisdictions contemplating the adoption of a sentencing guidelines scheme will need to know exactly how a guideline may enhance individualisation. One explanation returns us to the structure of the guideline which requires courts to consider individually a range of sentencing factors. A court sentencing without a guideline will have only submissions from the advocates to guide its application of principles and consideration of mitigating and aggravating factors. In contrast, the guidelines require sentencers to work their way systematically through all steps

[41] Galton, "Terms of Imprisonment" (1895) 1338(2) *Nature* 174, 176.

[42] As a judicial colleague observed, there may be other causes of penal clustering. For example, in applying a one-third reduction for a prompt guilty plea, a court may well begin with a provisional sentence divisible by three (and then subtract one third). Two-thirds of any integer divisible by three yields an even number.

[43] It is also possible that a court may consciously round up or down in the interests of expediency or because certain sentences may seem more easily explained in open court.

[44] See J. Pina-Sánchez and R. Linacre, "Sentence Consistency in England and Wales: Evidence from the Crown Court Sentencing Survey" (2013) 53 *British Journal of Criminology* 1118 and J. Pina-Sánchez and R. Linacre, "Enhancing Consistency in Sentencing: Exploring the Effects of Guidelines in England and Wales" (2014) 30 *Journal of Quantitative Criminology* 731.

of the guideline. This requirement may sensitise courts to important differences between cases, resulting in the use of a higher number of unique sentences. A second possibility is that sentencers simply take longer to sentence a case when applying a guideline with nine steps. The guideline requires courts to consider each step, and to review its lists of aggravating and mitigating factors. This may provoke a deeper consideration of the case characteristics, generating greater discrimination between cases—and more unique sentence outcomes. Replication of the enhanced individualisation effect with the other offence-specific guidelines would confirm or disconfirm these explanatory hypotheses.

Finally, the trends on the use of sentencing factors suggest another general lesson. By identifying the principal aggravating and mitigating sentencing factors a guideline promotes a more consistent approach to considering these factors, and as we found with respect to the assault guideline, enumerating factors leads courts to take more factors into account, and thus to further promote an individualised approach to sentencing.

Data limitations and future research

These data derive from the Crown Court; the Sentencing Council has yet to collect comparable data in the magistrates' courts. The findings, therefore, cannot automatically be generalised to the lower level of jurisdiction. However, we see no reason to suppose that the clustering and the improvement in individualisation for assault sentences reported here would not be replicated in the magistrates' courts: The guidelines apply to all courts.

This research, and the preceding studies on consistency, represents a small step towards a comprehensive understanding of the impact of the English sentencing guidelines. More research is needed,[45] and we also echo others' appeals for qualitative research into the application of the guidelines, including analysis of sentencing remarks.[46] Future research needs to employ other measures of individualisation in order to explore the effects of guidelines on sentencing decisions. As recently as eight years ago courts in England and Wales sentenced offenders guided only by appellate decisions and the submissions of advocates. Today, they "must follow"[47] detailed offence-specific and "generic" guidelines. Understanding the impact of this more structured approach to sentencing remains a priority for sentencing scholars, and indeed the Sentencing Council itself. The other countries considering adopting some form of sentencing guidelines would also benefit from research upon the first jurisdiction to introduce statutorily based, offence-specific sentencing guidelines.

[45] Ultimately, the question of whether guidelines improve or undermine principled sentencing is an empirical one, answered only after research which conforms to the *scientific* approach involving falsifiable hypotheses (see K. Popper *The Logic of Scientific Discovery*. (London: Routledge, 2002). While there no shortage of armchair critics and advocates of guidelines, scientific research which conforms to Popper's essential criterion of falsifiability is rare.

[46] See R. Allen, *The Sentencing Council for England and Wales: Brake or accelerator on the use of prison?* (London: Transform Justice, 2017), p.21 and N. Padfield, "Exploring the Success of Sentencing Guidelines" in Ashworth and Roberts (eds), *Sentencing Guidelines: Exploring the English Model* (2013).

[47] Coroners and Justice Act 2009 s.125(1).