

# Predictive Evidence and Unpredictable Freedom

Amit Pundik ([amitp@tauex.tau.ac.il](mailto:amitp@tauex.tau.ac.il))

## Introduction

When determining in criminal proceedings whether an individual performed a certain culpable action, predictive evidence is often ignored.<sup>1</sup> Most apparently, and with only few exceptions, base-rates are excluded.<sup>2</sup> Using such evidence in court also seems *intuitively inappropriate*. For example, using the high rate of crimes involving illegal firearms in a certain neighbourhood to support the conviction of an individual resident in a crime involving an illegal firearm (henceforth, the “crime-rates scenario”) seems highly objectionable. The objection to base-rates is not only aimed at the sufficiency of such evidence (on the grounds that “crime-rates are insufficient on their own to prove that the individual is guilty”). The objection also requires that such evidence should not be used at all in determining the individual’s guilt: that crime-rates should be *inadmissible* in criminal proceedings.<sup>3</sup> The hostility of criminal fact-finding toward predictive evidence is also apparent in the deeply-rooted suspicion of bad character and previous convictions.<sup>4</sup>

While I share the view that the various accounts that seek to justify this hostility toward predictive evidence have been unsuccessful,<sup>5</sup> I do not defend this view here. Instead, I seek to propose an alternative account, which is admittedly both counterintuitive and demanding in its metaphysical commitments, yet successfully provides a unifying justification for why different types of predictive evidence should not be admitted in criminal fact-finding. I would suggest that

---

<sup>1</sup> I rely on Uviller’s distinction between trace and predictive evidence: the former results from a past event that leaves some traces in the present (eg eyewitnesses, fingerprints), while the latter ‘looks forward from an established event or trait to predict the likely repetition of its occurrence’. See R Uviller, ‘Evidence of Character to Prove Conduct: Illusion, Illogic, and Injustice in the Courtroom’ (1982) 130 Penn L Rev 845, 847.

<sup>2</sup> JJ Koehler, ‘When do Courts Think Base Rate Statistics are Relevant?’ (2002) 42 Jurimetrics J 373.

<sup>3</sup> This intuitive objection to admissibility distinguishes this example from the lottery and preface paradoxes in epistemology and the gate-crasher and prisoners paradoxes in legal theory. I have argued elsewhere that the latter are confusing and unhelpful – see A Pundik, ‘Freedom and Generalisation’ (2017) 37 OJLS 189.

<sup>4</sup> ‘English law’s suspicion of bad character and extraneous misconduct evidence has been cultivated for many centuries. It is deeply embedded in English judicial culture and institutions, and has frequently been actively propounded and celebrated’, P Roberts and AAS Zuckerman, *Criminal Evidence* (2<sup>nd</sup> ed, OUP 2010) 586.

<sup>5</sup> F Schoeman, ‘Statistical vs. Direct Evidence’ (1987) 21 *Noûs* 179; M Redmayne, ‘Exploring the Proof Paradoxes’ (2008) 14 LT 281.

the fact-finding practices used to determine culpability in criminal proceedings *implicitly adhere* to the view that culpable conduct requires *free will that is necessarily unpredictable*. While theorists of free will disagree on when an action can be considered free, they tend to agree that it is possible to predict a free action, at least to some degree of confidence. Contrary to this dominant view, in this paper I suggest that criminal fact-finding adheres to a theory of free will that includes a necessary condition of unpredictability, according to which free actions *cannot have probabilities* (henceforth, “the unpredictability condition”). This condition means that an *accurate* assessment of what an agent is likely to do *freely* is not merely epistemically unfeasible but metaphysically impossible. It is not only the lack of sufficient information that prevents an accurate prediction of how an agent will act freely: free actions cannot be predicted because their probability does not exist.

While I tend to think that, if free will exists, it is necessarily unpredictable, I do not pursue this claim here. Nor do I claim that the unpredictability condition is formally or consciously adopted by any existing legislation or judgment. I only argue that this condition is able to provide the sought-after justification for excluding predictive evidence.

In this paper, I assume that criminal punishment is constrained by culpability, at least if criminal law seeks to avoid punishing those who are not culpable for their actions. Hence, criminal proceedings constitute the legal context in which the role of attributing culpability is at its clearest. This constraint does not imply retributivism – namely, that punishment is inflicted because it is deserved. Instead, any theory of punishment that considers culpability to be a *constraint* on other legitimate goals of punishment should refrain from knowingly convicting the innocent.<sup>6</sup> I also assume, like most theorists of free will, that acting freely is a necessary condition of culpability.<sup>7</sup> While some might hold that our practices of attributing culpability do not require us to settle the metaphysical problem of free will,<sup>8</sup> I share the position that the distinction between justified and unjustified attribution of culpability – which any theory of

---

<sup>6</sup> One notable example of such a theory is Hart’s mixed theory, which accepts the retributivist constraint (‘only those who have broken the law—and voluntarily broken it—may be punished’) while rejecting retributivism as the ‘General Justifying Aim of the system’, HLA Hart, *Punishment and Responsibility* (2<sup>nd</sup> ed, OUP 2008) 9.

<sup>7</sup> A notable exception is that of semi-compatibilist theories (see the text accompanying footnote 12).

<sup>8</sup> P Strawson, ‘Freedom and Resentment’ (1962) 48 Proceedings of the British Academy 1.

culpability needs to make – is likely to rely on (or bring through the back door) notions very similar to “freedom” and “control”.<sup>9</sup>

The scope of my discussion is restricted in two respects. First, while I believe that my claim is applicable more widely, to legal and non-legal practices of determining culpability alike, I focus here on legal practices because they are easier to identify. Second, some culpable actions may cause the agent to perform further actions that may be both predictable and culpable (getting drunk voluntarily and then driving dangerously). The agent’s culpability for the latter seems to be *derived* from their culpability for the former. When, how and why culpability for one action is derived from another are complicated issues to address, and it is particularly questionable whether the agent’s culpability goes beyond their culpability for the first action. Be that as it may, such derivatively-culpable actions are outside the scope of this paper.

Section 1 explains the unpredictability condition by contrasting it to existing theories of free will and making some preliminary remarks on what a theory of free will that includes this condition might look like. Section 2 explains how the unpredictability condition can be used to provide a unifying justification for excluding predictive evidence.<sup>10</sup> Section 3 defends the suitability of the unpredictability condition to criminal fact-finding by criticising the suitability of the competing view, according to which being subject to causal influence enables the prediction of human conduct without rendering it unfree.

## 1. The Unpredictability Condition of Free Will

Since a free will theory that includes the unpredictability condition denies that free actions have probabilities, it is necessarily *incompatibilist*.<sup>11</sup> For compatibilist theories of free will, an action may be free even if it is determined by antecedent causal factors, so relying on these factors to predict the agent’s action poses no greater threat to freedom than determinism itself (which poses

---

<sup>9</sup> V Tadros, *Criminal Responsibility* (OUP 2005) 69.

<sup>10</sup> This section rehearses the argument I made in Pundik (n 3). Given the complexity of the issues involved (causation, free will and so on), I chose to repeat the argument itself in full but to remove some of the more nuanced qualifications. Readers who are not familiar with that paper and are left with some concerns about the claims made might find replies in there; and readers who are already familiar with the argument might want to move straight to the next section.

<sup>11</sup> My previous work was criticised by Federico Picinali for not specifying the theory of free will to which it is committed. See F Picinali, ‘Generalisations, Causal Relationships and Moral Responsibility’ (2016) 20 *Intl J of Evidence and Proof* 121. I hope this section rectifies this shortcoming.

none). By contrast, if free actions are *necessarily* unpredictable, an action that is determined by causal factors outside the agent's control *cannot* be free. This is because these causal factors would make the action at least potentially predictable (the possibility of prediction, and its accuracy, would depend on the state of our knowledge).

A similar point would apply to semi-compatibilist theories, such as Fischer and Ravizza's influential theory that distinguishes between "regulative control", which is incompatible with determinism but is not required for culpability, and "guidance control", which suffices for culpability and is based on the agent's responsiveness to reasons.<sup>12</sup> While semi-compatibilists uphold a stark division between *free* actions and *culpable* actions and could thus accept that free actions are unpredictable, they still hold, like "full" compatibilists, that *culpable* actions may be predictable (eg because they may be determined by causal factors that do not undermine guidance control). By contrast, in the following sections I claim that the unpredictability condition is able to justify the hostility of criminal fact-finding toward predictive evidence by looking into how *culpable* actions are proven. Consequently, my claim, if successful, would also suggest that the theory of free will to which criminal fact-finding adheres is libertarian rather than compatibilist or semi-compatibilist.

A constitutive feature of libertarian theories of free will is the claim that, if the agent's action were (fully) determined by antecedent causal factors outside their control, they would be neither free to do, nor culpable for doing, what they did. Yet, libertarians tend to accept the view that the agent's *free* actions have *objective probabilities*,<sup>13</sup> and that position is rarely challenged.<sup>14</sup> According to the objective interpretation of probability, the underlying reality itself is indeterministic. One common way to understand what objective probabilities are is to think of the indeterminism as lying in the cause itself.<sup>15</sup> Consider the following probabilistic generalisation: dropping a glass from a certain height onto a wooden floor will cause it to break in 50 per cent of cases. According to this understanding, dropping the glass is a genuinely indeterministic event:

---

<sup>12</sup> J Fischer and M Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (CUP 1998) 31–34.

<sup>13</sup> P van Inwagen, 'Free Will Remains a Mystery' (2000) 14 *Phil Perspectives* 1, 14–18; T O'Connor, *Persons and Causes: The Metaphysics of Free Will* (OUP 2000) 97; T O'Connor, 'Agent-Causal Power' in T Handfield (ed), *Dispositions and Causes* (OUP 2009) 189, 197.

<sup>14</sup> For exceptions, see L Vicens, 'Objective Probabilities of Free Choice' (2016) 93 *Res Philosophica* 1; G Sela, 'Torts as Self-Defense' (DPhil Thesis, University of Oxford 2017).

<sup>15</sup> D Lewis, *Philosophical Papers*, vol 2 (OUP 1986) ch 19. Another alternative (known as probabilistic causation and discussed in section 3) is that the indeterminism lies in the causal *relation*.

even had we known all the relevant facts (the particular fragility of the glass, the distance from the floor and so on) and the applicable laws of nature, it would have still been impossible to know, before the glass hit the floor, whether it would break in that instance.

And the same logic applies to human conduct: under the objective interpretation, it is impossible to predict with certainty how an agent will act, even if we know all that could possibly be known, because the underlying reality itself is indeterministic. However, just as it is possible to predict that the probability the glass will break is 50 per cent, it is also possible to predict with some degree of confidence what the agent will do. The better the prediction becomes, the closer it will be to the objective probability of that action. It is unsurprising that this is the dominant view among libertarians, because it enables them to account for the common practice of predicting what course of action an agent is likely to choose *freely*. If free actions have objective probabilities, they could be subject to causal influence, which would affect the probability that the agent will choose to act in a certain way without rendering the action unfree.

However, if criminal fact-finding adheres to the unpredictability condition, it cannot be based on any theory of freedom that holds that free actions have objective probabilities.<sup>16</sup> While this adherence requires a unique type of libertarian theory of free will,<sup>17</sup> and rules out most contemporary libertarian theories,<sup>18</sup> nothing in the commitment to libertarianism requires acceptance of the view that free actions have objective probabilities.<sup>19</sup>

While offering a comprehensive account of a libertarian theory that includes the unpredictability condition is outside the scope of this paper, I would like to make a few tentative remarks to suggest what such a theory might look like. Notably, the unpredictability condition is a necessary rather than sufficient condition of freedom. Consequently, some contemporary compatibilist

---

<sup>16</sup> Compatibilism does not imply that determinism is true, only that it poses no threat to freedom. Consequently, it is logically open for compatibilists to accept that the world is indeterministic and that free actions have objective probabilities. However, such a route is unlikely to be taken by compatibilists because it would deny them their luck-based objections to libertarianism. For the challenge luck poses to libertarianism, see N Levy, *Hard Luck* (OUP 2011) ch 3 (though Levy defines himself as a ‘disappointed compatibilist’ and argues that luck undermines compatibilism as well, *ibid.*, 2 and ch 4).

<sup>17</sup> A Kantian version of such a theory may be found in H Bergson, *Time and Free Will: An Essay on the Immediate Data of Consciousness* (tr. Frank Lubecki Pogson, Kessinger Publishing Company 1910). An alternative version may be based on the semantics of counterfactuals: counterfactual propositions about the action an agent would have carried out freely under different circumstances have no truth values. See Sela (n 14).

<sup>18</sup> n 13.

<sup>19</sup> Vicens (n 14). For O’Connor’s reply, see T O’Connor, ‘Probability and Freedom: A Reply to Vicens’, (2016) 93 *Res Philosophica* 289.

theories specify conditions that are unrelated to determinism, such as responsiveness to reasons,<sup>20</sup> and a libertarian theory that includes the unpredictability condition is likely to incorporate such conditions as well.<sup>21</sup>

A libertarian theory that includes the unpredictability condition need not deny that, in many cases, human conduct is predictable. For example, if a person drinks coffee almost every morning, we can predict with a high degree of confidence that they are going to drink coffee tomorrow morning as well. While such actions may be predictable, such a theory would deny that they are free in the sense that libertarians attribute to the term "freedom". Drinking coffee may certainly be free in various significant senses: it may be free from external interventions such as coercion, deception or manipulation; it may be free from political or social interference (drinking coffee is neither illegal nor socially unacceptable); it may be free from shortage of coffee beans, etc. However, that the action is free in any of these senses does not entail that it is free in the sense libertarians attribute to the notion of freedom and consider necessary for culpability, namely that a free action is one that is undetermined by causal factors beyond the agent's control. If it is the person's genetic composition or caffeine addiction that will determine their drinking of coffee tomorrow morning, all the aforementioned senses of freedom may hold, yet libertarians would insist that the individual's coffee-drinking was unfree. A theory that includes unpredictability could accept that many actions are indeed predictable, but suggest that they are predictable precisely *because of* their being unfree: the same causal factors that make an action predictable also render it unfree.<sup>22</sup>

Notably, the types and tokens of actions that may be free under the unpredictability condition are considerably fewer than under other theories of freedom (including many actions that are commonly regarded as 'free', such as drinking coffee). This implication of the unpredictability condition might be perceived as unattractive: *prima facie*, the more actions one can perform freely, the better off one is because one has "more" freedom. Putting aside the methodological

---

<sup>20</sup> Fischer and Ravizza (n 12).

<sup>21</sup> 'Being able to appreciate and act for reasons does not suffice for being a free agent. But such an ability is necessary for free will', Randolph Clarke, *Libertarian Accounts of Free Will* (OUP 2003) 15.

<sup>22</sup> Sela (n 13) suggests a more sophisticated explanation along similar lines. According to his 'Tuesday Freedom' approach, some types of action are carried out freely only *occasionally*: on most mornings, the agent drinks their coffee without exercising their free will, yet on Tuesday mornings, the agent drinks (or does not drink) their coffee freely. Predictions are based on the pattern created by the unfree occasions, and are hence irrelevant for predicting how the agent will act *freely*.

question of whether having an unattractive implication serves as an argument against a certain theory, I would like to suggest that this implication is not problematic, because having more opportunities to act freely does not necessarily make the agent better-off. First, the extent of the loss, if any, from not having “more” freedom seems to depend more on the quality of free actions than on their quantity: making a few important decisions in life freely matters more than drinking coffee freely every morning. Second, many predictable actions are *not even open to moral evaluation*. Drinking coffee may not cause any harm to others, violate any right etc, and consequently no blame or praise could be attributed to the agent for acting this way. In these cases, even if such actions are unfree, the range of actions for which the agent may be blameworthy or praiseworthy remains unaffected. Lastly, acting freely is not cost-free for the agent, because it might require time, effort and deliberation on their part. If it turns out that fewer of our actions are free, rather than this outcome being a source of disappointment it may actually be liberating.

A libertarian theory that includes the unpredictability condition would need to account for the role of the agent’s subjective reasons in their free actions. One option would be to accept that an action that was determined by the agent’s reasons cannot be free (because only undetermined actions may be free), but insist that such an action may still be culpable because its culpability may be derived from the agent’s previous free actions that resulted in acquiring or keeping these subjective reasons.<sup>23</sup> Since derivative culpability is outside the scope of this paper, I would like to focus on an alternative view, also common among libertarians, according to which *reasons are not causes* and reason-explanations are *non-causal* in nature.<sup>24</sup> Such a view need not deny that the agent may experience some motivations or subjective reasons for action as more powerful than others (perhaps because they cohere better with their beliefs, values and goals).<sup>25</sup> For example, an agent could freely choose to donate to charity *mainly* because they want to help others, and *also* because they want to claim a tax relief. However, such a view would deny that the difference in strength reflects a difference in the causal influence that each reason exerts on the agent; it would reject Davidson’s influential claim that reason-explanations are causal

---

<sup>23</sup> See, most notably, Kane’s notion of self-forming actions (SFAs), R Kane, *The Significance of Free Will* (OUP 1996) 74.

<sup>24</sup> See, for example, C Ginet, ‘Reasons Explanation of Action: An Incompatibilist Account’ (1989) 3 *Phil Perspectives* 17.

<sup>25</sup> I thank Timothy O’Connor for pressing me on this point.

explanations.<sup>26</sup> Consequently, from this perspective, the agent is not “pushed” by some reasons to act in one way and by other reasons to act in another way until some reasons prevail. Instead of being pushed around, when acting freely, it is the agent who chooses on which of their reasons to act.<sup>27</sup>

Under this view, the role given to reasons would probably amount to that of *enablers* of free decisions.<sup>28</sup> Free actions result from 'torn decisions':<sup>29</sup> for an action to be free, the agent needs to have reasons for each option between which they deliberate. For example, for an agent's decision to donate to charity to be free, they need to have reasons to keep the money for themselves. The conflicting reasons *enable* the agent's choice between different courses of action, because an agent cannot treat an action as a potential course of action they might take, if they have no reason whatsoever to perform that action. If the agent does not have equally strong reasons to keep the money and is thus not torn between the options, their decision to donate is determined by their reasons to donate and hence cannot be free (but if their decision is determined by a habit or previous decision to donate regularly, they may well be praiseworthy for developing this habit or making the previous decision). Understanding reasons as enablers rather than causes is consistent with the claim that free actions have no objective probabilities. That an agent is torn between two options does not mean that the objective probability of their choosing either is 50 per cent; deliberating and ultimately choosing between conflicting reasons is a prerequisite of acting freely, not a measure of probabilities.

Furthermore, free actions may still have rational explanations under this view, in the sense that they could be explained by the agent's subjective reasons for action.<sup>30</sup> Consider Mr Broyer, who stands outside a bank and deliberates between two options: going inside to rob it, or walking

---

<sup>26</sup> D Davidson, 'Actions, Reasons, and Causes' (1963) 60 J Phil 685.

<sup>27</sup> The sentence above assumes an agent-causalist libertarian theory rather than a non-causalist or event-causalist one (for the differences, see Clarke (n 21) ch 2). However, my preference for agent-causalist theory is based on considerations unrelated to predictability (following Clarke (n 21) chs 3-7) and it might be possible to include the unpredictability condition in a non-causalist or event-causalist theory as well.

<sup>28</sup> For the distinction between causes and enablers, see LB Lombard, 'Causes, Enablers and the Counterfactual Analysis' (1990) 59 Phil Studies 195

<sup>29</sup> The term 'torn decision' is taken from M Balaguer, *Free Will as an Open Scientific Problem* (MIT Press 2009) ch 3.

<sup>30</sup> The discussion in the text is based on Clarke's discussion of indeterminism and contrastive rational explanation (Clarke (n 21) 39–49). Clarke's discussion is conducted in the context of event-causalist theories of free will, but is equally applicable to agent-causalist theories.



away.<sup>31</sup> Under libertarian theories, if Broyer's action is free, it cannot be determined by factors outside his control. Until Broyer acts, even an omniscient bystander (be it God or an evil neuroscientist) cannot know what Broyer will do.<sup>32</sup> The unpredictability condition adds that, if Broyer's action is free, it also lacks an objective probability. Not only can an omniscient bystander not know what Broyer will do, they cannot even estimate rationally, because any estimation as to how Broyer might act would be equally good (or, more precisely, equally bad).

In reality, Broyer eventually opted to enter the bank to rob it. Under the unpredictability condition, while it was entirely unpredictable that he would freely go into the bank, his action was not whimsical or capricious, because it is possible to *explain* why he did so, based on his reasons for acting this way (eg he badly needed the money he intended to steal). What makes free actions so unique is that, had Broyer walked away, it would have *also* been possible to explain why he did so, based on other reasons of his (eg wanting to obey the law).<sup>33</sup> After all, he stood outside the bank and deliberated because he had strong reasons for either options.<sup>34</sup>

What *cannot* be explained about Broyer's action is why he went inside *rather than* walking away.<sup>35</sup> None of the causal factors outside Broyer's control, and *not even his reasons*, determined this outcome. What determined his going inside rather than walking away was *Broyer himself, as an agent*, enacting his will to cause the action. To generalise this example, for an action to be free, the agent needs to have a rational explanation for each of the courses of action they may freely take; yet, once the agent has acted, there is no *contrastive explanation* for why they have

---

<sup>31</sup> This example is inspired by the Israeli case of CrimA 9849/05 *State of Israel v Broyer*, 66 PD 726 (2006).

<sup>32</sup> The theological problem of foreknowledge and free will might differ from the problem of determinism and free will, but the former is used here only for the purpose of illustrating the unpredictability condition.

<sup>33</sup> I do not assume that only choices that comply with some thicker norms of rationality (eg Kant's Categorical Imperative) may be free.

<sup>34</sup> Questions concerning the conditions required for a mental process to constitute 'deliberation' and the norms with which such a process should comply are outside the scope of this paper.

<sup>35</sup> Clarke suggests that, in some cases, it may be possible to provide a contrastive explanation (eg if the agent preceded his action with a judgment about which course of action would be best to take). See Randolph Clarke, 'Reflections on an Argument from Luck' (2004) 32 *Phil Topics* 47, 52. I am concerned that this suggestion might lead to an infinite regress (for what contrastively explains the agent's preceding judgment?). Furthermore, the unpredictability condition is incompatible with this suggestion. While Clarke stipulates that the agent's previous judgment was among the causes of their action, according to the unpredictability condition, the later decision is free only if it had no objective probability, so the previous judgment could not causally affect it.

freely chosen one course of action over the other courses of action between which they deliberated.<sup>36</sup>

## 2. The Unpredictability Condition and Criminal Fact-finding

Recall the crime-rates scenario, in which the prosecution seeks to adduce the high rate of crimes involving illegal firearms in a certain neighbourhood to support the conviction of a resident in a crime involving an illegal firearm. The literature contains various accounts that seek to justify the exclusion of such predictive evidence. The first kind of strategy, which has received most scholarly attention, aims to identify an epistemic deficiency in the inference made from predictive evidence to the specific case. The inference is lacking – in weight,<sup>37</sup> appropriate causal connection,<sup>38</sup> case-specificity,<sup>39</sup> ability to provide the best explanation,<sup>40</sup> immunity to the problem of the reference class,<sup>41</sup> or sensitivity to the truth.<sup>42</sup> I am unconvinced by these epistemic accounts, because I think that not only does each one suffer from its own problems,<sup>43</sup> they also share some common deficiencies.<sup>44</sup> For example, why should the very same inference that is condemned as epistemically objectionable nevertheless be good enough for prediction purposes? If the inference suffers from some epistemic deficiency, this deficiency arises not only in the context of conviction but also in that of prediction.

---

<sup>36</sup> For theories of contrastive explanations more generally, see B van Fraassen, *The Scientific Image* (OUP 1980) 97–157; P Lipton, 'Contrastive Explanation', in D Knowles (ed), *Explanation and its Limits* (Royal Institute of Philosophy 1990) 247–266.

<sup>37</sup> LJ Cohen, *The Probable and the Provable* (Clarendon Press 1977) 74.

<sup>38</sup> JJ Thomson, 'Liability and Individualized Evidence' (1986) 49(3) L & CP 199.

<sup>39</sup> A Stein, *Foundations of Evidence Law* (OUP 2005) ch 3.

<sup>40</sup> M Dant, 'Gambling on the Truth: The Use of Purely Statistical Evidence as a Basis for Civil Liability' (1988) Colum JL & Soc Prob 31; MS Pardo and RJ Allen, 'Juridical Proof and the Best Explanation' (2008) L & Phil 223.

<sup>41</sup> RJ Allen and MS Pardo, 'The Problematic Value of Mathematical Models of Evidence' (2007) 36 JLS 107.

<sup>42</sup> D Enoch, L Spectre and T Fisher, 'Statistical Evidence, Sensitivity, and the Legal Value of Knowledge' (2012) 40 Phil & Pub Aff 197. The reference is to the epistemic explanation appearing in the first part of their paper, though in the second, they argue that epistemic considerations do not suffice to exclude predictive evidence, and later propose an alternative account based on primary incentives.

<sup>43</sup> A Pundik, "What is Wrong with Statistical Evidence? The Attempts to Establish an Epistemic Deficiency" (2008) 27 CJQ 461. See also Schoeman (n 5) and Redmayne (n 5).

<sup>44</sup> A Pundik, 'The Epistemology of Statistical Evidence' (2011) 15 Intl J of Evidence and Proof 117.

The second kind of strategy seeks to identify something in the *legal context* that makes some uses of predictive evidence objectionable, such as the rituality of the legal process,<sup>45</sup> the over-transparency of standards of proof,<sup>46</sup> equality between litigants,<sup>47</sup> and the individuality and autonomy of the litigant against whom the evidence is used.<sup>48</sup> Proponents of this type of account share the view that, even if such evidence may be useful in other contexts (science, policymaking and so on), its use in legal fact-finding conflicts with fundamental values of the legal system. I believe that, while there are specific problems with each of these accounts,<sup>49</sup> they capture something significant about predictive evidence because their strategy easily explains why the appropriateness of using this evidence depends fundamentally on the purpose for which it is used. In previous work,<sup>50</sup> I have suggested a contextualist account that is based on culpability. According to my culpability account, some types of generalisation about human conduct presuppose that the individual's conduct was determined by a certain *causal* factor that rendered their conduct unfree. By contrast, in the context of attributing culpability, it is necessary to presuppose the exact opposite: that the accused was free to determine their own conduct. Using these types of generalisation to determine culpability is objectionable, because it involves *contradicting* presuppositions about the individual's conduct.

The culpability account first argues that inferences about human conduct, drawn for either prediction or conviction purposes, require reliance on *causal* generalisations. In general, inferences from a known to an unknown empirical fact involve a generalisation about types.<sup>51</sup> In some cases, the reference to the generalisation is made explicitly. For example, inferring that Socrates is mortal from our knowledge that human beings are mortal refers explicitly to a

---

<sup>45</sup> LH Tribe, 'Trial by Mathematics: Precision and Ritual in the Legal Process' (1971) 84 Harv L Rev 1329.

<sup>46</sup> C Nesson, 'The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts' (1985) 98 Harv L Rev 1357.

<sup>47</sup> Stein (n 39) 105.

<sup>48</sup> DT Wasserman, 'The Morality of Statistical Proof and the Risk of Mistaken Liability' (1992) 13 Cardozo L Rev 935; AAS Zuckerman, 'Law, Fact or Justice?' (1986) Boston U L Rev 487.

<sup>49</sup> Schoeman (n 5). For criticism of Nesson and Tribe's accounts, see D Shaviro, 'Statistical-Probability Evidence and the Appearance of Justice' (1989) 103 Harv L Rev 530. For criticism of Wasserman's, see A Pundik, 'Statistical Evidence and Individual Litigants: A Reconsideration of Wasserman's Argument from Autonomy' (2008) 12 Intl J of Evidence and Proof 303. For criticism of Stein's, see, eg, A Pundik, 'Epistemology and The Law of Evidence: Four Doubts about Alex Stein's Foundations of Evidence Law' (2006) 25 CJQ 504.

<sup>50</sup> n 10.

<sup>51</sup> Schauer, for example, holds that 'the avoidance of generalizations is, with few or no qualifications, simply not possible at all'. FF Schauer, *Profiles, Probabilities and Stereotypes* (Belknap 2003) 101.

generalisation about human beings as a type. However, in many cases the generalisation is implicit in the inference. Consider, for example, an inference from the fact that a person reacted allergically to a certain cat to the fact that this individual is likely to react allergically to that same cat in future. This knowledge implies one or more generalisations that could serve as the basis for the inference (for example, the *type* of person who once reacted allergically to cats is likely to continue to react allergically). The important point is that drawing an inference from one empirical fact to another presupposes a generalisation about *types* of fact that connects the fact from which the inference begins and the fact with which the inference ends. Without this presupposition, the inference is invalid because it remains unclear what licenses the move from the first fact to the second.

I contend that inferences drawn from group membership to human conduct require a *causal* generalisation – that is, a generalisation that reflects a causal connection between the type of fact from which the inference begins and the type of fact the inference seeks to establish. If an inference is based on a non-causal generalisation, a mere correlation, it is unlicensed and thus invalid.<sup>52</sup> The causal relation can operate either directly or through a common cause. Inferring that a smoker is likelier to contract cancer than a non-smoker is based on a causal generalisation that smoking is a cause of (lung) cancer. By contrast, inferring that a Coca-Cola drinker is likelier to contract cancer than a non-drinker involves a causal generalisation that reflects a common cause. It is living in a hot country that is the common cause of both Coca-Cola drinking and (skin) cancer. Furthermore, the culpability account does not require us to *specify* the (direct or indirect) causal generalisation; it only requires that the *existence* of such a causal generalisation be presupposed.

Consider the opposite stance, according to which a *mere correlation* between two types of fact can suffice to infer an unknown from a known fact, even if there is no causal connection between the types of fact, not even indirectly. Such a stance would still require that the generalisation on which a valid inference is based satisfy certain conditions or standards, such as statistical significance. The difficulty with such a stance is that it renders the rejection of *spurious* correlations more difficult. Spurious correlations are those that do not reflect *any* actual connection (be they causal or not) between the two *types* of fact. Consider the almost-perfect

---

<sup>52</sup> This claim is part of the Common Cause Principle; see H Reichenbach, *The Direction of Time* (2<sup>nd</sup> ed, U California Press 1991) 158–159; F Arntzenius, 'The Common Cause Principle' (1992) 2 PSA 227.

correlation between the divorce rate in Maine and per capita consumption of margarine in the United States.<sup>53</sup> The lack of any actual connection between these facts means that this spurious correlation does not hold outside the group of initially-observed cases. It would hence be a mistake to infer anything about the consumption of margarine from the divorce rate (or vice versa) in a year that is not included in the group of years within which the spurious correlation was identified. Drawing any inference from a spurious correlation to an unobserved case is therefore unlicensed and misleading, whatever the purpose of the inquiry is (be it to obtain knowledge, provide an explanation or make a prediction about unobserved cases, for instance). Identifying a reliable process to ensure that a given correlation is not spurious is therefore essential, because spurious correlations are so widespread – indeed, they are *bound* to be ever-present. Since each specific case consists of innumerable details (most of which are, of course, unimportant), one could sift through a vast number of facts until one finds a group in which the identified fact correlates with the fact that one seeks to establish. For example, one might find a correlation between a certain type of action and the second (or third) letter of the person’s great-aunt’s surname.

If one accepts that inferences require *causal* generalisations, one can apply methods to distinguish between causal and non-causal connection<sup>54</sup> to identify which generalisations are spurious. However, if one denies that inferences require *causal* generalisations, one ought to find how to distinguish between informative and spurious correlations. Note that mere statistical significance will not do, because testing sufficiently large numbers of variables using sufficiently large databases would eventually generate statistically significant (yet spurious) generalisations. One might hope that such absurd, albeit statistically significant, correlations simply did not exist. But this hope relies on the belief that statistically-significant correlations need to “make sense” – that is, that it would be possible to *explain* why this correlation holds; and what would such an explanation be if not causal or causal-like?

One might object with the following counterexample: if there are ten balls in a jar, nine of which are blue, it is possible to infer that the probability of a randomly-chosen ball’s being blue is 90 per cent, without presupposing any causal connection between “being in that jar” and “being

---

<sup>53</sup> Tyler Vigen, ‘Spurious Correlations’ <[www.tylervigen.com](http://www.tylervigen.com)> accessed 12 August 2018.

<sup>54</sup> Various sophisticated methods have been proposed, such as the Markov Condition and Bayesian Nets. For a detailed introduction, see J Williamson, *Bayesian Nets and Causality* (OUP 2005).

blue”. However, this counterexample involves no *factual* inference, because the move from “nine out of the ten balls in a jar are blue” to “the probability of a randomly-chosen ball’s being blue is 90 per cent” does not add new information (that is, it is purely analytical). By contrast, predictions of human conduct are not analytical because the generalisation is based on a group of observed cases, from which the fact-finder is invited to infer something about *an unobserved* case (crime-rates are calculated based on *other* residents and brought in to support the contention that *this resident* committed a similar crime).<sup>55</sup>

But even if inferences about human conduct require reliance on causal generalisations, why cannot free actions be proven with such generalisations? Starting with a simple example, assume that Richard is exposed to radiation of a particular kind, which affects his nervous system, resulting in blotches all over his skin and an irresistible urge to attack everyone around him. Assume further that *every* person exposed to this radiation develops these symptoms. When Richard is admitted to hospital, it seems unproblematic to infer from the blotches that, given the opportunity, he will go berserk and should therefore be restrained. However, inferring from these marks that a violent action that had taken place before Richard arrived at the hospital was committed by him (rather than by someone else), for the purpose of convicting him of a violent offence, seems intuitively problematic.

According to the culpability account, this inference should not be used for the purpose of determining culpability, because it leads to a contradiction. To infer from Richard’s skin marks that he had acted violently, it is necessary to presuppose a *causal* generalisation: either one caused the other or they both have a common cause. In this example, the radiation caused both Richard’s blotches and his violent conduct. However, Richard’s acting violently may be culpable only if he acted *freely*. The culpability account is based on a libertarian theory of free will, which holds that people do not act freely when their conduct is determined by antecedent conditions outside their control. Establishing Richard’s guilt by inferring from his skin marks that it was he who acted violently is, therefore, contradictory: Richard’s conduct is treated as free and unfree at the same time.

Blaming Richard for a violent action, having inferred his conduct from the blotches, is problematic, since such an inference cannot be used without dissolving his culpability. Similarly,

---

<sup>55</sup> For a more elaborated discussion of this counterexample, see Pundik (n 3) 182.

if the inference is used to predict that Richard *will* act violently, it is only at the price of implying that his violent conduct will not be culpable. This example also explains why the very same inference seems unproblematic when restraining him in the hospital. While inferring from the skin marks that Richard will act violently in the hospital presupposes that his conduct is determined (and hence unfree), this leads to no contradiction because, in the medical context, it is not necessary to presuppose that Richard's violent conduct will be culpable.

Moving to probabilistic generalisations, consider the following variation on the previous example. Assume that Stephen is exposed to another type of radiation, which affects the nervous system and always causes certain skin blotches but causes an irresistible urge to attack others, when the opportunity arises, in only 80 per cent of cases. There are at least two ways to understand how this generalisation reflects the underlying causal relations between the radiation and the agent's conduct. According to the subjective interpretation of probability, which is commonly considered the most suitable for legal purposes,<sup>56</sup> probabilistic generalisations reflect the limited state of our knowledge rather than the true nature of the world. While the generalisation about the radiation is probabilistic, it imperfectly reflects a reality that may be deterministic. If the world is indeed deterministic, Stephen belongs to one of two possible sub-groups. One possibility is that he belongs to the sub-group of people who possess an extra unknown variable, which, together with the radiation, determines that he will go berserk. The other possibility is that he belongs to the sub-group of people who do not possess the extra variable, in which case the exposure to the radiation will not cause him to go berserk.

If Stephen possesses the extra variable, supporting his conviction by inferring from the blotches on his skin that he was (80 per cent) likely to have acted violently is problematic. Similarly to deterministic generalisations, such an inference leads to a contradiction. His conduct is taken to be both free (in order to be culpable) and unfree (as, together with another unknown variable, his violent actions were determined by the radiation). To avoid the contradiction, either the evidence of skin marks has to be accepted as probative of the violent act's having been committed by Stephen, in which case he is not culpable; or it has to be deemed not probative, in which case it should be ignored.

---

<sup>56</sup> For criminal law, see L Alexander and KK Ferzan with S Morse, *Crime and Culpability: A Theory of Criminal Law* (CUP 2009) 31; for tort law, see S Perry, 'Risk, Harm, and Responsibility' in DG Owen (ed), *Philosophical Foundations of Tort Law* (Clarendon Press 1995) 321, 333-335; for health and safety regulation, see MD Adler 'Against 'Individual Risk': A Sympathetic Critique of Risk Assessment' (2005) 153 U Pa L Rev 1121, 1247.

If Stephen does not possess the extra variable, inferring from his skin marks that he was (80 per cent) likely to have acted violently is mistaken and hence misleading. This is because, if he belongs to the sub-group of people who were not caused to act violently by the radiation, then the probability that he acted violently is not affected by the exposure to the radiation. Inferring from the skin marks that he is more likely to have acted violently than he would have been, had he not had these marks, is therefore mistaken. In sum, this inference is either contradictory, because it requires inconsistent presuppositions, or misleading, because it is mistaken and yet is presented as informative.

Using this evidence to support Stephen's conviction is objectionable also under the objective interpretation.<sup>57</sup> According to this interpretation, the radiation works in a genuinely indeterministic manner and it is impossible to know *at the time of the exposure* whether Stephen will go berserk. However, if Stephen is put to trial, the important question is whether the violent action, *which is given*, was performed by Stephen or someone else.<sup>58</sup> If the genuinely indeterministic radiation ultimately caused Stephen to go berserk, then his violent conduct was determined and not under his control. In such a scenario, the subjective and objective interpretations diverge on the question of whether the radiation, together with all relevant factors, determined Stephen's violent conduct, or whether there was room for chance. However, under both interpretations, Stephen's violent conduct was caused by a factor not under his control and hence he was unfree and cannot be held culpable for it. By contrast, if the radiation did not ultimately cause Stephen to go berserk, then inferring from his skin marks that he is likelier to have behaved violently is, again, mistaken. Therefore, inferring from the skin marks that he was likelier to have acted violently is either inconsistent with his being culpable, or mistaken and hence misleading.

The culpability account is able to provide a unifying justification for the hostility of criminal fact-finding toward predictive evidence. Returning to the crime-rates scenario, for an inference from crime-rates to the resident's case to be valid, it is necessary to presuppose that there is a causal generalisation that licenses this inference, be it the dangerous character of the neighbourhood, its socio-economic conditions etc. Such causal factors are outside the control of the individual

---

<sup>57</sup> The discussion here is based on understanding the indeterminacy of the radiation as lying in the cause itself (n 15). Understanding it as lying in the causal relation and *partially* influencing the agent's conduct is discussed in the text accompanying footnote 73.

<sup>58</sup> This issue is also known as 'actual causation' – see the text following footnote 77.



resident.<sup>59</sup> Inferring from the crime-rates that the resident was likelier to have committed a crime involving an illegal firearm is either inconsistent with their being culpable, or mistaken. As a result, if the court draws such an inference, it implicitly concedes the presupposition that the accused did not act freely. In such a case, the court would also have to concede that the individual is not culpable (and should therefore be acquitted).<sup>60</sup> Alternatively, if the court seeks to avoid the implications of this inference, it ought to deem it irrelevant to the individual's conduct and exclude the evidence adduced to substantiate it.

The culpability account also supports the common law's suspicion of previous convictions and yields some criticism of recent reforms. The rules and case law governing the admissibility of previous convictions are vast and complex, and I cannot provide here a comprehensive analysis of them. However, applying the culpability account to previous convictions of child molestation may serve as an example of how such an analysis might look. Previous convictions of child molestation are admissible in both the United Kingdom and the United States.<sup>61</sup> While the admission of such previous convictions has been criticised on various grounds, such as being unconstitutional,<sup>62</sup> unfair,<sup>63</sup> and even truth-suppressing,<sup>64</sup> the connection to the issue of free will seems to have gone unnoticed. The culpability account would draw attention to the importance of identifying the exact generalisation involved and considering whether using it for conviction conflicts with other presuppositions made in criminal proceedings. Like any inference about human conduct, inferring from the accused's previous convictions that they are likelier to have committed the alleged similar offence(s) relies on a causal generalisation. These previous convictions may be probative because they indicate that the accused suffers from a condition, such as perversion, illness or addiction, that raises the probability of reoffending. According to

---

<sup>59</sup> One might respond that the resident may still have some control over how to respond to these causal factors, leaving their conduct both predictable and free. I discuss this response in section 3.

<sup>60</sup> That convicting an accused should not be based on contradictory presuppositions should not be confused with the stronger claim that *every* case of practical decision-making is subject to *all* epistemic norms, a claim I do not endorse. Nor is it assumed that holding contradictory beliefs is, in itself, morally wrong – only that it is wrong to rely on contradictory beliefs to treat someone as culpable.

<sup>61</sup> For the United Kingdom, see the Criminal Justice Act 2003, c 44, pt 11, ch 1, s 103, and for the United States, see Rule 414 of the Federal Rules of Evidence.

<sup>62</sup> M Sheft, 'Federal Rule of Evidence 413: A Dangerous New Frontier' (1995) 33 Am Crim L Rev 57.

<sup>63</sup> J McCandless, 'Prior Bad Acts and Two Bad Rules: The Fundamental Unfairness of Federal Rules of Evidence 413 and 414' (1997) 5 Wm & Mary Bill Rts J 689, 694.

<sup>64</sup> M Cowley and J Colyer, 'Asymmetries in Prior Conviction Reasoning: Truth Suppression Effects in Child Protection Contexts' (2010) 16 Psychology, Crime & L 211.

the culpability account, if these previous convictions are indeed probative, it might only be at the price of exposing that the accused's conduct is unfree and thus nonculpable.

One might raise the concern that the culpability account rejects the use of *all* inferences, or at least numerous inferences that are intuitively unobjectionable. For example, supporting the accused's conviction of murder with the fact that the victim's bloodstains were found on the former's clothing relies on a generalisation by which individuals with the victim's blood on their clothes are likelier to have killed the victim in question than individuals whose clothes are not stained with the victim's blood. If drawing such an inference contradicts the presumption that the defendant acted freely, then my approach would deem legal fact-finding all but impossible.

However, the culpability account objects only to inferences that are based on causal generalisations in which the *direction* of the causal connection runs *from* the characteristic the person shares with other people to whom the generalisation applies *to* the culpable conduct. It is only then that the conduct might have been caused by an antecedent that renders it unfree and thus unsuitable for the determination of culpability. Unlike crime-rates, it is possible to use the evidence of bloodstains without presupposing a common antecedent factor among people with blood-stained clothes that causes them to commit a murder. Using bloodstains as evidence is hence not problematic, because the defendant's conduct (in this case, murder) was *not* caused by the bloodstains; rather, *it is their own free action that caused their commonality*, the murderous action that caused their clothes to be stained with blood.<sup>65</sup>

Another proving-too-much objection focuses on *opportunity*. If two people are found at the scene of a murder by stabbing, and one of them is completely paralysed, it seems intuitive to infer that the able-bodied person is the likelier of the two to be the murderer. The availability of an adequate opportunity seems to make it likelier that the person who had that opportunity acted culpably. It could thus be argued, once again, that if the culpability account objects to such inferences, it must be flawed.

As with the bloodstains example, the question remains whether these inferences are based on causal generalisations running *from* the characteristic the accused shares with other people *to* the culpable conduct. It is difficult to accept that ruling out the paralysed individual and suspecting

---

<sup>65</sup> This example also illustrates an important difference between the argument of this paper and existing attempts to give an account of what is wrong with using statistical evidence in court. While some accounts focus on what caused the *creation of the evidence*, eg Thomson (n 38), this paper focuses on what caused the *individual's conduct*.

the able-bodied of the crime is based on the causal generalisation that being healthy (or being able to move) *causes* people to commit murder. Rather, people immobilised through paralysis are outside the group of potential suspects to begin with, because they are *incapable* of committing the crime. Being able, or having an adequate opportunity, to commit the crime is a necessary condition of committing that crime, but is not the cause (just as the presence of the victim is a necessary condition of the event, but is not the cause of their death). The culpability account hence assumes a distinction between *causes* and *enablers*,<sup>66</sup> and accepts that being a cause is not reducible to merely being a necessary condition.

The issue of motive is more complicated. Evidence of motive is usually admissible in court,<sup>67</sup> and Redmayne even holds that ‘excluding motive evidence is counter-intuitive’.<sup>68</sup> For example, that the accused’s wife had an affair and their marriage broke down ‘shows that he had a motive (albeit an irrational motive) for killing her’.<sup>69</sup> I do not share Redmayne’s intuition (for example, I am not convinced that evidence of the accused’s being poor should be used to support their conviction of theft). More importantly, it is commonly thought that the general approach of substantive criminal law to motive is that the accused’s motives are irrelevant.<sup>70</sup> The culpability account would suggest that criminal evidence should also be wary of motives. One way in which motives may be probative of the accused’s conduct is by reflecting external influences that were outside their control and rendered their conduct unfree. If poverty is probative of stealing because poverty *causes* (some) poor people to steal, the culpability account can raise the same objection levelled against the radiation that causes (some) people to go berserk. A similar objection can be

---

<sup>66</sup> See Lombard (n 28).

<sup>67</sup> ‘When motive is relevant, evidence tending to show its existence is usually admissible, subject to exclusion if the risk of unfair prejudice is too great’, DP Leonard, ‘Character and Motive in Evidence Law’ (2001) 34 Loy LA L Rev 439, 439-40.

<sup>68</sup> M Redmayne, *Character in the Criminal Trial* (OUP 2015) 70. See also DN Husak, ‘Motive and Criminal Liability’ (1989) 8 Crim Just Ethics 3.

<sup>69</sup> *R v Phillips* [2003] EWCA Crim 1379.

<sup>70</sup> ‘Hardly any part of penal law is more definitely settled than that motive is irrelevant’, Jerome Hall, *General Principles of Criminal law* (2<sup>nd</sup> ed, Bobbs-Merrill 1960) 88; ‘It has been uniformly accepted in Anglo-American jurisprudence that motive is neither an element of a crime, nor a defense to its existence’, Theodore Sachs, ‘Criminal Law – Humanitarian Motive As a Defense To Homicide – State v. Sander, (N.H. 1950)’ (1950) 48 Mich L Rev 1199.

made if adultery is probative of murder because it causes (some) cuckolded husbands to kill their wives.<sup>71</sup>

Another way in which motives could be probative is by tracking the individual's subjective reasons for action. A theory of free will that includes the unpredictability condition may hold that such reasons are not causes but merely enablers.<sup>72</sup> If the prosecution needs to prove the existence of such an enabler, the culpability account can raise no objection against the evidence of motive. For example, if the accused claims in his defence that he did not steal or kill his wife, because he had no reason to do so, evidence of motive may be used to refute this claim and show that he faced the choice of whether to commit the alleged crime. However, if evidence of motive is adduced to establish that the accused was likelier to commit the crime by invoking the aforementioned causal generalisations, the culpability account can object because motives cannot be probative of what the accused had eventually chosen to do (for only causal generalisations can be probative). The culpability account thus draws attention to the generalisation underlying the inference from motive and the importance of scrutinising its effect on the accused's freedom and culpability.

### 3. Against The Partial-Influence View

The unpredictability condition yields a binary view, according to which human conduct could be either free (and thus unpredictable) or determined by causal factors outside the agent's control (and thus predictable to some degree). There is a widely-held view that denies this dichotomy. According to the partial-influence view, being subject to *partial* causal influence enables the prediction of human conduct without rendering it unfree. For example, when a person grows up in a crime-infested neighbourhood and then, as an adult, commits a crime, their criminal conduct may be free even if their growing up in that neighbourhood causally influenced their conduct to some degree, as long as it did not fully determine it. This view seems both intuitive and theoretically attractive because it enables us to predict what course of action the agent is likely to choose *freely*: the agent's action is predictable to some degree of probability, because it is

---

<sup>71</sup> Interestingly, the same generalisation that makes adultery probative could also be used to justify the defence of provocation, thereby acknowledging, in my view, that the underlying causal relation undermines the accused's freedom.

<sup>72</sup> See the text accompanying n 28.

partially influenced by some causal factors, but it may still be free, because it is not *determined* by these factors.

In this section, I claim that rejecting the culpability account by resorting to the partial-influence view is more problematic than it seems. Even if its proponents could provide an alternative justification for the hostility of criminal fact-finding toward predictive evidence, this view is also unsuitable for criminal proceedings for other reasons. This section is divided according to the different concepts that may be used to account for the partial-influence view: probabilistic causation, degrees of freedom and degrees of culpability. I outline each one in turn, briefly note some difficulties in their philosophical foundations and describe their unsuitability for criminal proceedings.

Notably, the partial-influence view is consistent not only with compatibilist theories, which hold that even *fully* determined conduct may be free, but also with most libertarian theories, which accept that free actions may have objective probabilities that may be causally influenced by factors outside the agent's control. Consequently, any challenge to the suitability of this view to criminal proceedings is also a challenge to the suitability of both compatibilist and libertarian theories that accept that free action may have either subjective or objective probabilities.

### *Probabilistic Causation*

The first way to explicate the partial-influence view is through the concept of probabilistic causation: the cause (growing up in a crime-infested neighbourhood) merely raises the probability of the effect (the agent's committing of a crime), without necessitating this outcome, thereby leaving the agent free to refrain from criminal conduct. The agent's conduct is hence *both* predictable (because of their background) *and* free, even under most libertarian theories of free will (because it is not determined by causal factors).

Indeed, when the culpability account was described earlier, the indeterminacy was placed in the cause itself, namely the *relata*. Recall the example of dropping a glass, which causes it to break in 50 per cent of cases. The indeterminacy under this understanding is analogous to flipping an imaginary coin, where the outcome (heads or tails) determines whether a *particular instance* of dropping would necessitate a particular glass to break. This imaginary coin, also known as an

objective chance,<sup>73</sup> serves as an activation mechanism that randomly separates between cases in which the dropping *fully* determines the breaking and those in which the dropping has *no effect* (in terms of breaking the glass). By contrast, theories of probabilistic causation place the indeterminacy in the *relation* between cause and effect, the dropping and the breaking: the cause *raises* the probability that the effect will occur.<sup>74</sup> The indeterminacy does not render the dropping fully effective only in *some* cases (in which the dropped glass breaks) and entirely ineffective in others (in which the dropped glass remains unbroken). Instead, the indeterminacy limits the effectiveness of the causal relation between the dropping and the breaking in all cases: a causal relation means that the dropping *raised the probability* of this breaking rather than necessitated or determined it. When a dropped glass breaks, the dropping is the cause of its breaking, just as is the case under deterministic causation. By contrast, when a different dropped glass does not break, the effect of the glass being broken does not exist and hence there is no causal relation (probabilistic or otherwise) in this case because a relation requires (at least) two relata. What was it, then, that determined that a certain dropped glass would break rather than not break? According to theories of probabilistic causation, the answer is that *nothing did*, not even an objective chance (otherwise the indeterminacy is pushed from the causal relation back to the cause itself). Placing the indeterminacy in the causal relation means that there is a causal explanation for why a certain glass broke (because it was dropped), but in most cases there is no *contrastive* explanation for why it broke rather than remaining intact despite being dropped.<sup>75</sup>

The idea of causation as probability-raising has been criticised for being subject to numerous counterexamples in which the cause actually *reduces* the probability that the effect will occur.<sup>76</sup> For example, assume that smoking causes lung cancer and is also correlated with being poor: poor people are likelier to smoke. Assume further that pollution from oil refineries raises the probability of lung cancer more significantly than smoking. If most people in a given city live in affluent neighbourhoods near an oil refinery, then smoking actually *reduces* the probability of

---

<sup>73</sup> Lewis (n 15).

<sup>74</sup> For a general introduction, see J Williamson, 'Probabilistic Theories' in H Beebe, C Hitchcock and P Menzies (eds), *The Oxford Handbook of Causation* (OUP 2009) ch 9.

<sup>75</sup> For some exceptions, see Clarke (n 35).

<sup>76</sup> For description and response, see C Hitchcock, 'Do All and Only Causes Raise the Probabilities of Effects?' in J Collins, N Hall and LA Paul (eds), *Causation and Counterfactuals* (MIT Press 2004).

contracting lung cancer (because smokers are less likely to live near an oil refinery). This example is an instance of what became known as Simpson's Paradox.<sup>77</sup>

More importantly, even if the latter objection could be answered (eg by refining the reference class), it focuses on the ability of theories of probabilistic causation to account for the causal relation on the *type* level (a connection between types of events, such as the activity of smoking and the medical condition of lung cancer). Probabilistic causation faces further difficulties in accounting for causal relation at the *token* level (a connection between the singular event of dropping a specific glass and the singular event of its breaking). It is particularly unsuitable for identifying *actual causation* – that is, the actual cause(s) of a specific effect that has already occurred – because there are forceful counterexamples in which an *actual cause reduced* the probability of the effect. Consider the following example, in which Assassin A is a gifted shooter and his chances of hitting a target are 90 per cent. However, he prefers someone else to do the dirty work, so once he notices that Assassin B is also aiming at the target, he decides to wait before he shoots, to see what Assassin B will do (even though only one of them can attempt the assassination in practice because the target will take cover after the first shot). Assassin B is very dedicated to his job, but he is not as gifted as his counterpart: his chances of hitting the target stand at just 50 per cent. Unfortunately for the target, Assassin B decides to shoot and shoots well, so the assassination attempt is successful. While B's shooting is clearly the cause of the target's death, theories of probabilistic causation yield the awkward conclusion that it *cannot* be the cause, because they hold that causes raise the probability of their effects, while B's shooting *reduced* the probability of the target's death from 90 to 50 per cent. While the scholarship on actual causation is vast and includes various attempts to address such cases,<sup>78</sup> I share the view that they are mostly unsuccessful.<sup>79</sup>

If theories of probabilistic causation are unable to account for actual causation, they are unsuitable for criminal proceedings, which focus mostly on the token level (the singular criminal case) and seek to identify the cause(s) of a specific outcome that already occurred. While the

---

<sup>77</sup> For a description of the paradox and attempted solutions, see N Cartwright, 'Causation Laws and Effective Strategies' (1979) 13 *Noûs* 419; Brian Skyrms, *Causation Necessity* (Yale UP 1980).

<sup>78</sup> See, in particular, David Lewis' extension of his counterfactual theory to probabilistic causation. D Lewis, 'Postscripts to 'Causation'' in *Philosophical Papers*, vol 2 (OUP 1986).

<sup>79</sup> See, for example, P Menzies, 'Probabilistic Causation and Causal Processes: A Critique of Lewis' (1989) 56 *PSA* 642. Even Lewis himself eventually abandoned his extension. See D Lewis, 'Causation as Influence' (extended version) in Collins, Hall and Paul (n 76) 75.

application of probabilistic causation to legal contexts has been discussed mainly within the context of tortious liability,<sup>80</sup> similar points could be applied to the criminal context. If theories of probabilistic causation were to be accepted in criminal proceedings, radical and counterintuitive reforms might be required. Assassin B, who, we intuitively conclude, caused the target's death, could not be convicted of murder under these theories, because his shooting reduced the probability of death and thus could not be its cause. I therefore doubt that rejecting the culpability account by resorting to the partial-influence view is likely to succeed if this view is explicated with probabilistic causation.

### *Degrees of Freedom*

Perhaps a more tenable explication of the partial-influence view is that the agent's freedom is influenced by some causal factors outside their control, but only *to some degree*, leaving them with a less-than-maximum extent of freedom. There is something intuitively plausible about this explication: human beings are different in so many ways and are subject to such a wide range of personal and environmental circumstances that believing they all enjoy the exact same degree of freedom seems unrealistic. Notably, this explication is not committed to the view of probabilistic causation: growing up in a crime-infested neighbourhood does not merely raise the probability of the agent's committing a crime. Instead, the causal relation between "growing up in that neighbourhood" and "committing a crime" can be deterministic: growing up there causes the agent to be *less* free, by denying them a certain *share* of their freedom.

Under this explication, actions may be simultaneously predictable and free. Their predictability is based on the causal factors that limit the agent's freedom. The stronger these factors are, the more predictable an action becomes: the stronger the influence of the neighbourhood on the resident, the more predictable their criminal conduct becomes. Yet, the action may still be free because these causal factors do not determine how the agent will act: if they eventually commit a crime, they still did so freely, *to some degree*.

My reservation about this explication is that it could quickly descend into a discriminatory and even racist worldview. If indeed different people have different degrees of freedom, surely young

---

<sup>80</sup> See, for example, Wright's defence of 'the actual causation requirement, which relieves a defendant of liability if his tortious conduct was not in fact a cause of the plaintiff's injury'. R Wright, 'Actual Causation vs. Probabilistic Linkage: The Bane of Economic Analysis' (1985) 14 JLS 435, 435.



black men in the United States could serve as an example of one collective that enjoys significantly less freedom than another, that of middle-aged white men (although opinions might differ as to the causes of, and responsibility for, this phenomenon). According to Michael Levin's troubling view, the cause of the higher crime-rate among black people is probably genetic, mediated by lower intelligence, impulsive temperament or higher serum testosterone levels.<sup>81</sup> Yet while such genetic influences may make black people more prone to crime, he clarifies, they do not render them unfree.<sup>82</sup> Instead, Levin concludes that 'blacks have less free will than whites'.<sup>83</sup> But, he argues, 'diminished responsibility does not imply greater leniency',<sup>84</sup> and he calls for harsher treatment of black people than white people (eg in addition to heavier punishments, he also suggests 'swifter administration of punishment to blacks, along with stricter limits on appeals').<sup>85</sup>

It is easy to object to Levin's conclusions on the basis that the lesser freedom young black men enjoy does not result from their genetic composition and the offensive mediating factors Levin purports (eg lower intelligence), but from the on-going discrimination by the state and hegemonic groups. It is also possible to object to Levin's calls for harsher treatment of black people and argue that the lesser freedom they enjoy should lead to more *lenient* treatment relative to other groups, particularly by the criminal justice system, which has arguably contributed to the higher rate of criminality among black people. However, it is important to note that, while these progressive responses may reach more palatable conclusions, they still accept Levin's fundamental claim that 'blacks have less free will than whites'.

Furthermore, the explication of degrees of freedom seems to me incompatible with criminal proceedings because the latter does not treat the question of whether black people are less free than white people as an *empirical question*. It is a matter of fact that some people have more options, better abilities and stronger willpower than others. However, criminal proceedings seem

---

<sup>81</sup> 'Race differences in IQ and temperament, variables which significantly affect criminal behavior, are significantly genetic in origin' and 'one mediating mechanism may be the race difference ... in serum testosterone, known to facilitate aggression', M Levin, *Why Race Matters: Race Differences and What They Mean* (Praeger 1997) 316 and 317 respectively.

<sup>82</sup> '[T]hat we are free when we do what we choose to do, although our choices are caused by unchosen genes, preserves the freedom of individuals whose unchosen genetic aggressiveness leads them to [choose] lawbreaking', *ibid*, 320.

<sup>83</sup> *Ibid*, 322.

<sup>84</sup> *Ibid*, 323.

<sup>85</sup> *Ibid*, 325.

to follow a normative assumption, according to which all members of society *should be treated* as either (fully) free or (entirely) unfree, at least in the context of determining culpability. A notable advantage of this binary approach to free will is that it treats all members of society as *equally* free. One might retort that, even if this advantage supports the view that criminal proceedings should adopt this normative assumption, it does not make it *true*: metaphysically, black people may have less freedom, whatever the law may choose to assume. Be that as it may, this paper focuses on the assumptions that criminal fact-finding would need to make about free will to justify its hostility toward predictive evidence, not on questions of whether these assumptions are true or not.

### *Degrees of Culpability*

Nelkin observes that '[w]e often judge some people to be more blameworthy than others for their actions'<sup>86</sup> and that 'we have intuitions [...] that difficulty is mitigating in blameworthy cases, and excuses based on difficulty play a large role in our moral and legal practices'.<sup>87</sup> Indeed, sentence mitigation seems to provide a paradigmatic case to support Nelkin's observation. For example, a paedophile's sentence might be mitigated by the fact that he was a victim of molestation in his childhood. If the partial-influence view is explicated with degrees of culpability, this mitigation acknowledges that his being a victim himself renders him worthy of less culpability than a paedophile who was not a victim of similar offences in his childhood. According to this explication, culpable actions may be predictable based on the causal factors outside the agent's control that reduce their culpability: being a victim of molestation in his childhood partially influenced the paedophile to molest other children, thereby rendering his actions predictable to some degree and reducing his culpability, but without undermining his freedom and excusing him altogether.

Notably, degrees of culpability do not have to be based on either probabilistic causation or degrees of freedom. The causal relation between the antecedent factors and the agent's culpability may connect between these factors and a certain *share* of the agent's culpability. Furthermore, degrees of culpability may be based on a binary concept of freedom. One might

---

<sup>86</sup> DK Nelkin, 'Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness' (2016) 50 *Noûs* 356, 356.

<sup>87</sup> *Ibid*, 370.

hold that causal factors outside the agent's control may reduce their degree of culpability without affecting their freedom because being subject to partial causal influence is consistent with being (fully) free (according to compatibilist or libertarian theories that accept that free conduct has subjective or objective probabilities). Alternatively, one could argue that the binary question of whether one is free or unfree is immaterial to culpability because the latter does not require free will (according to semi-compatibilist and reactive attitudes theories).<sup>88</sup>

Nelkin argues – persuasively – that libertarians have no advantage over compatibilist theories in accounting for degrees of culpability because existing libertarian theories lack any unique theoretical resource to do so.<sup>89</sup> Libertarians would account for degrees of culpability by connecting them to degrees of objective probability that free actions allegedly have: the higher this probability is, the stronger the external causal pressure on the agent to act in a particular way, thereby making it more difficult for them to resist and rendering them worthy of less culpability for yielding to the pressure. However, as Nelkin demonstrates, there are various counterexamples to this view. For example, a person who makes an extraordinary effort to overcome a temptation to act wrongly – thereby reducing the probability of the action significantly – but who eventually fails, and yields to the temptation, would be *more* culpable than a person who yields to the same temptation after making virtually no effort to resist it. Libertarians would reach this counterintuitive outcome because the objective probability of the former committing the culpable action is lower, yet they committed it nevertheless. Nelkin then suggests a compatibilist way to account for degrees of culpability and concludes that ‘attention to degrees of blameworthiness and praiseworthiness [...] can play an important part in a larger case for compatibilism’.<sup>90</sup>

At first glance, a libertarian theory that includes the unpredictability condition seems to be even more vulnerable to Nelkin’s criticism, because if free actions have no objective probabilities, this view is left with no theoretical resources to account for degrees of culpability. However, I think Nelkin’s conclusion is too hasty. First, that culpability comes in degrees could be denied altogether. It may be suggested that whenever the agent’s culpability seems to be greater, this is not because it is of a higher degree, but because the agent is *culpable for more than one thing*. The person who made great efforts to resist the temptation is *not more* (or less) culpable than

---

<sup>88</sup> nn 8 and 12

<sup>89</sup> Ibid, 358–60.

<sup>90</sup> Ibid, 374.

another who made virtually no effort. Rather, they are both *equally* culpable for the same thing: acting the way they ultimately did. It is the second person who is culpable for an additional thing: not making any effort to resist the temptation in the first place.

Second, even if culpability does come in degrees, libertarian theories could utilise whatever theoretical resource compatibilists would use to account for degrees of culpability, by connecting degrees of culpability to this component rather than to objective probabilities. As for the unpredictability condition, it is a necessary rather than sufficient condition of freedom, and a libertarian theory that includes it is likely to incorporate other conditions, which could explain, *inter alia*, why one agent is worthy of more culpability than another.<sup>91</sup> Consequently, such a theory need not deny the existence of degrees of culpability, only that the component that makes an action worthy of more culpability also makes it more (or less?) predictable.

Third, sentence mitigation could be accounted-for without any reference to the partial-influence view, no matter how this view is explicated. True – responses to culpable conduct are typically scalar: punishment, for example, could include a longer or shorter period of imprisonment or a heftier or lighter fine. However, it is possible to explain why the paedophile's childhood experience serves to mitigate the appropriate punishment without referring to this experience as a factor that causally influenced him to molest other children. To mention just a few alternative explanations: the shorter imprisonment could be related to the increased effect that imprisonment would have on him as a result of his experience, his vulnerability to becoming a victim again during imprisonment, an attempt to compensate him for his bad luck or maybe even pure mercy. While exploring the justification for sentence mitigation lies outside the scope of this paper, the important point is that sentencing mitigation need not be based on the view that the factors that justify mitigation are causal factors that partially influenced the perpetrator's conduct.

Last, and most important for the purpose of this paper, those who rely on the partial-influence view to account for sentence mitigation face a challenge to account for the conviction stage of the trial. If sentence mitigation is based on the view that evidence such as crime-rates reflects a partial *causal* influence, why not allow its admission at the conviction stage? After all, if the accused's personal background (be it their childhood experience or socioeconomic conditions, for instance) made their actions more predictable, this means that they are likelier to have committed

---

<sup>91</sup> See the text accompanying n 20.

the alleged crime. Proponents of the partial-influence view would hence need to explain why the prosecution should not be allowed to admit the very same evidence at the conviction stage to support its allegation that the accused has committed the offence.

## Conclusion

In this paper, I have proposed to account for the hostility of criminal fact-finding toward predictive evidence with the unpredictability condition. I have suggested that the fact-finding practices used to determine culpability in criminal proceedings implicitly adhere to the view that culpable conduct requires free will that is necessarily unpredictable.

While I have not defended the claim that free will is unpredictable, or that it even exists, the sceptic could nevertheless utilise my account to claim that it demonstrates the futility of any attempt to hold people culpable for what they did *freely*. This is because the unpredictability condition itself is so metaphysically demanding that it is unlikely to exist. Alternatively, when we know enough about a certain person, we can predict how they will act with high degree of confidence, so the unpredictability condition is hardly ever met, even if it does exist. Either way, the sceptic could posit that, if criminal fact-finding adheres to the unpredictability condition, it in fact adheres to the view that we rarely act freely, if ever at all, and we should hence dispense with any attempt to use criminal proceedings to attribute culpability to individuals.

While I find this sceptical claim forceful and tempting, it is important to remember, as some sceptics have emphasised,<sup>92</sup> that what is at stake here is not only our practices of attributing culpability, but also various issues that seem to depend on the existence of free will: praise, the ability to deliberate, being the ultimate source of one's actions and achievements and perhaps even the meaning of life.<sup>93</sup> The sceptical stance vis-à-vis free will is thus anything but the noncommittal or intuitive position to which one can retreat in the absence of a plausible theory of free will. Such a stance would require us to make radical and counterintuitive changes to our practices, attitudes and even way of life – probably more so than any theory of free will, including one that incorporates the unpredictability condition.

---

<sup>92</sup> Eg D Pereboom, *Free Will, Agency, and Meaning of Life* (OUP 2016).

<sup>93</sup> This list is based on Clarke (n 21), 7.