
Article:
Marder, ID and Pina-Sánchez, J (2020) Nudge the judge? Theorising the interaction between heuristics, sentencing guidelines and sentence clustering. Criminology and Criminal Justice, 20 (4). pp. 399-415. ISSN 1748-8958

https://doi.org/10.1177/1748895818818869

Nudge the judge? Theorising the interaction between heuristics, sentencing guidelines and sentence clustering

Abstract

Although it has long been acknowledged that heuristics influence judicial decision-making, researchers have yet to explore how sentencing guidelines might interact with heuristics to shape sentencing decisions. This article contributes to addressing this gap in the literature in three ways: firstly, by considering how heuristics might help produce the phenomenon of sentence clustering, in which a significant proportion of sentences are concentrated around a small number of outcomes; secondly, by reflecting on the role of sentencing guidelines as a feature of the environment within which sentencing decisions are made; and thirdly, by analysing the guidelines from Minnesota and from England and Wales, theorising how their content might interact with heuristics to make clustering more or less likely. Ultimately, we argue that sentencing guidelines likely affect the role played by heuristics in shaping sentencing decisions and, consequently, that their design should be informed by research evidence from the decision sciences.

Key words: Sentencing; judicial decision-making; sentencing guidelines; heuristics; sentence clustering; choice architecture; nudge theory

Acknowledgements: We would like to thank Prof. Julian Roberts, Anne Brunelle and the anonymous reviewers for their support and helpful comments.

Word count: 7922

Dr. Ian D. Marder (Lecturer in Criminology, Maynooth University)
Room 54, New House
Department of Law, Maynooth University
Co. Kildare, Ireland
ian.marder@mu.ie
+447912161264

Dr. Marder is Lecturer in Criminology at Maynooth University in Ireland. He completed his Ph.D. on the institutionalisation of restorative justice in the police at the University of Leeds in 2018. His research interests include practitioner discretion and decision-making, criminal justice policy and reform, and the implementation of restorative justice.

Dr. Jose Pina-Sánchez (Lecturer in Quantitative Criminology, University of Leeds)
School of Law, University of Leeds
LS29JT, UK
j.pinasanchez@leeds.ac.uk
+447909414923

Dr. Pina-Sánchez is Lecturer in Quantitative Criminology at the University of Leeds. Previously he worked at the London School of Economics as a Fellow in Statistics. His research interests span substantive and methodological areas such as sentencing, compliance with the law, survey research and statistical modelling.

1. Introduction

Sentencing is a complex cognitive task. Judges are required to recognise differences between offenders and between criminal cases, to identify and assign weight to a variety of aggravating and mitigating circumstances, and to select a sentence outcome from a wide range of available options (Ashworth, 2015), all while trying to reconcile several (often conflicting) principles and goals (Dhami, et al., 2015). We expect judges both to utilise a consistent decision-making process and to sentence offenders in accordance with each case's unique characteristics (Pina-Sánchez, 2015). However, empirical research consistently unearths evidence of sentence clustering, whereby terms of imprisonment are clustered around a small number of specific sentences (e.g. six, twelve or eighteen months) (Bottomley and Pease, 1986; Fitzmaurice and Pease, 1986; Jones and Rankin, 2015; Roberts, et al., 2018). This is problematic, insofar as it indicates that judges may be treating offences of different seriousness and offenders of varying culpability in a uniform manner, undermining the principle of individualisation. Although sentence clustering is a well-documented phenomenon, researchers have yet to develop sufficient theoretically- and empirically-informed hypotheses as to why it exists (Roberts, et al., 2018).

This article argues that we should draw on research from the decision sciences in order to help explain why sentence clustering occurs and how individualisation and other

normative goals might be promoted within sentencing. Specifically, we consider the possible role of heuristics – a group of cognitive shortcuts which humans use to make decisions (Shah and Oppenheimer, 2008) – in causing sentence clustering. We also examine sentencing guidelines in order to hypothesise how they might exacerbate or mitigate the phenomenon of clustering, depending on how their content interacts with judges' heuristics during the sentencing process. Dhami (2013a, 2013b) has observed that the design of sentencing guidelines is insufficiently informed by the psychology of decision making. By investigating one of the hitherto neglected dimensions of this complex policy area, we seek to add to the literature on how guidelines might aid sentencers in refining their decision-making processes.

This paper begins by explaining what heuristics are and exploring how they might contribute to sentence clustering. It then proposes that guidelines are increasingly a feature of the environment in which sentencing decisions are made, forming part of the 'choice architecture' (Thaler and Sunstein, 2009) of sentencing in jurisdictions where they are in force. Next, the paper presents an original analysis of two distinct sentencing guidelines from (1) Minnesota and (2) England and Wales, theorising how their content might interact with heuristics so as to exacerbate or mitigate clustering. It concludes by considering how 'nudge theory' (Thaler and Sunstein, 2009) could assist those who design sentencing guidelines to promote individualisation. Ultimately, we argue that, in

order to reduce the effects of heuristics on the sentencing process, policymakers should more strategically organise the environment in which judges make sentencing decisions.

2. Heuristics, sentencing and sentence clustering

Our ability to engage in rational decision-making is constrained by the parameters of our cognition. As Shah and Oppenheimer (2008: 207) argue, humans 'do not have unlimited processing capacity.' Rather, our rationality is 'bounded' (Simon, 1990) by our capacities to acknowledge and retain information, and to interpret and weight impartially the information which we do identify. We also tend to suppress doubt, neglect ambiguity, and accept or reject information depending on whether it conforms with our current beliefs (Kahneman, 2011). Essentially, we 'jump to conclusions' and use 'rules of thumb' when making decisions, without fully and fairly processesing all of the relevant, available information. These limitations of our processing abilities are exacerbated by our tendency to underestimate the likelihood that our decisions are based on imperfect cognition (Kahneman, 2011).

The consequences of this 'bounded rationality' are often minor, insofar as most of the decisions that we make in our day-to-day lives are relatively inconsequential (Simon, 1990). As key gatekeepers to (criminal) justice, however, sentencers make

decisions in their working lives which have significant, long-term implictions for offenders and victims, their families and wider society (Lipsky, 2010). It is necessary, therefore, to consider how these cognitive constraints might make it more or less likely that the goals and principles of sentencing will be realised in practice.

This paper focuses on how heuristics might cause or exacerbate the phenomenon of sentence clustering. Assuming that no two criminal cases are exactly alike, a perfectly individualised sentencing process would result in a smooth continuum of sentences which reflect all the nuances and differences between the cases brought before the court. Instead, sentencing data tends to suggest that a small number of specific sentences are given to a significant proportion of cases (Bottomley and Pease, 1986; Fitzmaurice and Pease, 1986; Roberts, et al., 2018). In a recent study of sentencing in England and Wales, for example, Roberts, et al (2018) found that just five sentences (one year; one and a half years; two years; six months; and eight months) accounted for around 40% of 6,743 assault cases sentenced to immediate custody in 2011; ten sentences accounted for over half (56%) of that sample. While it is not possible to determine objectively how much clustering is 'too much', the growing importance placed on individualised sentencing (e.g. Berry, 2016) means that these levels of clustering warrant further attention. Sentence clustering is antithetical to individualisation: the more clustering we observe, the less likely it is that sentence outcomes reflect the unique nature of each case or are proportionate, relative to each other (Berry, 2016; Roberts, et al., 2018). In other words, the presence of sentence

clustering may signify that varying levels of harm and culpability are being treated as equivalent by the courts.

We can turn to Shah and Oppernheimer's conceptualisation of heuristics (2008) in order to consider how they might contribute to this phenomenon. They define heuristics as cognitive shortcuts which reduce the amount of time and effort required in coming to a conclusion. These shortcuts include: failing to identify or acknowledge all of the relevant information; failing to store or recall information; weighting competing information arbitrarily; excluding relevant information when forming an overall impression of the alternatives; and examining only a subset of the potential decision options (see also Gigerenzer and Gaissmaier, 2011). According to Gilovic, et al. (2002), understanding decision-making in any context requires an acknowledgement of the role played by heuristics. Yet, in criminal justice, we often examine (and are quick to criticise) practitioners' decision-making without reflecting on the cognitive biases which structure their thinking. For example, we discuss racial and social profiling by the police without reference to the innate human propensity to make prejudicial assumptions based on the social-psychological processes of stereotype formation – what Smith and Alpert refer to as the 'illusory correlation phenomenon' (2007: 1273). The same is true in debates on sentencing, in which sentencers' decision-making is appraised and policies introduced to structure their discretion, without fully considering the underlying cognitive processes which structure their thinking (Dhami, 2013b).

We can easily theorise how heuristics might shape sentencing in ways which contribute to clustering. Consider, for example, the tendencies to take into account less information than is available and relevant, to forget some of the information which was initially selected for processing, and to weigh up fewer decision options than exist. In Shah and Oppenheimer's words (2008: 210-12), these heuristics might mean that sentencers use 'only a small subset of the available information when making judgments' and 'examine or compare fewer alternatives' than exist when selecting the outcome. As a result, clustering might occur or be exacerbated – and individiualisation might suffer – if differentiating circumstances of criminal cases are neglected or forgotten by sentencers and/or if some sentencing options are overlooked or rejected without due consideration as to their applicability.

The literature which applies heuristics to court decisions further illustrates how heuristics might cause or exacerbate sentence clustering. Several studies, for example, found that judges seldom identify and remember all the information which makes cases distinct. Dhami and Ayton (2001) observed that bail decisions often omit relevant factors, while Arce, et al. (2001) suggest that sentence disparities in similar cases might be attributed to differences in judges' (in)abilities to encode, store and retrieve all the relevant information. Von Helversen and Rieskamp (2009) found that prosecutors also consider only a small subset of the legally relevant and important factors when recommending sentences, arguing that their cognitive constraints might help explain why

important factors are often overlooked. The point is that judges likely fail to identify or remember all of the differentiating factors in each case, limiting their ability to distinguish between cases during sentencing and increasing the chances of differing cases incurring the same sentence outcome.

Peer and Gamliel (2013: 118) note that sentencing might also be shaped by heuristics because of 'the dominant effect of anchoring'. Anchoring is a type of heuristic in which decision-makers consider fewer outcome alternatives than exist, as their thinking becomes 'anchored' around a value with which they are presented (Mussweiler, et al., 2004). This value becomes an initial estimate which decision-makers then 'adjust' in order to find an acceptable outcome (Saks and Kidd, 1980). Anchoring is one of the most robustly studied and influential biases on numerical decision-making (Kahneman, 2011; Mussweiler, et al., 2004), and research suggests that it impacts judicial decisions. For example, Englich and Mussweiler (2001) found that experienced judges gave wildly different sentences to similar cases, depending on the sentence requested by prosecutors. Another study found that judges were 'anchored' by journalists' suggestions despite their legal irrelevance (Englich, et al., 2006), while Guthrie, et al. (2001) found that civil courts gave significantly lower damage awards when anchored by a defendant's motion to do so. Thus, if judges were exposed to a consistent anchor in the sentencing process, we might expect this to contribute to clustering: the anchor would become the initial estimate

which judges then adjust to determine the sentence in each case. As we shall see later, the starting points provided in sentence guidelines could act to trigger this heuristic.

We might expect judges to be skilled at identifying and remembering relevant factors and assessing varying outcomes for their applicability. After all, it is precisely this 'art' in which they are trained, and on which fair sentencing depends. Still, heuristics seem likely to influence sentencing in ways which limit the judges' ability to differentiate between cases. The stronger the influence of these heuristics on sentencing decisions, the more generic and uniform we might expect sentences to be (Alschuler, 2005). Conversely, the more we can reduce the influence of heuristics on sentencing decisions, then the more case characteristics and outcomes which judges might consider, and the more likely it should be that sentences come to reflect the unique features of each case, occasioning more individualised sentencing. To consider this subject further, we now explore the role which sentencing guidelines play in shaping sentencing decisions, before theorising how they might act to exacerbate or mitigate the cognitive barriers to individualisation.

3. Sentencing guidelines as choice architecture

For decades, empirical research from the decision sciences has shown how the environment in which decisions are made, interacts with heuristics and other cognitive

biases to shape those decisions. This research was summarised by Richard Thaler and his colleagues who coined the term 'choice architecture' to refer to the 'context in which people make decisions' (Thaler and Sunstein, 2009: 3; see also, Thaler, et al., 2013). They also used the term 'choice architects' in reference to the people who organise or design that context, and who therefore shape decision making. Importantly, Thaler and Sunstein (2009) argued that there is no neutral way to design choice architecture: environments necessarily influence decision making, irrespective of whether they are designed with any consideration as to their effects.

The choice architecture of sentencing includes a range of more or less tangible features, including the time pressures and legislative and precedential frameworks under which judges operate, and the representations from those involved in arguing cases, *inter alia*. We would also argue that, in jurisdictions where judges are compelled or expected to use sentencing guidelines to guide their decision making, these guidelines are part of the sentencing choice architecture. Given that sentencing guidelines are now proliferating around the world (Roberts, et al., 2018), therefore, it is important to consider how their interaction with heuristics might shape sentencing decisions.

Here, we investigate the potential interaction between heuristics and guidelines in two jurisdictions: (1) Minnesota and (2) England and Wales. These jurisdictions are the subject of recent research in this field (see, for example, Roberts, et al., 2018) for three reasons: a) their different guideline models are used to exemplify distinct approaches to

guideline design (Reitz, 2013); b) they are, as common law jurisdictions in the Anglo-American world, reasonably comparable; and c) they both publish enough numerical data to enable advanced quantitative analysis on sentencing outcomes.

In both Minnesota and England and Wales, guidelines are central to the sentencing process. Minnesota was the first American State to create legally binding guidelines which have since served as a model for several other States, including Pennsylvania and Oregon (Frase, 2005). They are designed in a 'grid' format: a two-dimensional structure, with crime seriousness along one axis (composed of eleven levels), and criminal history along the other (composed of seven levels). All offences must be assigned to one of the eleven offence seriousness levels, before the offender is given a 'Criminal History Score' running from '0' to '6 or more', depending on their prior convictions. The grid then provides either a presumptive stayed (i.e. suspended) sentence (for less serious offences) or a starting point and sentencing range at the intersection of each row and column (for more serious offences and recidivists). For example, a conviction for a seriousness level-six offence by an offender with a category-three criminal history, results in a 34- to 46-month term of imprisonment (Minnesota Sentencing Guidelines Commission, 2016); the same offence by an offender without a criminal record would only attract a 21-month suspended sentence. Importantly, the guidelines are mandatory and sentencers can only deviate from the prescribed ranges if 'substantial and compelling circumstances [are]

identified and articulated' as to why they must do so (Minnesota Sentencing Guidelines Commission, 2016: 2).

Guidelines in England and Wales, in contrast, are offence-specific. The first definitive sentencing guidelines came into force in 2007, addressed offences of assault and consisted primarily of narrative text (Ashworth and Roberts, 2013). In 2011, the Sentencing Council replaced the original assault guidelines with a new design, requiring sentencers to follow a series of structured steps (Dhami, 2013a, 2013b). This new format was used to revise all previous guidelines and has been used for all subsequent sentencing guidelines. Guidelines have since been issued for several further offences, including burglary, sexual and drug offences. At the time of writing, the most recently published guidelines relate to offences of manslaughter and child cruelty.[1] It is this most recent 'structured steps' format of England and Wales guideline to which this article refers.

Each sentencing guideline in England and Wales is subdivided into several specific offences.[2] Specific offences are then divided into three or four categories of seriousness, each of which is associated with its own starting point sentence and range. Judges determine the category by identifying and weighting factors relating to harm and culpability from an exhaustive list. For the offence of common assault, for example,

[1] See https://www.sentencingcouncil.org.uk/publications/?s&cat=definitive-guideline to access all of the published sentencing guidelines in England and Wales.

[2] For example, the assault guidelines cover: causing grievous bodily harm with intent; inflicting grievous bodily harm; assault occasioning actual bodily harm; assault with intent to resist arrest; assault on a police constable; and common assault.

eighteen such factors are listed, including a 'significant degree of premeditation' and a 'leading role in group or gang' (Sentencing Council, 2011: 24). Sentencers must then follow a series of structured steps (for common assault, there are nine in total) which require them to consider an array of aggravating and mitigating circumstances, any assistance which the offender provided the prosecution, reductions for guilty pleas and the totality principle, *inter alia*, when determining the final sentence.

The design of the England and Wales guidelines was influenced by the desire to structure the discretion of sentencers, while allowing for more judicial discretion than their American counterparts (Ashworth and Roberts, 2013). The law in England and Wales states that sentencing guidelines must be followed: Section 125(1) of the Coroners and Justice Act 2009 declares that all sentencers must 'follow any sentencing guidelines which are relevant to the offender's case […] unless the court is satisfied that it would be contrary to the interests of justice to do so'. Some commentators maintain that this legislation, when read alongside the step-structure format of the England and Wales guidelines, creates both a binding duty on the courts to follow the guidelines and an expectation that they will do so in practice (e.g. Roberts and Ashworth, 2015; Roberts and Bradford, 2015). According to Dhami (2013b), however, the term 'follow' is flexible enough that this law falls short of making the guidelines mandatory. Either way, as Dhami (2013a, 2013b) points out, the guidelines provide sentencers with considerable freedom to select and weight different factors as they see fit.

At this stage, it is useful to consider the empirical research on the different ways in which the sentencing guidelines are 'followed', as this term could denote adherence to either the *outcomes* or the *processes* outlined therein. This is important for our purposes because, the more evidence there is that the guidelines are followed by sentencers, then the more confidently we can state that they form a key part of the choice architecture of sentencing.

In relation to outcomes, adherence can be determined by measuring the proportion of cases in which sentences fall within or outside of the specified ranges. Minnesota provides regular reports on 'departure rates', most recently finding that, in the year 2016, around three-quarters of felony sentences fell within the guidelines' ranges (Minnesota Sentencing Guidelines Commission, 2017). A study in England and Wales (Roberts, 2013) found that, in 97% of cases of two assault offences, and in 92% of cases of another assault offence, sentences complied with the ranges provided in the guidelines. In addition, research from England and Wales suggest that discounts given for guilty pleas typically correspond with those mandated within the relevant guidelines (Roberts, 2013; Roberts and Bradford, 2015), suggesting that sentencers may make these decisions with reference to, or based on their knowledge of, the requirements therein. It is worth bearing in mind that low departure rates do not necessarily indicate an optimal sentencing process: low departure rates could conincide with high levels of clustering, for example, if judges

use only a small number of sentencing outcomes (albeit, ones which fell within the predetermined ranges for each offence).

Furthermore, although the low departure rates may indicate that judges largely adhere to prescribed outcomes, they do not necessarily mean that the outlined *procedures* are being followed in every case. With reference to the structured steps in the England and Wales guidelines, Roberts (2013: 111) notes that departure rates 'do not reveal the extent to which [judges] have "followed" the guidelines in terms of moving through the nine steps, taking into account the appropriate factors at the appropriate stage'. He concludes that, in the absence of research which examines exactly how and when judges engage with guidelines during sentencing, 'the full impact of the guidelines, and the way that courts use them, remains hidden from view' (Roberts, 2013: 111).

The most recent evidence from England and Wales suggests that sentencers might indeed be following the processes outlined in the guidelines. For example, Pina-Sánchez (2015) and Pina-Sánchez and Linacre (2014) found that sentencing consistency increased following a reformatting of the guidelines which explicitly sought to achieve this goal. The new format provides a much longer and clearer list of aggravating and mitigating factors than its predecessor, corresponding with findings by Roberts, et al. (2018) that sentencers reported taking more factors into account after this change was introduced. Roberts, et al. (2018) also found that, for offences of assault, clustering had decreased under the new guidelines, as measured by the proportion of offenders sentenced to the ten

or twenty most common outcomes. As the next section argues, we might expect to see lower clustering when multiple factors are listed in the guidelines, if judges refer back to them during sentencing. Further studies have also found that, as the England and Wales guidelines require, sentencers generally give more weight to 'Step One' factors (i.e. those indicating seriousness) than to the additional aggravating and mitigating factors provided in 'Step Two' (Fleetwood, et al., 2015; Irwin-Rogers and Perry, 2015; Lightowlers and Pina-Sánchez, 2017; Maslen, 2015; Pina-Sánchez and Grech, 2018). More recently, Pina-Sánchez, et al. (under review) found evidence that judges generally follow the step-sequence in the order prescribed within the guidelines. Specficially, they found that, for the most part, the case characteristics listed in Steps Two and Four of the guidelines were being used as originally intended.

There are further reasons that we might expect judges to engage with the guidelines. In Minnesota, the multitude of options presented on the grid may mean that judges, unless they are able to memorise all the calculations contained therein, must consult the guidelines regularly if they wish to ensure that their decisions are compliant. Judges in England and Wales might also need to refer back to the guidelines during sentencing in order to decide on the initial category of seriousness, consider each of the structured steps and reflect on the lists of aggravating and mitigating circumstances when determining the sentence – especially as these vary for the different offences. In both jurisdictions, moreover, sentencers are trained to follow the guidelines. In England and

Wales, the progressive introduction of new guidelines is accompanied by judicial training in their use (Roberts and Ashworth, 2015), meaning that, even if sentencers do not always consult the guidelines during sentencing, they should be aware of the guidelines' details and their own obligations therein. Finally, the fact that magistrates are not legally trained, need only sit a minimum of thirteen days per year, and tend not to oversee a large volume or variety of cases, might mean that they are unfamiliar with some offences' guidelines, and are therefore more likely to refer back to them for direction on a regular basis (Roberts and Ashworth, 2015). Given that magistrates tend to sentence around 95% of cases in the criminal courts (Office for National Statistics, 2017), this would mean that the guidelines play a significant role in shaping sentencing decisions in this jurisdiction.

Sentencing guidelines may not always be utilised or followed to the letter (or number). Still, it seems reasonable to conclude that *guidelines matter*: where in force, they are an important part of the choice architecture in which sentencing takes place. Given the influence of choice architecture over how people make decisions (Thaler and Sunstein, 2009), therefore, researchers must seek to assess both the likely interaction between heuristics and sentencing guidelines, and the implications of this relationship for achieving the goals of sentencing. We will now analyse two different guidelines in order to hypothesise how their design might intensify or reduce the influence of heuristics on sentencing, focusing on the implications for clustering and individualisation.

4. Analysing the architecture of sentencing guidelines

Research suggests that judges' heuristics might contribute to sentence clustering by preventing them from identifying and recalling all the relevant circumstances of the offence and offender and all the potentially applicable sentencing options. As a feature of the choice architecture which contextualises the sentencing decision, we should expect that sentencing guidelines interact with judges' heuristics in ways which strengthen or reduce these tendencies, depending on how the guidelines are designed.

In Minnesota, the standard sentencing grid[3] can be found on page 79 of the 146-page Sentencing Guidelines and Commentary (Minnesota Sentencing Guidelines Commission, 2016). Already, this design is problematic, if the designers sought to ensure adherence to the full detail of the guidelines: the decision sciences literature suggests that any guidance which is lengthy, convoluted and heaving with text and detail is unlikely to be fully utilised (Dhami, 2013a, 2013b; Hutton, 2013; Ruback and Wroblewski, 2001). Still, we can study the sentencing grid itself, as it is this feature of the document to which Minnesotan judges are most likely to refer when sentencing (Frase, 2005).

The grid's design requires judges to undertake three steps. The first and second steps, as described above, are to determine the crime seriousness and the criminal history

[3] The guidelines provide separate grids for sexual offences (p.81) and drug offences (p.83).

score, from which it is possible to identify the category into which each offence falls. The judges must then determine the final sentence based on the starting points and, in some cases, ranges provided in each offence category. There are no lists of aggravating and mitigating circumstances supplied next to the grid, nor any other explicit reminder or encouragement to individualise sentences.

An analysis of the Minnesota sentencing grid indicates that four of its features may act to exacerbate sentence clustering:

1) The first and second steps require judges to consider *only* the offence's severity and the offender's criminal history. Judges are therefore precluded (or, at least, discouraged) from taking other factors into account at this point (Frase, 2005).
2) The grid comprises a total of 77 offence categories, representing all the possible combinations of offence seriousness and criminal history.
3) 48 (62.3%) of these categories denote both starting points and sentence ranges; the remaining 29 (37.7%) categories relate to presumptive stayed sentences and only provide starting points.
4) The grid neither reminds nor encourages judges to consider aggravating and mitigating circumstances when determining whether, or by how much, the final sentence should deviate from the starting point.

These features of the grid might exacerbate clustering by preventing, discouraging or failing to remind or encourage judges to take multiple factors into account and consider the full range of sentencing options. First of all, judges are required only to consider two factors when selecting the category on which the sentence largely depends. Consequently, hundreds of offences are shoehorned into eleven groups and thousands of offenders are placed within just 77 categories, meaning that offences of varying seriousness and offenders of differing culpability will likely be assigned to the same categories. This is a common criticism of grid-based sentencing guidelines (Alschuler, 1991; Lowenthal, 1993; Schulhofer, 1991) and means that their design may exacerbate the tendency to exclude potentially relevant factors during sentencing.

Moreover, the grid does not encourage judges to consider the full range of sentencing outcomes. Firstly, we might expect that starting points, present in all 77 categories, will anchor the judges' thinking, risking uniform sentences that have not been reasoned exclusively according to the principles of sentencing (Isaacs, 2011). Secondly, although sentence ranges are provided directly underneath the starting point in two-thirds of offence categories (i.e. those representing immediate custody), the grid provides no incentives or reminders to adjust the starting point, and no aggravating or mitigating circumstances which judges could use to help them do so. The guidelines do provide a list of aggravating and mitigating circumstances which judges may use as *reasons to depart* from the predetermined ranges. Yet, this list is located some distance from the grid

(pp. 45-49), is woven inconspicuously into the narrative text which constitutes the bulk of the document and does not apply to the majority of cases that are sentenced within the ranges provided. These features of the Minnesota grid may compound the tendency to consider fewer factors and outcomes than are relevant and available, respectively.

As noted earlier, the England and Wales guidelines were developed with the express intention of affording judges more discretion than their Minnesotan counterparts (Sentencing Commission Working Group, 2008). Consequently, they diverge from those in Minnesota in several important ways, most notably by delineating a series of structured steps which sentencers must follow, but which are less prescriptive than the grid.

An analysis of one representative example of the England and Wales guidelines (the Assault Definitive Guideline [Sentencing Council, 2011]) suggests four ways that it may be more likely than the Minnesotan grid to mitigate for the impact of heuristics on the sentencing process:

1) It lists several aggravating and mitigating factors which judges are encouraged to consider during the first step (i.e. when an offence category is determined). This is divided into factors relating to harm and to culpability. For example, for the offence of common assault, the guidelines list three factors indicating higher harm and one indicating lower harm, as well as nine factors indicating higher culpability and five indicating lower culpability.

2) All offence categories provide both starting points and ranges. For example, the starting point for a category one common assault is a 'High level community order', while the range is specified as 'Low level community order – 26 weeks' custody' (Sentencing Council, 2011: 24).
3) In 'Step-Two', once a category is determined, judges are encouraged to consider a long list of potentially relevant aggravating and mitigating factors. Sentencers are asked to 'identify whether any combination of these, or other relevant factors, should result in an upward or downward adjustment from the starting point' (Sentencing Council, 2011: 25). The factors listed vary for different offences. For common assault, for example, nineteen aggravating factors and eleven mitigating factors are listed; it is also stated that this list is '**non-exhaustive**' (emphasis in original).
4) Following these two steps, judges are required to adjust the sentence in accordance with several further explicitly stated considerations, including any assistance the offender provided to the prosecution, reductions for a guilty plea, and dangerousness, *inter alia*.

In theory, these features of the England and Wales guidelines should act either to mitigate for the influence of heuristics on sentencing or, at least, exacerbate it to a lesser degree than the Minnesotan grid. For example, sentence ranges are detailed in every case where

a starting point is provided, reminding judges that they can deviate from the starting point. Moreover, sentencers who refer to these guidelines will encounter dozens of factors to consider when selecting the offence category and deciding whether and by how much to depart from starting points, potentially encouraging or reminding them to take multiple factors into account. They are even reminded, in bold, that there may be unwritten, relevant factors which they should consider. These design features could help sentencers to identify and recall more of the pertinent circumstances and consider more sentence outcomes, relative to the Minnesotan grid. This hypothesis is consistent with Shah and Oppenheimer's belief that decision-makers can be encouraged to identify and make use of additional information which is 'processed from an external source' (2008: 207) – to put it simply, if they are reminded and encouraged to do so.

This is not to say that we should necessarily expect more clustering in Minnesota than in England and Wales, as many other factors may contribute to this phenomenon. For example, levels of clustering may in part be a product of the training, culture or composition of the judiciaries. Additionally, the starting points in England and Wales will still act to anchor judges' thinking and may even be more likely to occasion sentence clustering than those in Minnesota. In England and Wales, starting points for terms of imprisonment are almost always expressed in whole years or in multiples of three months. This might encourage sentencers to use certain 'round' numbers (e.g. six, twelve or eighteen months) to a greater degree than the Minnesotan guidelines, in which both the

starting points and the sentence ranges are always expressed in months, and in which numbers of months which do not represent full, half or quarter years are commonly used.[4] If anchoring is indeed the most influential heuristic, we might expect to find more clustering overall in in England and Wales than in Minnesota, despite the various other differences between the guidelines. Anchoring caused by starting points and sentence ranges might also help to explain the finding by Roberts, et al. (2018) that sentencers in England and Wales prefer certain 'round' numbers.

Having illustrated how the interaction between heuristics and sentencing guidelines might impact on clustering, we can now turn to 'nudge theory' to consider how evidence from the decision sciences might be used to enhance individualisation and other sentencing goals.

5. Conclusion: nudging judges through sentencing guidelines?

[4] Of the 48 (out of 77) categories in the Minnesota standard grid which have both starting points and sentence ranges, only six provide a starting point and sentence range which both represent either full, half or quarter years. For example, the category for someone who commits murder in the second or third degree and has six or more previous offences provides a starting point of 240 months (12 years) and a range of 204 months (17 years) to 288 months (24 years). Starting points and/or sentence ranges in the remaining 42 categories all include numbers of months which are not divisible by three. For example, the category for a 1st degree assault committed by someone with no previous convictions provides a starting point of 86 months and a range of 74 to 103 months.

More empirical research is required before we can conclude exactly how and to what degree heuristics influence sentencing. To the extent that heuristics seem likely to contribute to sentence clustering, however, it is worth exploring whether 'nudge theory' may help the 'choice architects' who design sentencing guidelines to achieve greater individualisation in sentencing. Nudge theory posits that choice architecture can be designed in a manner which reduces the negative influence of heuristics on decision making, without restricting the choices available to decision makers (Thaler and Sunstein, 2009). This could be useful in the context of sentencing, where judges often resist efforts to constrain their discretion (Dhami, 2013a), but where it is nonetheless important to structure their decision making to maximise the chances that the goals of sentencing are realised. Nudge theory could help policymakers to design sentencing guidelines which improve decision quality, without prompting the resistance associated with compulsory, restrictive or prescriptive measures.

It is possible to conceptualise some of the existing features of the England and Wales guidelines as nudges. For example, both the lists of aggravating and mitigating circumstances and the notice that these lists are not exhaustive, could remind sentencers to consider more relevant factors. This, in turn, might help judges to individualise sentences, without policymakers being prescriptive in terms of which factors should be considered and how they should be weighted. These design features might help explain both the decrease in clustering and the increase in the number of factors which judges

reported considering, after this new model of guidelines was introduced in England and Wales (Roberts, et al., 2018). Nudges might also help to encourage more proportionate sentencing. Fleetwood et al. (2015) examined the impact of new drug offence guidelines in England and Wales which urge judges to link the culpability of drug traffickers to their 'role' in the process. Their analysis found that this change precipitated a fall in sentence lengths for offenders who had been coerced or pressured into trafficking. Future research must seek to establish empirically the manner and extent to which judges follow to the offence-specific (and, for that matter, the generic[5]) guidelines in order to help determine whether provisions such as this can influence sentencing in practice.

Future studies could also experiment with different kinds of nudges in order to promote almost any goal which we might wish sentencing to achieve. For example, they could be used to promote a greater emphasis on evidence-based sentencing (Lösel, 2012; Marcus, 2006). In the United States, where sentencing reform has gained bipartisan support in recent years (Karstedt, et al., 2018), guidelines could be used to nudge judges into greater use of non-custodial sentences. Researchers could make use of experimental methods and hypothetical cases to test the impact of different types of nudges on judicial decisions (Dhami, 2002; Dhami, et al., 2015). As nudge theory and the decision sciences become increasingly influential across public policy circles, criminologists may wish to

[5] England and Wales has sentencing guidelines which are not offence specific, but which outline principles for the sentencing of young people, for the imposition of community and custodial sentences and for assessing sentence reductions for guilty pleas, *inter alia*.

consider how the knowledge located in these fields might help us to achieve hitherto intractable problems in our own work.

It may be that judges can be nudged into undertaking a more analytical process than the more intuitive one which might otherwise occur (Kahneman, 2011; Krasnostein and Freiber, 2013). This is not to say that sentencing operates on intuition alone; Dhami, et al. (2015) have argued that sentencing is a 'quasirational' process, located somewhere in between intuition and analysis. Nor is it to suggest that heuristics and intuition are inherently 'irrational', or lead always to incorrect decisions (Alison, et al., 2011). Rather, it is to recognise that the use of intuition and the influence of heuristics in sentencing might make this process insufficiently precise to achieve individualisation (Casey, et al., 2013; Peer and Gamliel, 2013). Research has long established that practitioners in public institutions tend to over-simplify complex situations and humans' needs, when deciding how to use their powers to administer benefits and sanctions (Lipsky, 2010; McConville, et al., 1991). This results in decisions which systematically and predictably fail to respond to the unique nature of individual cases (Sunstein, et al., 2002). Given the proliferation of sentencing guidelines around the world, we concur with Dhami (2013a, 2013b): we must turn to evidence from the decision sciences to ensure that guidelines are always designed in ways which promote important sentencing goals.

Bibliography

Alison L, Barrett E and Almond L (2011) The interpretation and utilisation of offender profiles: a critical review of 'traditional' approaches to profiling. In: Adler J and Gray J (eds) *Forensic Psychology: Concepts, Debates and Practice.* Abingdon: Routledge, pp.38-57.

Alschuler A (1991) The failure of sentencing guidelines: a plea for less aggregation. *University of Chicago Law Review* 58(3): 901-951.

Alschuler A (2005) Disparity: the normative and empirical failure of the federal guidelines. *Standford Law Review* 58(1): 85-117.

Arce R, Farina F, Novo M and Seijo D (2001) Judges' decision-making from within. In: Roesch R, Corrado R and Dempster R (eds) *Psychology in the Court: International Advances with Knowledge.* New York City: Routledge, pp.195-206.

Ashworth A (2015) *Sentencing and Criminal Justice.* Cambridge: Cambridge University Press.

Ashworth A and Roberts J (2013). The origins and evolution of sentencing guidelines in England and Wales. In: Ashworth A and Roberts J (eds) *Sentencing Guidelines: Exploring the English Model.* Oxford: Oxford University Press, pp.16-29.

Berry W (2016) Remembering Furman's comparative proportionality: a response to Smith and Staihar. *Iowa Law Review* 65(1): 65-73.

Bottomley K and Pease K (1986). *Crime and Punishment: Interpreting the Data.* Maidenhead: Open University Press.

Casey P, Burke K and Leben S (2013) Minding the court: enhancing the decision-making process. *International Journal for Court Administration* 5(1): 45–54.

Dhami M (2002) Do bail information schemes really affect bail decisions? *The Howard Journal of Criminal Justice* 41(3): 245–262.

Dhami M (2013a) Sentencing guidelines in England and Wales: missed opportunities? *Law and Contemporary Problems* 76(1): 289-307.

Dhami M (2013b) A `decision science' perspective on the old and new sentencing guidelines in England and Wales. In: Ashworth A and Roberts J (eds) *Sentencing Guidelines: Exploring the English Model.* Oxford: Oxford University Press, pp.165-181.

Dhami M and Ayton P (2001) Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making* 14(2): 141-168.

Dhami M, Belton A and Goodman-Delahunty J (2015) Quasirational models of sentencing. *Journal of Applied Research in Memory and Cognition* 4(3): 239-247.

Englich B and Mussweiler T (2001) Sentencing under uncertainty: anchoring effects in the courtroom. *Journal of Applied Social Psychology* 31(7): 1535-1551.

Englich B, Mussweiler T and Strack F (2006) Playing dice with criminal sentences: the influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin* 32(2): 188-200.

Fitzmaurice C and Pease K (1986) *The Psychology of Judicial Sentencing.* Manchester: Manchester University Press.

Fleetwood J, Radcliffe P and Stevens A (2015) Shorter sentences for drug mules: the early impact of the sentencing guidelines in England and Wales. *Drugs: Education, Prevention and Policy* 22(5): 428-436.

Frase R (2005) Sentencing guidelines in Minnesota, 1978-2003. *Crime and Justice* 32: 131-219.

Gigerenzer G and Gaissmaier W (2011) Heuristic decision making. *Annual Review of Psychology* 62(1): 451-482.

Gilovic T, Griffin D and Kahneman D (eds) (2002) *Heuristics and Biases: The Psychology of Intuitive Judgment.* Cambridge: Cambridge University Press.

Guthrie C, Rachlinski J and Wistrich A (2001) Inside the judicial mind. *Cornell Law Review* 86(4): 777-830.

Hutton N (2013) The definitive guideline on assault offences: the performance of justice. In: Ashworth A and Roberts J (eds) *Sentencing Guidelines: Exploring the English Model.* Oxford: Oxford University Press, pp.86-103.

Irwin-Rogers K and Perry T (2015) Exploring the impact of sentencing factors on sentencing domestic burglary. In: Roberts J (ed) *Sentencing Guidelines: Exploring Sentencing Practice in England and Wales.* Basingstoke: Palgrave, pp.213-239.

Isaacs D (2011) Baseline framing in sentencing. *The Yale Law Journal* 121(2): 428-458.

Jones C and Rankin M (2015) Justice as a rounding error: evidence of subconscious bias in second degree murder sentences in Canada. *Osgoode Hall Law Journal* 52(1): 109-140

Kahneman D (2011) *Thinking Fast and Slow.* London: Penguin.

Karstedt S, Bergin T and Koch M (2018) Critical junctures and conditions of change: exploring the fall of prison populations in US States. *Social and Legal Studies.* Epub ahead of print 11 January 2018. DOI: 10.1177/0964663917747342.

Krasnostein S and Freiberg A (2013) Pursuing consistency in an individualistic sentencing framework: if you know where you're going, how do you know when you've got there?. *Law and Contemporary Problems* 76(1): 265-288.

Lightowlers C and Pina-Sánchez J (2017) Intoxication and assault: an analysis of Crown Court sentencing practices in England and Wales. *British Journal of Criminology* 58(1): 132–154.

Lipsky M (2010) *Street Level Bureaucracy: Dilemmas of the Individual in Public Service.* New York: Russel Sage Foundation.

Lösel F (2012) What works in correctional treatment and rehabilitation for young adults? In: Lösel F and Farrington D (eds) *Young Adult Offenders: Lost in Transition?.* Abingdon: Routledge, pp.74-112.

Lowenthal G (1993) Mandatory sentencing laws: undermining the effectiveness of determinate sentencing reform. *California Law Review* 81(1): 61-123.

Marcus M (2006) Smart Sentencing: Public Safety, Public Trust and Confidence through Evidence-based Dispositions. Available at: http://cdm16501.contentdm.oclc.org/cdm/ref/collection/criminal/id/161 (accessed 17 June 2016).

Maslen H (2015) Penitence and persistence: how should sentencing factors interact? In: Roberts J (ed) *Sentencing Guidelines: Exploring Sentencing Practice in England and Wales.* Basingstoke: Palgrave, pp.173-193.

McConville M, Sanders A and Leng R (1991) *The Case for the Prosecution.* London: Routledge.

Minnesota Sentencing Guidelines Commission (2016) Minnesota Sentencing Guidelines and Commentary. Available at: http://mn.gov/msgc-stat/documents/2016%20Guidelines/2016%20MN%20Sentencing%20Guidelines%20and%20Commentary.pdf (accessed 01 May 2018).

Minnesota Sentencing Guidelines Commission (2017) 2016 Sentencing Practices: Annual Summary Statistics for Felony Offenders. Available at: https://www.leg.state.mn.us/docs/2017/other/171185.pdf (accessed 01 May 2018).

Mussweiler T, Englich B and Strack F (2004) Anchoring effect. In: Pohl R (ed) *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgment and Memory.* Hove: Psychology Press, pp.183-200.

Office for National Statistics (2017) Overview of crime and justice statistics. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/methodologies/overviewofcrimeandjusticestatistics (accessed 01 May 2018).

Peer E and Gamliel E (2013) Heuristics and biases in judicial decisions. *Court Review* 49: 114-118.

Pina-Sánchez J (2015) Defining and measuring consistency in sentencing. In: Roberts J (ed) *Exploring Sentencing Practice in England and Wales.* Basingstoke: Palgrave Macmillan, pp.76-92.

Pina-Sánchez J and Linacre R (2014) Enhancing consistency in sentencing: exploring the effects of guidelines in England and Wales. *Journal of Quantitative Criminology* 30(4): 731-748.

Pina-Sánchez J, Brunton-Smith I and Guangquan L (under review) Mind the step: a more insightful and robust framework for the analysis of the sentencing process in England and Wales under the new sentencing guidelines. *Criminology and Criminal Justice.*

Pina-Sánchez J and Grech D (2017) Location and sentencing: to what extent do contextual factors explain between court disparities? *British Journal of Criminology* 58(3): 529-549.

Reitz K (2013) Comparing sentencing guidelines: do US systems have anything worthwhile to offer England and Wales? In: Ashworth A and Roberts J (eds)

Sentencing Guidelines: Exploring the English Model. Oxford: Oxford University Press, pp.182-201.

Roberts J (2011) Sentencing guidelines and judicial discretion: evolution of the duty of courts to comply in England and Wales. *British Journal of Criminology* 51(6): 997–1013.

Roberts J (2013) Complying with sentencing guidelines: latest findings from the Crown Court sentencing survey. In: Ashworth A and Roberts J (eds) *Sentencing Guidelines: Exploring the English Model.* Oxford: Oxford University Press, pp.104-121.

Roberts J and Ashworth A (2015) The evolution of sentencing policy and practice in England and Wales, 2003–2015. *Crime and Justice* 45(1): 307-358.

Roberts J and Bradford B (2015) Sentence reductions for a guilty plea in England and Wales: exploring new empirical trends. *Journal of Empirical Legal Studies* 12(2): 187–210.

Roberts J, Pina-Sánchez J and Marder I (2018) Individualisation at sentencing: the effects of guidelines and 'preferred' numbers. *Criminal Law Review* 2: 123-136.

Ruback R and Wroblewski J (2001) The federal sentencing guidelines: psychological and policy reasons for simplification. *Psychology, Public Policy and Law* 7(4) 739-775.

Saks M and Kidd R (1980) Human information processing and adjudication: trial by heuristics. *Law and Society Review* 15(1): 123-160.

Schulhofer S (1991) Assessing the federal sentencing process: the problem is uniformity, not disparity. *American Criminal Law Review* 29(3): 833-873.

Sentencing Commission Working Group (2008) *Sentencing guidelines in England and Wales: an evolutionary approach*. Sentencing Commision Working Group: London.

Sentencing Council (2011) Assault: Definitive Guideline. Available at: https://www.sentencingcouncil.org.uk/wp-content/uploads/Assault_definitive_guideline_-_Crown_Court.pdf (accessed 03 May 2018).

Shah A and Oppenheimer D (2008) Heuristics made easy: an effort-reduction framework. *Psychological Bulletin* 134(2): 207-222.

Simon H (1990) Invariants of human behaviour. *Annual Review of Psychology,* 41(1): 1-19.

Smith M and Alpert G (2007) Explaining police bias: a theory of social conditioning and illusory correlation. *Criminal Justice and Behaviour* 34(10): 1262-1283.

Sunstein C, Kahneman D, Ritov I and Schkade D (2002) Predictably incoherent judgments. *Stanford Law Review* 54: 1153-1216.

Thaler R and Sunstein C (2009) *Nudge: Improving Decisions about Health, Wealth and Happiness.* London: Penguin.

Thaler R, Sunstein C and Balz J (2013) Choice architecture. In: Shafir E (ed) *The Behavioral Foundations of Public Policy.* Princeton: Princeton University Press, pp.428-439.

Von Helversen B and Rieskamp J (2009) Predicting sentences for low-level crimes: comparing models of human judgment. *Journal of Experimental Psychology: Applied* 15(4): 375-395.