



CAN SELF-DETERMINED ACTIONS BE PREDICTABLE?

AMIT PUNDIK

Tel Aviv University

Original scientific article – Received: 30/05/2019 Accepted: 15/09/2019

ABSTRACT

This paper examines Lockie's theory of libertarian self-determinism in light of the question of prediction: "Can we know (or justifiably believe) how an agent will act, or is likely to act, freely?" I argue that, when Lockie's theory is taken to its full logical extent, free actions cannot be predicted to any degree of accuracy because, even if they have probabilities, these cannot be known. However, I suggest that this implication of his theory is actually advantageous, because it is able to explain and justify an important feature of the practices we use to determine whether someone has acted culpably: our hostility to the use of predictive evidence.

Keywords: *Free will, causation, objective probability, determinism, criminal responsibility, Dennett, prediction, Lockie*

1. Introduction

Some philosophers arrive at the free will question from an ontological starting point (for example, "What kind of freedom exists, if any?" or "What are its conditions?"). Others arrive from an ethical starting point (for example, "How should I treat myself or others when we fail to do what we ought to?"). By contrast, Lockie takes a refreshing epistemic stance — based on the forceful transcendental argument for libertarianism that his book presents — and questions how epistemic norms affect the arguments we can use to support metaphysical claims about free will. Here, I would like to focus on the specific account of libertarianism that he proposes: his theory of self-determination (outlined mainly in Chapter 9). I will examine

his account in light of a different epistemic question, that of prediction: “Can we know (or justifiably believe) how an agent will act, or is likely to act, *freely*?” I would argue that, when Lockie’s self-determinism is taken to its full logical extent, free actions cannot be predicted to any degree of accuracy on the basis of anything other than previous free actions, because even if they have probabilities, these cannot be known. While Lockie himself seems to accept the view that free actions may be predictable, I argue that this view cannot be accommodated with other parts of his theory, hence he needs to choose between freedom and predictability. Furthermore, I would like to suggest that this implication of his theory, that free actions cannot be predicted, is actually advantageous: it is able to explain an important feature of the practices we use to determine whether someone has acted culpably—particularly, though not only, in criminal trials.

I start by arguing that, to be epistemically warranted, predictions need to rely on causal generalisations. I then turn to Lockie’s self-determinism and examine whether the agent’s character traits, reasons, and objective probabilities, or maybe even the agent as a whole, may be used to anchor such causal generalisations. Lastly, I briefly explain the hostility of Common Law to predictive evidence and suggest that libertarian theories that renounce the idea that free actions have discoverable objective probabilities are able to account for this hostility.

2. Why Predictions Require Causal Generalisations

Inferences from a known to an unknown empirical fact involve a generalisation about types. Schauer, for example, holds that “the avoidance of generalizations is, with few or no qualifications, simply not possible at all” (Schauer 2003, 101). 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 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 to 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 from a known to an unknown empirical fact 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 (this claim is part of the Common Cause Principle, see Reichenbach 1999, 157-166; Arntzenius 1992). The causal relation can operate either directly or through a common cause. For instance, 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. I do not argue that, for the inference to be valid, it is necessary to specify the (direct or indirect) causal generalisation; I only argue that the existence of such a causal generalisation has to 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, without making any commitment about the existence or kind of causal connection between these types of fact. 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 correlation between the number of people who drowned by falling into a swimming pool during a given period of years and the number of films in which Nicolas Cage appeared, in that same period (Vigen 2015). 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 number of people who drowned from the number of Nicholas Cage films (or vice versa) in a year that is not included in the group of years within which the 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). 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 to identify which generalisations are spurious (for the various sophisticated methods that have been proposed, such as the Markov Condition and Bayesian Nets, see Williamson 2005). 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 wish that such absurd, albeit statistically-significant, correlations simply did not exist. But this wish relies on the assumption 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 challenge this argument using counterexamples in which an inference from a known to an unknown fact is made without presupposing a causal connection between the types of fact. For example, if there are ten balls in a jar, of which nine are blue, it might be possible to infer that the probability of a randomly-chosen ball's being blue is 90%, without presupposing any causal connection between “being in that jar” and “being blue”.

However, even if not all factual inferences require a causal connection to be presupposed, the inferences drawn in legal fact-finding almost always do. Denying an underlying causal connection is easier when the generalisation is extracted from a group of cases to which the specific case at hand belongs. It is important to note that the randomly-chosen ball is, itself, one of the ten balls in the jar. It might thus be possible to draw some inferences about it without presupposing anything about the relation between the types of fact. While such inferences raise a set of difficult problems (Hájek 2009), these differ in kind from those involved in drawing inferences from generalisations that do not include the case at hand. To infer the probability that a randomly-chosen ball will be blue from the proportion of blue balls in another jar, it is necessary to presuppose that there is some substantial relation between “being in a jar” and “being blue”.

And, again, if this substantial relation is not causal or causal-like, what else could it be?

3. The Predictability of Self-determined Actions

Instructive hints on the predictability of self-determined actions can be found in Lockie's discussion of Dennett's character-based example of a person who is unable to torture an innocent for \$10 (Lockie 2018, 216). Let us ignore Milgram's experiments and assume, with Dennett, that this is, indeed, a paradigmatic example of an action (or omission) that is determined. Dennett uses this example to argue that the fact that the person's actions are determined (in this example, by his moral nature) does not undermine freedom. Dennett seems to rely on an intuition that this person acts freely when he does not torture. I do not share this intuition. My view is that the person's omission, if so determined, is *unfree* because they had no reason to torture (and it therefore seems to me that they deserve no praise for this omission — though I will not pursue this point here). Lockie, however, agrees with Dennett that the person's omission to torture for \$10 is free. He explains: "I may be unable to deviate from a path that I, my moral nature, has determined" (*ibid.*). However, he insists that Dennett's example fails to establish compatibilism, because the individual could still be free even if so determined: "Ethical responsibility [...] is something that requires freedom from determination by the Big Bang and laws of nature precisely in order to preserve the possibility of self-attributable axiological determination: of the agent determining acts in accordance with his moral nature and responsiveness to moral (and other) reasons" (*ibid.*).

I would argue that Lockie's agreement with Dennett cannot be settled with the rest of his theory, hence Lockie needs to accept that the person's omission to torture is unfree. More generally, I would suggest that Lockie has no theoretical resources to explain how actions may be both free and predictable based on anything other than the agent's own previous free actions. I would seek to establish this claim by discussing a related question: while Lockie discusses the matter of whether this omission could be both free *and* determined (by the person's moral nature), I would like to question whether this omission could be both free and *predictable*. I assume that, if any human actions and omissions are predictable, a person's omission to torture for \$10 must be one of them, even if nothing is known about the previous actions of that person. As I have already argued, this prediction must rely on a causal generalisation. Consequently, the question revolves around what the *relatum* of the causal connection that this generalisation reflects might be.

The first possibility is that the person's "moral nature" is the relatum, causing the agent to refrain from torturing an innocent. Moral nature itself could be the relatum, or it could consist of some propensities, traits, and so on, one of which determines the agent's omission. Whatever the exact relatum is, under this option, moral nature (or one of its components) is *ontologically distinct* from the agent. Such a view could easily explain the predictability of this omission: to have a certain moral nature either implies or consists of the *predictable tendency* of the agent to act in a certain way. For example, to be a kind-natured person is to have a higher likelihood of performing kind actions (compared to another person who is of an unkind nature). However, Lockie rightly rejects moral nature as the relatum that causes the agent's omission. He asks: "whatever the conative part was that determined your choice, was this in turn determined by natural law or by chance?", to which he answers: "persons, not their parts, determine choices" (*ibid.*, 196). Lockie makes this move to fend off Hobbes' regress objection: if some part of the person determined the choice, what determined that part? Consequently, Lockie objects to "an ontologically real, prior and separable item in the chooser called an act of will — historically: a 'desire', or sometimes 'volition'" (*ibid.*) As a result, according to Lockie, neither moral nature, nor any other part of the agent's character, can be the causal relatum on which predictions would be based.

The second possibility is that the predictability of the refusal to torture is based on the person's *reasons*. Perhaps these could explain why it is so predictable that they would refuse to torture an innocent for \$10. While their reason for doing so is weak (\$10), they have plenty of forceful reasons to avoid torturing. But Lockie seems to take the view that reasons are not causes: "For [reasons] to play a role, they don't associatively cause action and cognition. They enter a *mind* and become active in interaction in that mind. The mind (the agent, the person, the self) decides — in active assimilation, accommodation and equilibration of *that agent's* reasons" (*ibid.*, 207). Yet, if reasons have no independent causal power, they cannot constitute the relatum in the causal generalisations needed to make our predictions warranted.

Moving to the third possibility, perhaps the predictability of self-determined actions could be rooted in objective probabilities. One helpful way to understand objective probabilities, for our purposes, is as *free-standing ontological entities* (otherwise, it is unclear how they could be the relatum in the causal generalisations on which warranted predictions are based). Consider the predictability of the radioactive decay of a certain unstable atom. Let us assume that this atom, X, has a probability, Y, of decaying in the next second. Let us assume further that this is not a

subjective probability — even if we know everything that could possibly be known about this atom and the laws of nature, we would still be unable to predict *with certainty* whether this atom will decay in the next second. However, we can still predict that X will decay *with a certain probability*. After we observe enough cases of such atoms, we could generalise that the probability that X will decay in the next second is Y. The closer our prediction gets to the objective probability, the more accurate it is. This prediction is warranted because it is based on a causal generalisation in which objective probabilities are either the relatum or part of it: the objective probability is ontologically distinct from the other parts of the causal relations. It either causes the effect directly or it allows other potential causes to bring about the effect (or prevents them from doing so).

If free actions have such ontologically-distinct objective probabilities, they could warrant our predictions. If there is an objective probability that the agent will torture an innocent for \$10, and that probability is zero or close to zero, it could anchor our prediction. Under this view, predicting that the person is not going to torture is warranted, because it is based on a causal generalisation in which objective probabilities are part of the causal relatum that, together with the agent, determines the action. The person's inability to torture is basically a prediction that reflects this close-to-zero objective probability that they *would* torture.

However, if objective probabilities are such *free-standing ontological entities*, they cannot help make the agent's action self-determined. On the contrary, they would get in the way. Based on James, Lockie distinguishes between positive and negative chances. A positive chance is “a true generator of randomness in the world”, which is inserted “into, or over, or at, the origin of our acts” (*ibid.*, 197). Such a chance is “destructive of freedom and responsibility” because “responsibility for our actions as remained to us would be just that degree of determination of our actions as was robust enough to survive this chaos, this noise, these gremlins. The more noise, the less we would determine action — the less we could be said to act at all” (*ibid.*, 197-8). By contrast, a negative chance “is simply an absence of determination by any positive force *external to the agent himself*” (*ibid.*, 198, original emphasis). So Lockie rules out the type of chance required for rooting our predictions in objective probabilities, *at least when they are understood as free-standing ontological entities*.

Let us take stock: predicting the agent's free actions requires a causal generalisation, but neither the agent's character traits and propensities, nor their reasons, nor even ontologically-separable objective probabilities could be used as the relatum. The fourth and last possibility I can think of is that the relatum consists of the *agent himself*, as a whole. Being an

agent-causalist myself, I find this view very plausible — but the question is how such a view would account for predictions. Can such an agent have objective probabilities that are *integral* to them, making them who they are, rather than being “ontologically real, prior and separable items”? I would argue that, even if such objective probabilities were somehow possible, it would not matter for any practical purpose; in particular, it would not warrant our predictions. This is because such objective probabilities, even if they do exist, *cannot be known*.

When distinguishing the agent’s character from their reasons, Lockie emphasises that “[t]he person’s character is, in a significant sense, *ontologically unique*, prior and fundamental” (*ibid.*, 207). I take it that, when Lockie refers to “the person’s character” he means *the agent as a whole*, in an attempt to distinguish the agent from their reasons: “... what it is to have the character of the one isn’t just to be built up out of (‘bundled out of’) different, and different-strength, ‘reasons’ — it is to be ontologically different; it is to be a different person” (*ibid.*, 206). It is little wonder that Lockie emphasises uniqueness: if an agent is self-determined, and not subject to any natural law, what would be the basis for assuming that one self-determined agent will determine themselves similarly to another, if there is no external law to govern their conduct? But *if the agent, person, or character is unique, how could we predict what the agent will do freely, if we cannot draw any inference from observations about other similar people?*

One might respond that we could still predict that the person is not going to torture an innocent for \$10, based on his previous actions. However, this response does not help to reconcile Lockie’s agreement with Dennett with the rest of his theory, because Dennett’s example seems to work even in cases in which the prediction that the person will refuse to torture for \$10 is not based on their previous actions. I can step into my classroom on the first day of the academic year, knowing virtually nothing about the 150-or-so students there, and yet predict that they would not torture an innocent for \$10. If predictions are warranted only based on my knowledge of previous actions, then it is unclear how this prediction is warranted. More generally, setting aside cases in which a free action is predicted based on the agent’s previous actions, Lockie has no resources to explain how actions can be both free *and* predictable.

4. Is Unpredictability of Free Actions a Disadvantage for Lockie's Theory?

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* (van Inwagen 2000, 14-18; O'Connor 2000, 97; 2009, 197), and that position is rarely challenged (for exceptions, see Vicens 2016; Sela 2017). If Lockie's self-determined actions cannot be predictable, as I argued in the previous section, this implication of his theory might be viewed as a serious problem, even by those who are sympathetic to his libertarian inclinations. By contrast, my view is that accepting unpredictability as a necessary condition of free will may assist libertarian theories to overcome some of the common objections levelled against them.¹ In the remainder of this paper, I would like to suggest that accepting this implication of Lockie's self-determinism has the advantage of being able to explain an important feature of the practices used to determine whether someone has acted culpably — particularly, but not exclusively, in criminal trials.

While I believe that the following analysis is applicable more widely, to legal and non-legal practices of determining culpability alike, I focus here on the former because they include explicit and well-specified rules. I take legal practices in criminal trials to offer the most suitable case study because criminal punishment is clearly constrained by culpability, at least if criminal law seeks to avoid punishing those who are not culpable for their actions. 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.² Hence, criminal proceedings constitute the clearest context in which culpability is attributed. I also assume, like many theorists of free will, that acting freely or with some kind of control is a necessary condition of culpability. While some might hold that our practices of attributing culpability do not require us to settle the metaphysical problem of free will (Strawson 1962), I share the position that the distinction between justified and unjustified attribution of culpability — which any theory of culpability

¹ For example, unpredictability may assist libertarians to overcome van Inwagen's rollback argument. See Bernáth and Tózsér (2019).

² 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" (Hart 2008, 9).

seems to need — is likely to rely on (or bring through the back door) notions very similar to “freedom” and “control” (Tadros 2005, 69).

The scope of my discussion is restricted in one important respect. 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. I will therefore not discuss here evidence of planning, preparation, and motive (because, in such cases, the evidence may be probative of the alleged crime by establishing an earlier free decision that caused both the creation of the predictive evidence and the later commission of the crime).

When determining, in criminal proceedings, whether an individual performed a certain culpable action, predictive evidence is often ignored.³ Most apparently, and with only a few exceptions, base-rates are excluded (Koehler 2002). Using such evidence in court also seems *intuitively problematic*. 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 not be used at all in determining the individual’s guilt: that crime-rates be *inadmissible* in criminal proceedings.⁴ The hostility of criminal fact-finding toward predictive evidence is also apparent in the deeply-rooted suspicion of bad character and previous convictions.⁵

³ 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 (e.g. eyewitnesses, fingerprints), while the latter “looks forward from an established event or trait to predict the likely repetition of its occurrence” (Uviller 1982, 847).

⁴ 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 Pundik (2017, 192-193).

⁵ “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” (Roberts and Zuckerman 2010, 586).

Legal scholarship contains various accounts that seek to justify the exclusion of such predictive evidence. The first kind of strategy, which has received the 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 (Cohen 1977, 74); appropriate causal connection (Thomson 1986); case-specificity (Stein 2005, 64-106); ability to provide the best explanation (Dant 1988; Allen and Pardo 2008); immunity to the problem of the reference class (Allen and Pardo 2007); or sensitivity to the truth (Enoch et al. 2012).⁶ I am unconvinced by these epistemic accounts, because I think that not only does each one suffer from its own problems (Pundik 2008a), but they all share some common deficiencies (Pundik 2011; see also Schoeman 1987 and Redmayne 2008). 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.

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 (Tribe 1971), the over-transparency of standards of proof (Nesson 1985), equality between litigants (Stein 2005, 105), and the individuality and autonomy of the litigant against whom the evidence is used (Wasserman 1992; Zuckerman 1986). 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,⁷ 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,⁸ I have suggested a contextualist account that is based on culpability. According to this “culpability account”, some types of

⁶ The reference is to the epistemic explanation appearing in the first part of their paper, although, 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.

⁷ See Schoeman (1987). For criticism of Nesson and Tribe’s accounts, see Shaviro (1989). For criticism of Wasserman’s, see Pundik (2008b). For criticism of Stein’s, see, e.g., Pundik (2006).

⁸ This section rehearses the argument I made in Pundik (2017). 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.

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.

In Section 2, I argued that inferences about human conduct require reliance on causal generalisations. But, even if they do, why can free actions not 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 blotches 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 blotches 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 the blotches on his skin 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, 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 blotches 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 relation 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,⁹ 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 the blotches 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.

If Stephen does not possess the extra variable, inferring from his blotches 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 blotches that he is more likely to have acted violently than he would have been, had he not presented these marks, is therefore mistaken. In sum, this inference is either contradictory, because

⁹ For criminal law, see Alexander et al. (2009, 31); for tort law, see Perry (1995, 333-335); for health and safety regulation, see Adler (2005, 1247).

it requires inconsistent presuppositions, or it is 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.¹⁰ 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 a given*, was performed by Stephen or someone else. 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 the blotches on his skin that he is likelier to have behaved violently is, again, mistaken. Therefore, inferring from the blotches 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, such as the dangerous character of the neighbourhood, its socio-economic conditions, and so on. Such causal factors are outside the control of the individual resident. 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).¹¹ Alternatively, if the court seeks to avoid the implications of this inference, it ought to deem it

¹⁰ The discussion here is based on understanding the indeterminacy of the radiation as lying in the cause itself (Lewis 1986).

¹¹ 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.

irrelevant to the individual's conduct and exclude the evidence adduced to substantiate it.

The culpability account also supports common law's traditional 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 a comprehensive analysis of them here. 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.¹² While the admission of such previous convictions has been criticised on various grounds, such as being unconstitutional (Sheft 1995), unfair (McCandless 1997, 694), and even truth-suppressing (Cowley and Colyer 2010), 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. For example, 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 the culpability account, if these previous convictions are indeed probative, it might be at the price of exposing that the accused's conduct is unfree and thus nonculpable.

One might retort that my analysis stands in contrast to a common intuitive view of criminal responsibility. While the analysis implies that the agent's conduct is either fully determined or entirely unaffected, the practices of assigning criminal responsibility often seem to assume that an agent can be *partially* causally influenced. The agent is treated as causally influenced by some factor, but only to some degree, leaving them with a less-than-maximum extent of criminal responsibility. For example, a paedophile's sentence might be mitigated by the fact that he was a victim of molestation in his childhood. According to this view, the mitigation acknowledges that his childhood experience causally influenced the way he currently acts, yet it left him sufficiently responsible for molesting children now that he is an adult.

¹² 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.

The difficulty with this view of criminal responsibility is that it fails to account for the conviction stage of the trial, which seeks a binary outcome: the accused is either guilty of the alleged crime or not. Finding him guilty requires that he is culpable of committing the crime, which, in turn, requires that he acted freely. Free action is thus a precondition of criminal responsibility, and, when undermined by a defence such as insanity or duress, the accused is found *not* guilty rather than *less* guilty.

One means of explaining away the intuitive force of this view of criminal responsibility is to note that, while the question of guilt is binary, the consequences of conviction are typically scalar. The punishment could include a longer or shorter period of imprisonment or a heftier or lighter fine. It is at the sentencing stage that the paedophile's childhood experience is taken into consideration. However, there could be various explanations for why this experience serves to mitigate the appropriate punishment that make no reference to a partial causal influence. To mention just a few alternatives, there would be the increased effect that punishment would have on him as a result of his experience, his vulnerability to becoming a victim again during imprisonment, or maybe even the attempt to compensate him for his bad fortune in childhood.

Whatever the justification may be, it need not rely on a causal generalisation, according to which his childhood experience causally influenced him to commit the alleged offence. Moreover, if such a causal generalisation is used at the sentencing stage, it becomes difficult 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. The challenge here is not only to identify a solid objection to the use of such evidence in criminal trials (which is more difficult than it might seem), but also to explain why the same objection is not equally applicable at the sentencing stage. While exploring the justification for such mitigation lies outside the scope of this paper, it suffices to note that taking into account the paedophile's childhood experiences at the sentencing stage need not be based on his being less responsible for molesting the children he did. Therefore, my analysis does not stand in contrast to current sentencing practices. On the contrary, the "partial causal influences" view stands in contrast to our binary practices of *conviction*. Proponents of such a view would thus need to explain how freedom and criminal responsibility work, in their understanding.

5. Conclusion

Accepting that many free actions are necessarily unpredictable might be viewed as implausible and counterintuitive, even by libertarians such as Lockie. Yet, it seems that Lockie's self-determinism cannot be settled with the predictability of free actions. While I tend to think that, if free will exists, it is necessarily unpredictable, I did not pursue this claim here. Rather, I suggested that a theory of free will in which free actions are necessarily unpredictable is able to provide the sought-after justification for excluding predictive evidence. So even if Lockie's theory implies that self-determined actions are necessarily unpredictable, this might not be a bad thing after all ...

REFERENCES

- Adler, M. 2005. Against 'individual risk': A sympathetic critique of risk assessment. *University of Pennsylvania Law Review* 153: 1121-1150.
- Alexander, L., K. Ferzan, and S. Morse. 2009. *Crime and Culpability: A Theory of Criminal Law*. Cambridge: Cambridge University Press.
- Allen, R., and M. Pardo. 2008. Juridical proof and the best explanation. *Law and Philosophy* 27: 223-268.
- Allen, R., and M. Pardo. 2007. The problematic value of mathematical models of evidence. *Journal of Legal Studies* 36: 107-140.
- Arntzenius, F. 1992. The common cause principle. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1992: 227-237.
- Bernáth, L., and J. Tózsér. 2019. Rolling back the rollback argument. Unpublished manuscript (on file with author).
- Cohen, L. 1977. *The Probable and the Provable*. Oxford: Clarendon Press.
- Cowley, M., and J. Colyer. 2010. Asymmetries in prior conviction reasoning: Truth suppression effects in child protection contexts. *Psychology, Crime and Law* 16: 211-231.
- Dant, D. 1988. Gambling on the truth: The use of purely statistical evidence as a basis for civil liability. *Columbia Journal of Law and Social Problems* 22: 31-70.
- Enoch, D., L. Spectre, and T. Fisher. 2012. Statistical evidence, sensitivity, and the legal value of knowledge. *Philosophy and Public Affairs* 40: 197-224.
- Goodman, N. 1983. *Fact, Fiction, and Forecast* (4th ed). Cambridge, MA: Harvard University Press.

- Hájek, A. 2009. Fifteen arguments against hypothetical frequentism. *Erkenntnis* 70: 211-235.
- Hart, H. 2008. *Punishment and Responsibility* (2nd ed). Oxford: Oxford University Press.
- Koehler, J. 2002. When do courts think base rate statistics are relevant? *Jurimetrics Journal* 42: 373-402.
- Lewis, D. 1986. A Subjectivist's Guide to Objective Chance. In his *Philosophical Papers* (vol. 2). Oxford: Oxford University Press.
- Lockie, R. 2018. *Free Will and Epistemology: A Defence of The Transcendental Argument for Freedom*. London: Bloomsbury.
- McCandless, J. 1997. Prior bad acts and two bad rules: The fundamental unfairness of federal rules of evidence 413 and 414. *William & Mary Bill of Rights Journal* 5: 689-715.
- Nesson, C. 1985. The evidence or the event? On judicial proof and the acceptability of verdicts. *Harvard Law Review* 98: 1357-1392.
- O'Connor, T. 2000. *Persons and Causes: The Metaphysics of Free Will*. Oxford: Oxford University Press.
- O'Connor, T. 2009. Agent-Causal Power. In *Dispositions and Causes*, ed. T. Handfield, 189-214. Oxford: Oxford University Press.
- Perry, S. 1995. Risk, Harm, and Responsibility. In *Philosophical Foundations of Tort Law*, ed. D. Owen, 321-346. Oxford: Clarendon Press.
- Pundik, A. 2006. Epistemology and the law of evidence: Four doubts about Alex Stein's foundations of evidence Law. *Civil Justice Quarterly* 25: 504-28.
- Pundik, A. 2008a. What is wrong with statistical evidence? The attempts to establish an epistemic deficiency. *Civil Justice Quarterly* 27: 461-493.
- Pundik, A. 2008b. Statistical evidence and individual litigants: A reconsideration of Wasserman's argument from autonomy. *International Journal of Evidence and Proof* 12: 303-324.
- Pundik, A. 2011. The epistemology of statistical evidence. *International Journal of Evidence and Proof* 15: 117-143.
- Pundik A. 2017. Freedom and generalisation. *Oxford Journal of Legal Studies* 37: 189-121
- Redmayne, M. 2008. Exploring the proof paradoxes. *Legal Theory* 14: 281-309.
- Reichenbach, H. 1999. *The Direction of Time*. Mineola, NY: Dover Publications.
- Roberts, P., and A. Zuckerman. 2010. *Criminal Evidence* (2nd ed). Oxford: Oxford University Press.
- Schauer, F. 2003. *Profiles, Probabilities and Stereotypes*. Cambridge: Harvard University Press.
- Schoeman, F. 1987. Statistical vs. direct evidence. *Noûs* 21: 179-198.

- Sela, G. 2017. Torts as Self-Defense. DPhil Thesis. University of Oxford.
- Shaviro, D. 1989. Statistical-probability evidence and the appearance of justice. *Harvard Law Review* 103: 530-554.
- Sheft, M. 1995. Federal rule of evidence 413: A dangerous new frontier. *American Criminal Law Review* 33: 57-87.
- Stein, A. 2005. *Foundations of Evidence Law*. Oxford: Oxford University Press.
- Strawson, P. 1962. Freedom and resentment. *Proceedings of the British Academy* 48: 1-25.
- Tadros, V. 2005. *Criminal Responsibility*. Oxford: Oxford University Press.
- Thomson, J. 1986. Liability and individualized evidence. *Law and Contemporary Problems* 49: 199-219.
- Tribe, L. 1971. Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review* 84: 1329-1393.
- Uviller, R. 1982. Evidence of character to prove conduct: Illusion, illogic, and injustice in the courtroom. *Pennsylvania Law Review* 130: 845-891.
- van Inwagen, P. 2000. Free will remains a mystery: The eighth philosophical perspectives lecture. *Philosophical Perspectives* 14: 1-19.
- Vicens, L. 2016. Objective probabilities of free choice. *Res Philosophica* 93: 125-315.
- Vigen, T. 2015. Spurious Correlations. Online at: <http://www.tylervigen.com/spurious-correlations>. Accessed 12/16/2019.
- Wasserman, D. 1992. The morality of statistical proof and the risk of mistaken liability. *Cardozo Law Review* 13: 935-76.
- Williamson, J. 2005. *Bayesian Nets and Causality*. Oxford: Oxford University Press.
- Zuckerman, A. 1986. Law, fact or justice? *Boston University Law Review* 66: 487-508.

