

Federico Picinali

Generalisations, causal relationships, and moral responsibility

**Article (Accepted version)
(Refereed)**

Original citation:

Picinali, Federico (2016) *Generalisations, causal relationships, and moral responsibility*.
[International Journal of Evidence and Proof](#) . ISSN 1365-7127

© 2016 Sage

This version available at: <http://eprints.lse.ac.uk/65028/>

Available in LSE Research Online: January 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Generalisations, Causal Relationships, and Moral Responsibility

*Federico Picinali**

1. Introduction

In Chapter 4 of his book *Character in the Criminal Trial*, Mike Redmayne addresses a series of non-epistemic arguments against the use of bad character evidence. Mike is particularly successful in debunking arguments focusing on the defendant's autonomy. According to these arguments, it is morally wrong to prove the defendant's guilt by relying on behavioural generalisations¹ about groups of which the defendant is a member; in particular, if the defendant has previous convictions, it is morally wrong to rely on generalisations concerning the recidivism rate of people with previous convictions. Using this information – the arguments claim – is 'inconsistent with the law's commitment to treat the defendant as an autonomous individual, free to determine and alter his conduct at each moment.'² In particular, using previous convictions as just described amounts to treating the defendant 'as if her past conduct determines her present conduct'³ and is 'dismissive of his capacity to revise, or act against, his bad character.'⁴ While Mike is firm in denying that the use of previous convictions has a moral cost, his thoughts concerning the use of behavioural generalisations that do not express recidivism rates are more tentative.⁵

* Law Department, London School of Economics and Political Science. I am indebted to Jules Holroyd for the invaluable exchanges I had with her during the research for this paper and for her very helpful comments on a previous draft. I especially thank Amit Pundik for helping me to avoid misrepresentations of his view.

¹ With the expression 'behavioural generalisation' I am referring to generalisations concerning the behaviour of individuals. More precisely, behaviour must feature in the second term of the generalisation for it to be 'behavioural' – although it may also feature in the first term. Consider, for example, a generalisation stating that 'cyclists in London are very likely to wear helmets'. The first term is 'being a London cyclist'. The second term is the behaviour consisting in 'wearing an helmet'. Notably, the second term of a generalisation is its focus, being the fact 'in question'. The first term, instead, represents a 'starting point'. On the definition of 'generalisation' see Picinali (2012) at 199-201.

² Wasserman (1991) at 943.

³ Duff et al. (2007) at 114.

⁴ Ho (2008) at 300.

⁵ Redmayne (2015) at 77-86. An example of such generalisations used by Mike is: 'young men who have served in the armed forces are three times more likely to commit crimes of violence than members of the

Inspired by the early drafts of Chapter 4 – which I had the honour to read – I researched the question as to whether the use of these generalisations in criminal trials is morally problematic. As a result of this research, I published an article⁶ where I concluded that I had been unable to find or formulate a decisive moral argument against their use. At the time of writing the article, however, I was not aware of – and was not capable of formulating myself – a beguiling argument concerning the use of behavioural generalisations, formulated and defended by Amit Pundik in a recent research paper.⁷ Had I been aware of this argument, I would have certainly addressed it in my work; I am confident that the same holds true for Mike and his Chapter 4. Thus, in addressing Pundik’s argument here I intend both to follow the trajectory of Mike’s work and to provide additional support for my earlier conclusion.

In section 2 of this article I present Pundik’s argument. In section 3 I raise three general concerns with the argument. In section 4 I formulate some specific criticisms, based on the analysis of the essential structures of bad character inferences. In fact, these criticisms restate, and elaborate on, some of the points already advanced in section 3. The conclusion of the article is that the problems raised by Pundik’s argument, although intriguing, should not affect the use of behavioural generalisations in criminal trials.

2. Generalisations and individual freedom: where does the problem lie?

Consider a generalisation stating that ‘people with gambling debts are very likely to steal’. This generalisation is ‘behavioural’, in that it concerns the behaviour of individuals,⁸ e.g., whether or not they steal. Pundik premises his argument about behavioural generalisations with the claim that only generalisations reflecting a causal relationship should be used to draw inferences about the defendant’s behaviour. There are three important clarifications to make with regard to this premise. First, the notion of ‘causation’ that Pundik seems to

general public’ (at 78). Mike also hypothesizes generalisations expressing the statistical incidence of misconduct in a racial or an ethnic group. These generalisations are not about recidivism rates, as they do not concern a group of offenders.

⁶ Picinali (2015).

⁷ Pundik (2015).

⁸ For a more accurate definition of ‘behavioural generalisation’ see n 1 above.

appeal to is not the strong notion of 'but-for' causation. It is, instead, the weaker notion of 'necessitation' or 'determination', according to which an event – e.g. having gambling debts – causes another – e.g. stealing – if the latter is necessitated or determined by the former. For the presence of the causal relationship it is not also required that the latter event would not have obtained in the absence of the former. Second, the causal relationship thus understood may link the two terms of the generalisation – e.g., having gambling debts and stealing – or each term to a common cause, not appearing in the generalisation – e.g., the fact of being a prodigal person. Third, Pundik remarks that it is not necessary that we identify the causal relationship at play. However, for the generalisation to be used in a criminal trial, it must be such that it would be reasonable to expect that an underlying causal relationship exists. The generalisation in the example seems to meet this requirement, as it would be reasonable to expect that a causal relationship exists either between the fact of having gambling debts and the fact of stealing, or between an ulterior fact – e.g., being a prodigal person – and each term. If we didn't posit this requirement and claimed, instead, that the presence of a statistical correlation is sufficient, nothing would prevent the use of 'spurious' generalisations – such as an hypothetical generalisation reflecting the striking correlation between the number of people who drowned by falling into a pool between 1999 and 2009 in the US, and the number of movies that Nicolas Cage appeared in, in each of those years.⁹ Generalisations of this sort – Pundik points out – are not hard to produce, if one relies on a sufficiently large set of variables and a sufficiently large database. Indeed, it would be possible to come across spurious behavioural generalisations that may be relevant to criminal fact finding. However, these generalisations are called 'spurious' for a reason: we are unable to make sense of them; we cannot discern a meaningful connection between their terms, other than that indicated by the correlation. Therefore, it would be unsafe and, possibly, irrational for us to rely on them in order to draw inferences. Now, I am not interested in assessing Pundik's premise on causation, or

⁹ Cf. <http://www.tylervigen.com/spurious-correlations> (last accessed 27 October 2015). This ridiculous generalization may state that 'movies featuring Nicholas Cage increase the risk of accidental death by drowning in swimming pools in the US'.

discussing it further. I accept it for the sake of argument, because the more interesting claims lie ahead. It is sufficient for me to recognise that many of the behavioural generalisations mentioned in the scholarly debates presented in Mike's Chapter 4 seem indeed to have an underlying causal relationship.¹⁰

Criminal trials – Pundik claims – are chiefly concerned with attributing moral responsibility¹¹ to an individual. According to Pundik, in order for us to conclude that an individual is morally responsible, we must prove, or at least presuppose, that she was free to determine her own behaviour. This implies that we should not prove the individual's responsibility by resorting to generalisations that deny her freedom: to do so would be irrational. But what generalisations deny our freedom? This – Pundik remarks – will depend on what we posit as the necessary conditions for acting freely, that is, for exercising free will. Notoriously, incompatibilists argue that if the world were deterministic – in particular, if facts of the past together with the laws of nature entailed every truth about our future actions – we could not have free will. Among the incompatibilists are the 'hard determinists', who believe that *in fact* the world is deterministic. As a result, they conclude that we don't have free will. Pundik leaves this school of thought aside,¹² to concentrate only on the incompatibilist 'libertarians', who claim that determinism is, indeed, false. Compatibilist thinkers, instead, argue that even if our actions are determined by antecedent events we may still enjoy free will. The bottom line is that, depending on whether we are incompatibilists or compatibilists, the fact that our actions are not determined by antecedent events, respectively is or is not a necessary condition of free will.

¹⁰ Consider, for example, generalisations stating that 'people who committed burglary in the past are *n* likely to commit burglary again', or that 'people who are violent towards objects are *n* likely to be violent towards individuals', or that 'people who served in the armed forces are *n* likely to commit violent crimes.' In section 4 I will make some remarks concerning the causal relationships that seem to underlie these generalisations.

¹¹ Pundik uses the term 'culpability' rather than 'moral responsibility'. However – based on his paper and on our exchanges – I am confident that we mean the same thing. I prefer to use the latter expression, as in the literature the term 'culpability' is sometimes used to refer to *mens rea* elements only. Also, Pundik recognizes that his argument may not apply at all, or may not apply with the same strength, to crimes that do not track moral wrongs – e.g. some regulatory offences. All references to 'responsibility' in the paper should be read as references to 'moral responsibility'.

¹² In fact, a 'hard determinist' may reject the whole enterprise of criminal justice, at least to the extent that it is based on a retributivist rationale.

Now, consider a fanciful generalisation stating that ‘at sunset every person with a particular pattern of skin marks is subjected to an urge to attack everyone around her’.¹³ Assume that the generalisation is reliable and that a scientific study of the phenomenon has proven that both the pattern of skin marks and the urge are caused by the exposure to a toxic substance; thus, the generalisation is not spurious. John is tried for attacking Steve at sunset and the police discover that he has the relevant pattern of skin marks. The prosecution wants to rely on the above generalisation in order to infer that John attacked Steve. Given this scenario, Pundik makes two claims about the admissibility of the generalisation and of the related inference. 1) *If we were libertarians, we should admit this evidence only on the condition of acquitting John for not being responsible.* In fact, the generalisation and the inference presuppose that John’s behaviour was determined by an antecedent event. Therefore, if the inference is warranted – that is, if the generalisation applies to John’s case¹⁴ – John cannot have acted freely; thus, he cannot be responsible for the crime. If the inference is not warranted, this should be sufficient reason to exclude it, together with the underlying generalisation. 2) *If we were compatibilists and the evidence were such that ‘its use presupposes a causal factor which is one of the causal factors rendering behaviour unfree according to the respective compatibilist criteria,’¹⁵ then we should admit the evidence only on the condition of acquitting John for not being responsible.* Even though, as compatibilists, we would claim that indeterminism is not a necessary condition for free will, depending on our account of freedom we may maintain that some causal connections between antecedent events and individual behaviour deny the freedom of such behaviour. For instance – Pundik exemplifies – a compatibilist may argue that an agent has acted freely only if she would have acted differently had she wanted to.¹⁶ If it were the case that the causal

¹³ This is roughly the same example used by Pundik at 15 ff.

¹⁴ For the inference to be warranted it is necessary both that the underlying generalisation is reliable – which is the case *ex hypothesi* – and that John’s case falls within the group of cases to which the generalization applies. As I will point out later, concluding that the generalisation applies to the defendant’s case, does not entail the conclusion that the defendant acted as the generalisation states. This conclusion will depend on how high is the probability expressed in the generalisation. In John’s case, the generalisation being universal, if we conclude that it applies to his case we are also bound to conclude that he acted as the generalisation states.

¹⁵ Pundik (2015) at 19.

¹⁶ *Ibid.*

connection between the exposure to the substance and the urge to attack bypasses the desires/intentions of the individual – in other words, that the individual would attack others irrespective of her actual desires/intentions – then our compatibilist would claim that the generalisation and the inference are only admissible on the condition of acquitting. If the inference is warranted – that is, if the generalisation applies to John’s case – it shows that John was not acting freely under this compatibilist account of freedom and, therefore, that he was not responsible. If the inference is not warranted, this is sufficient reason to exclude it, together with the underlying generalisation.

Pundik also notes that the two claims hold regardless of whether the generalisation at issue is universal – e.g., ‘every person...is subjected to an urge’ – or not – e.g., ‘most persons...are subjected to an urge’. In fact, he stresses that his theory is mainly concerned with generalisations of the latter kind – given that this is the kind of generalisations normally used in criminal trials. Whether the generalisation is universal or not, the prosecution’s argument would necessarily imply that the causal connection was at play in the defendant’s case – in other words, that the defendant’s case falls within the class of cases where the connection is operative. Thus, it would be irrational to rely on the generalisation and the respective inference in order to prove that the defendant is responsible: if the causal connection is operative in John’s case, he has not acted freely and, therefore, he is not responsible. Finally, according to Pundik the two claims would apply even if the evidence were put forward by the defence; however in this case it would not be irrational for the defendant to rely on the generalisation were she aiming to show, precisely, that she was not responsible because she was not acting freely.

The elegance of Pundik’s argument rests in its being noncommittal as to whether determinism is true or false, and as to whether incompatibilism is right or wrong. It remains to be seen whether, aside from being elegant, it is also a convincing argument.

3.General concerns with the argument

Is Free Will Necessary for Moral Responsibility?

Although Pundik's argument is apparently noncommittal as to whether incompatibilism is right or wrong, it does presuppose that free will is necessary for moral responsibility. This is a contentious point, and one that is denied by prominent compatibilist theories. These theories proceed by 'confession and avoidance': libertarians are right in claiming that determinism and free will are not compatible; however, they are wrong in claiming that the truth of determinism denies responsibility. For these compatibilists – known as semi-compatibilists – it is conceptually mistaken to posit free will as a necessary condition for responsibility and, therefore, it would be conceptually mistaken to argue that there is an inconsistency with using generalisations that deny our free will in order to prove our moral responsibility.

Take, for example, Jay Wallace's compatibilist theory.¹⁷ Wallace defines free will as the 'availability of a range of alternate possibilities, holding fixed the laws of nature and the facts about the past.'¹⁸ This definition is by no means peculiar to his account; it is, indeed, widely accepted.¹⁹ Wallace agrees that free will cannot exist if determinism is true. And yet, he contends that the truth of determinism would not deny our responsibility, because the latter does not require free will. Responsibility, instead, requires the possession of two rational powers of 'reflective self-control': that of grasping and applying moral reasons and that of controlling behaviour in light of them – where the power to control behaviour involves a capacity for critical reflection, a capacity to make choices as a result of deliberation, and a capacity to translate choices into action.²⁰ As Wallace convincingly argues, these powers are not denied by the absence of alternative possibilities, which determinism entails.²¹

¹⁷ See Wallace (1994), in particular Chapters 5 and 6. Wallace builds on Strawson (1962), famously claiming that moral responsibility must be understood in the context of our practice of holding reactive attitudes in response to people's behaviour, and that this practice would not – and should not – be threatened by the possible discovery that determinism is true.

¹⁸ Id. at 3.

¹⁹ Cf. O'Connor (2010), stating at 1 that "'Free Will" is a philosophical term of art for a particular sort of capacity of rational agents to choose a course of action from among various alternatives.' Consider also the definition of 'free will' underpinning the famous 'consequence argument' in van Inwagen (1975).

²⁰ Wallace (1994) at 157-159.

²¹ Id. in particular at 147-153 and 180-186.

Take now John Fischer's and Mark Ravizza's theory.²² Fischer and Ravizza distinguish between two types of control over our agency.²³ 'Regulative' control consists in the freedom to choose between alternative possibilities. This 'freedom to do otherwise' – i.e., 'free will' according to the definition given above – is not required for moral responsibility. 'Guidance' control, instead, is necessary for moral responsibility. Fischer's and Ravizza's conception of guidance control is complex and I cannot do justice to it here.²⁴ However, giving a rough account of this conception is useful, as it helps the reader to appreciate some of the points that I will make later. According to Fischer and Ravizza, an agent exercises guidance control if her action issues from a mechanism – that is, a set of attitudes and of other 'agential characteristics' that are relevant to the action²⁵ – which is both owned by the agent and moderately reasons-responsive. The mechanism is owned by the agent if she takes responsibility for its operation – i.e., she recognises its potential effects on the world around her and she accepts that she may be the target of reactive attitudes as a result of her action. The mechanism is moderately reasons-responsive if, under various counterfactual scenarios where the agent is presented with new, coherent reasons, the mechanism would recognise and process these reasons, and possibly – based on one or more of them – would cause the agent to act differently from how she actually acted. The exercise of guidance control, thus understood, does not require the presence of free will – that is, the ability to do otherwise in the actual scenario.

Pundik's argument has no purchase with respect to these compatibilist accounts,²⁶ as it posits free will as a necessary condition for responsibility.²⁷ Considering the impressive influence on the academic debate and the apparent plausibility of both these accounts, the power of Pundik's argument seems considerably reduced by this realisation. Now, Pundik

²² See Fischer and Ravizza (1998).

²³ See *id.* at 30-34.

²⁴ See *id.* Chapters 2, 3, and 8.

²⁵ On this understanding of 'mechanism' see also McKenna and Coates (2015) at 47.

²⁶ Indeed, there are other prominent compatibilist accounts that reject the claim according to which the freedom to do otherwise is necessary for moral responsibility. See, for instance, Strawson (1962), Frankfurt (1971), and Arpaly (2002).

²⁷ To be fair, Pundik briefly discusses some of these accounts in footnote 57 of his manuscript. The discussion is, however, inadequate.

would have two simple strategies to respond to this concern. The first strategy would be to reassert that both of his two claims – addressed, respectively, to the libertarian and to the compatibilist – are about free will, however using a definition of ‘free will’ that is different from the common definition given above. It remains to be seen what this definition may be and how could it work both for the first and the second claim. The second strategy would be to reframe his second claim as follows: *If we were compatibilists and the evidence were such that its use presupposes a causal factor which is one of the causal factors that denies moral responsibility according to the respective compatibilist criteria, then we should admit the evidence only at the condition of acquitting John for not being responsible.* In other words, the strategy involves expunging the reference to free will,²⁸ so as to make the claim applicable also to theories such as Wallace’s, Fischer’s and Ravizza’s. Considering this strategy is useful, because the strategy emphasises a flaw of Pundik’s claim that was already present in its original version. Compatibilists have taken different views with respect to the necessary conditions for responsibility. Libertarians, on the other hand, at least agree that the presence of free will is one such condition – which gives to Pundik’s first claim its extensive reach and intuitive power. Given the variety of compatibilist theories out there and the many kinds of causal antecedents that may characterise behavioural generalisations – e.g., objects, mental states, medical conditions, actions, natural events, social conditions – it seems that both versions of Pundik’s second claim are general to the point that they may appear uninteresting. Unless the claim is made more specific – elaborating on the relevant conditions for responsibility and on the causal factors that would deny them – and/or an effort is made to find a common denominator between the main compatibilist theories, the purport of the claim is just too vague to make it a useful directive. Of course, the risk involved in specifying the claim is that of making it inapplicable to some prominent compatibilist accounts.

Insufficient Causes and Indeterministic Choices

²⁸ Here I am working under the assumption that when Pundik uses the attribute ‘unfree’ in his second claim, he means ‘without free will’. He gives no indication that the reader should interpret the term otherwise.

Pundik's first claim – addressed to the libertarian – is not immune to criticism either. The causal relationship underlying a behavioural generalisation may be such that the antecedent is not a sufficient cause: an additional causal factor is needed for the behaviour to ensue.²⁹ This factor may be precisely the individual's indeterministic – i.e., not entailed by antecedent facts – choice to engage in that behaviour. Indeed, since libertarians assume indeterminism, it is entirely consistent for them to recognise the possibility of an indeterministic choice on the part of the agent, at the same time positing the truth of a behavioural generalisation with an underlying causal relationship. Consider the generalisation stating that '95% of people with disease X act violently'. Assume that there is a causal connection between having the disease and acting violently. In other words, in 95% of cases the disease causes violent behaviour. This does not exclude that the behaviour would not ensue but for the individual's choice to engage in that behaviour. And, unless we are able to identify a sufficient cause for that choice, we must take the choice to be indeterministic, and thus *free* in the sense expressed above.³⁰ Notably, we could conceive of a scenario of this sort even if the generalisation were universal: what matters is not the probability expressed in the generalisation, but whether the antecedent to which the generalisation explicitly or implicitly refers is a sufficient cause.³¹ Now, under these circumstances many libertarians would have nothing to object to the use of the generalisation in order to prove the defendant's responsibility. Moral responsibility is not denied by the generalisation, nor by the inference built upon it, because one of the causal

²⁹ In fact, it does not matter for this objection whether the additional causal factor is always needed. What matters is that it may be necessary in the defendant's case.

³⁰ As explained later, for this criticism to succeed it is not even necessary that the choice be indeterministic. It would be sufficient that the deliberation preceding the choice involve an indeterministic mental event.

³¹ However, it would seem that this scenario is more likely to occur with probabilistic behavioural generalisations; indeed, that it occurs with any such generalisation, provided that the individual is not denied an opportunity to choose whether to act. The ulterior causal factor (i.e., the choice of the individual) may be conceived of as an additional relevant trait of the individual case. This trait would allow us to identify a more specific reference class than that of the original generalisation – i.e., the class of people with both the 'original' trait and the trait consisting in 'choosing to act'. Such reference class would yield a generalisation with a higher probability; possibly a probability close to 1 – given that under normal circumstances a choice to act brings about the action. Cf. Picinali (2012). A process of 'gradual specification' of this sort would not apply if the original generalisation were already universal. However, it is possible that also in this case there are one or more hidden causal factors that contribute to bringing about the relevant behaviour. In any case, it is doubtful whether universal behavioural generalisations exist.

antecedents of the agent's behaviour would be her indeterministic choice to act. The presence of a causal link between choice and action is considered unproblematic by many libertarians, when the choice is not itself entailed by causal antecedents. Notably, some libertarians would reach the same conclusion even if the choice were determined, provided that the preceding deliberation involves an indeterministic mental event – e.g., if it is undetermined what set of the agent's beliefs will come to her mind during deliberation.³²

It derives from this objection that Pundik's first claim is imperfect. In order to avoid the objection the claim should be rephrased so as to apply only if: 1) the causal antecedent of the relationship underlying the generalisation is a sufficient cause in the defendant's case; 2) or it can be excluded that an indeterministic choice to act is within the sufficient set of causes that brought about the defendant's behaviour *and*, if there is a deterministic choice within such set, that the deliberation preceding the choice involves an indeterministic event. The problem with this revised version of the claim is that its application requires an explanatory inquiry that the court – if anyone at all – is simply unable to undertake. True, there may be cases where we can reasonably exclude that the defendant made any choice or engaged in any deliberation at all – e.g., a case of automatism due to an epileptic seizure. In these clear cases a libertarian court may confidently say that Pundik's first claim applies. However, these cases are already covered by well-known defences, which casts doubts on the need to be guided by Pundik's 'metaphysical' claim in the first place. Imagine that the prosecution were so daft to argue that the defendant is responsible for assault because a generalisation, stating that 'people with an epileptic fit are almost certain to hit people around them,' applies to her case. Were the prosecution to advance such a self-defeating argument, there certainly would be no need for giving the defendant the additional safeguard of Pundik's first claim, nor would the claim add anything to our understanding of the defendant's situation. If the generalisation applies to the defendant, it means that she has acted as an automaton and therefore isn't responsible. What is more, outside the clear

³² For examples of libertarian theories that support the claims in this paragraph, consider the theories discussed in Clarke and Capes (2013) at 9-14, and in Levy and McKenna (2009) at 121-122. See also Kane (1996).

cases where we can confidently say that the defendant didn't make a choice or deliberate, Pundik's first claim could not be applied without tackling the very question as to whether our world is deterministic or not. Only if we concluded that it is, we may safely exclude the presence of an indeterministic choice on the part of the defendant or of an indeterministic event in her deliberation. Needless to say, this question is intractable, at least at the current state of our knowledge. Surely we don't expect the court or the fact finder in a criminal trial to make judgments about the truth of such metaphysical theses.³³

Pundik may contend that the objection mounted here is based on too fanciful a state of affairs – i.e., that the causal antecedent is not sufficient and that another causal antecedent is the indeterministic choice of the agent or a deterministic choice brought about by at least an indeterministic mental event. Therefore, Pundik may conclude that the objection does not pose any reasonable problem for its first claim. To this I would reply that states of affairs of this sort are far from fanciful, as I will show in section 4.³⁴ And I would add that, in any case, the debate about responsibility and free will is a cradle of fanciful hypotheticals and objections: one should not participate in this debate – or base her theory on it – expecting to be immune from such argumentative tools.

A Procedural Concern: Not an Argument About Admissibility

In the previous subsections I attempted to show that it is far from clear that libertarians and compatibilists would have difficulties accommodating generalisations of the sort that Pundik problematises. Now I would like to air an additional general concern with Pundik's argument, which refers to the role that the argument is meant to play in the criminal process of proof. According to Pundik, his argument concerns the question of admissibility. He maintains that if the prosecution advances an inference based on a behavioural generalisation in order to prove the defendant's guilt – and if the conditions presupposed by either his first or his second claim apply – the court should admit this evidence only at

³³ Finally, were we able to answer the question of determinism and were we to conclude that determinism is true, incompatibilists would argue that no one is morally responsible, not just the defendant in the example.

³⁴ See also n 31 above.

the condition of acquitting the defendant. If the court believes that the inference is not warranted – i.e., that the generalisation does not apply to the defendant’s case – it should not admit the evidence precisely on this ground. The same directives should apply when the evidence is presented by the defence in order to show that the defendant was not acting freely, the only difference being that in this case the defence would not be advancing an inconsistent argument.

Pundik does not seem to hold that the court should exclude the evidence depending on whether such evidence would be inconsistent with the case of the party advancing it – as it may be if the prosecution were such party. In other words, the decision on admissibility should not be governed by the goal to foster or preserve consistency in the party’s case. This, I believe, is a sensible position to take: it would be unreasonable to exclude useful evidence only because it does not help the case of the party producing it.³⁵ Pundik seems to posit a different criterion for the decision on admissibility: the evidence should be admitted only if it is found that the generalisation applies to the defendant’s case.³⁶ However, once the evidence is admitted on this ground, the die is cast: the defendant should be acquitted.

The problem with this approach is that it conflates the question of admissibility with the question of fact-finding, and – in the case of a jury system like the English and Welsh – the role of the court with the role of the jury. Were I to accept that a defendant may not be responsible if a particular behavioural generalisation applied to her case, I would still maintain that the question of admissibility does not involve drawing a conclusion as to whether the generalisation actually applies to the defendant’s case. Nor does it involve drawing a conclusion as to what are the consequences of this for the finding of guilt.³⁷ These judgments are part of fact-finding; in the jury system, they are not for the court to make.

³⁵ The regulation of disclosure may be taken as an indication that the consistency of each party’s behaviour is not a value protected by the system. Under section 3(1)(a) of the Criminal Procedure and Investigations Act 1996, the prosecution is required to disclose material that ‘might reasonably be considered capable of undermining’ its own case.

³⁶ In fact, Pundik’s claims leave room for the possibility that the court finds that the generalisation applies to the defendant’s case and yet decides not to admit it. It is not clear from Pundik’s theory what other factors may the court consider in reaching such decision.

³⁷ On this second point see n 14 above: claiming that a generalisation applies to the defendant’s case, does not yet mean that the defendant engaged in the behaviour stated in the generalisation.

The decision whether to admit or exclude the evidence should depend on other criteria (e.g., the apparent relevance of the evidence, its potential for prejudicial effect, fairness). Possibly, should the evidence be presented by an expert, a special test of reliability may be imposed. But it shouldn't be necessary for a party *to prove* that the generalisation applies to the defendant in order to have it admitted at trial; nor should the question of admissibility directly determine the outcome of the trial, lest we undermine the distinction between admissibility and fact-finding, and hence the division of labour between the court and the jury.³⁸ It is only with fact-finding that, considering all the available evidence, a decision can be made on whether the generalisation does or does not apply to the defendant's case and on what are the consequences that this judgment has for the finding of guilt. Moreover, it is important to consider that most of the examples used so far involved generalisations expressing high probabilities – in particular, probabilities that many would deem sufficient to meet the criminal standard of proof. It may well be the case that the parties present weaker generalisations. If so, concluding that the generalisation applies to the defendant's case may not be decisive.³⁹ We should also determine whether the respective probability meets the reasonable doubt standard – or, if the inference is presented by the defence, whether it is sufficient to raise a reasonable doubt.⁴⁰ These are determinations that normally fall within the province of the fact finder.

In light of these objections, a more convincing version of Pundik's theory may be devised for a jury system such as the English and Welsh. According to this version, the problem raised by Pundik is not to be addressed at the time of deciding whether to admit the evidence; it is to be addressed at the time of instructing the jury. If we accepted that a generalisation denies the responsibility of the group of people to which it applies – as shown above, it is far from clear that Pundik's argument is useful to identify generalisations

³⁸ This is not to deny that a court may sometimes heavily engage in fact-finding, for instance when it has to decide on an application for 'no case to answer'.

³⁹ As pointed out in n 14 above, claiming that the generalisation applies to the defendant's case merely means that the defendant falls in the group of people to which the generalisation refers. It does not entail that the defendant engaged in the behaviour stated in the generalisation.

⁴⁰ In fact, provided that Pundik's claims and their premises are sound, it seems that the crucial question is the latter: if the probability is sufficient to raise a reasonable doubt concerning the freedom of the defendant at the relevant time, the defendant should be acquitted.

of this kind – we may want the jury to be instructed that only at the condition of acquitting the defendant is it possible to conclude that the generalisation applies to her case.

4.Responsibility and the structures of character inferences

In this section I put forward more specific criticisms of Pundik’s argument, which restate, and elaborate on, points that I previously advanced in a general form. The aim of these criticisms is to reinforce the claim that, whether someone is a libertarian or a compatibilist, the generalisations with which Pundik takes issue do not deny moral responsibility simply in virtue of their status as generalisations that presuppose an underlying causal relationship.

As mentioned in the introduction, Chapter 4 of Mike’s *Character in the Criminal Trial* addresses two types of character inferences. The first type is that of inferences relying on generalisations expressing recidivism rates. The inference is between the defendant’s past conviction(s) and the criminal behaviour for which the defendant is tried. The inference relies upon a generalisation expressing the recidivism rate of individuals with past conviction(s) – where the recidivism rate would obviously refer to the criminal behaviour at issue. The second type is that of inferences relying on generalisations that do not express recidivism rates. Here the inference is between a particular trait of the defendant – the membership in a racial, ethnic, social group; the fact that she engaged in particular non-criminal behaviour; the fact of having a particular job, a particular medical condition etc. – and the criminal behaviour for which the defendant is tried. The inference relies on a generalisation expressing the statistical incidence of such criminal behaviour within the group of people sharing the relevant trait. Both types of inferences are ‘character inferences’ as they rely on evidence of previous misconduct on the part of the defendant and/or on information concerning the defendant’s disposition towards misconduct.⁴¹

In what follows I work with a different, but coextensive, partition of character inferences, that between inferences starting from the defendant’s previous behaviour – be it criminal or

⁴¹ Cf. section 98 of the Criminal Justice Act 2003.

not – and inferences starting from a trait of the defendant other than behaviour – e.g. her membership in a certain ethnic group.⁴² This partition will allow me to make some – admittedly tentative – remarks on the essential structures of character inferences and on the possible causal relationship underlying them. In light of these remarks I will try to show that it isn't at all granted that Pundik's two claims would apply to these inferences. However, given that the two types of character inferences that I have identified seem to exhaust the potential range of inferences to which Pundik's claims may apply, the question is raised as to what possible applications are left open. As I will argue, Pundik's claims seem to apply only to cases where there is no need for them, since the exculpatory work is done by other, more obvious, factors.

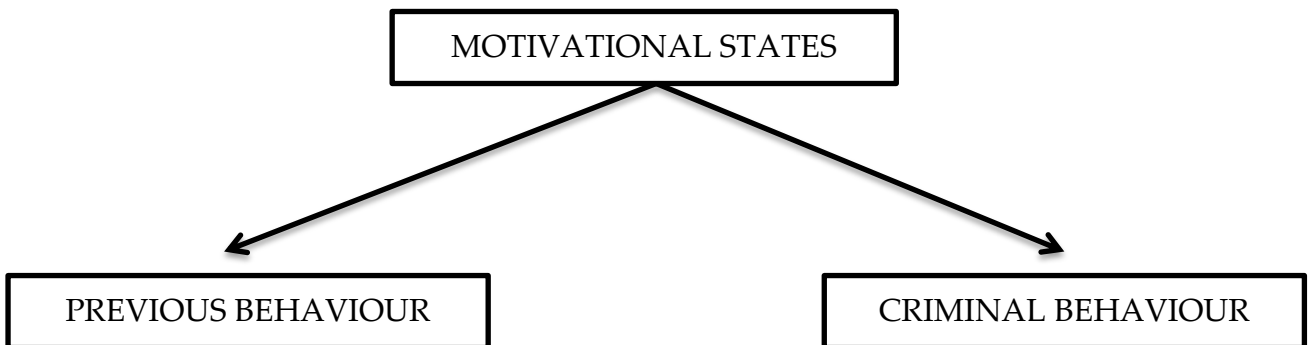
Inferences from behaviour

The defendant behaved in a particular way in the past – e.g., she committed crime X. She is now tried for committing crime Y. There is a reliable generalisation stating that '95% of people who committed crime X later commit crime Y.'⁴³ The prosecution relies on this generalisation to claim that the fact finder should infer from the defendant's previous behaviour that she committed crime Y. In 'inferences from behaviour', it seems reasonable to expect that if the generalisation is reliable, it reflects an underlying causal relationship. A plausible relationship is that between a common cause and each term of the generalisation, where the common cause consists in the individual's motivational states – e.g., her beliefs, desires, intentions. These states seem the best candidate to explain the previous and the subsequent behaviour – e.g., a previous and a subsequent sexual offence may be explained by reference to the individual's desire for sexual gratification and to her belief that sexual gratification is best achieved if the other party is not consenting.⁴⁴ The causal relationship is, therefore, *triangular*, and can be pictured as follows.

⁴² To clarify, being a member is not behaviour, while *becoming* a member is. I will say more on the relevance of this distinction later.

⁴³ The assumption, of course, is that the percentage is different (in particular, lower) in the general population. Otherwise the previous conviction would be irrelevant and the evidence would be excluded on this ground.

Picture 1



The question to ask now is whether Pundik's claims would have any purchase with respect to a behavioural generalisation of this kind. In other words, would compatibilists and/or libertarians argue that this generalisation and the respective inference deny the individual's responsibility? It seems obvious that compatibilists like Wallace, Fischer and Ravizza would answer the question in the negative, thus contending that Pundik's second claim does not apply to 'inferences from behaviour'. The presence of a causal link between the motivational states and the two terms of the generalisation does not deny their requirements for moral responsibility.

Consider the sketch of Wallace's conception of 'reflective self-control', given in section 3. The causal link between motivational states and behaviour does not deny the agent's grasp of the relevant moral reasons, nor the agent's capacity to critically reflect on these reasons, deliberate on them, make choices as a result of such deliberation, and translate these choices into action. After all, the agent's motivational states may be precisely the outcome of exercising the powers of reflective self-control. What is more, even if such states were themselves entailed by causal antecedents – as it would happen in a deterministic world – Wallace would still contend that reflective self-control could be exercised. Consider now the sketch of Fischer's and Ravizza's conception of 'guidance control'. The causal link between motivational states and behaviour does not deny that the action may have issued from a mechanism that is both owned by the agent and moderately reasons-responsive. In fact, the

agent may be taking responsibility for the operation of the mechanism – in the sense described above. Also, the causal link gives no reason to exclude that there may be a counterfactual scenario where, if the agent is presented with a new and coherent reason, the mechanism would recognise and process this reason, and possibly – based on it – cause the agent to act differently from how she actually acted. Again, these requirements may be satisfied even if the motivational states were themselves determined.

Libertarians too may answer the question in the negative, thus contending that Pundik's first claim does not apply to 'inferences from behaviour'. Even in the presence of a causal link between motivational states and behaviour, a libertarian may argue that responsibility for that behaviour is not denied if these motivational states are not themselves entailed by causal antecedents. More precisely, the agent may be responsible when, e.g.: it is undetermined what set of the agent's beliefs will come to her mind during deliberation; or the agent acts based on a preference or desire that is not itself determined; or the agent's intention is undetermined, because it results from a qualified 'effort of will' to resist the temptation to act.⁴⁵ The bottom line is that a libertarian too may remain unimpressed with Pundik's theory, as applied to 'inferences from behaviour'.

Inferences from a trait other than behaviour (for brevity, 'OB inferences')

The defendant is a member of an ethnic group, the Xs. She is now tried for committing crime Y. There is a reliable generalisation stating that '95% of Xs commit crime Y at least once in their lifetime.'⁴⁶ The prosecution relies on this generalisation to claim that the fact finder should infer from the defendant's membership in the ethnic group of the Xs that she committed crime Y. In 'OB inferences' such as this one, it seems reasonable to expect that if the generalisation is reliable, it reflects an underlying causal relationship. Given that the first term of the generalisation is not behaviour, a triangular relationship with the motivational states of the individual as a common cause seems implausible – after all, as the

⁴⁵ See the works referenced in n 32 above.

⁴⁶ Again, the assumption is that the percentage is different (in particular, lower) in the general population. Otherwise the fact of being an E would be irrelevant and the evidence would be excluded on this ground.

example shows, the first term may not be something that the individual can bring about. It is plausible, instead, to argue that the relationship is *linear*: the membership in the ethnic group causes certain motivational states that, in turn, cause the criminal behaviour.⁴⁷ Of course, the motivational states may be caused by the membership, together with concomitant antecedents – such as the social conditions that the ethnic group enjoys or endures. Here, I am making no claim that all ‘OB inferences’ involve generalisations with this underlying causal pattern. In particular, generalisations linking physical traits to behaviour – such as the fanciful generalisation in John’s case above – may well present a triangular pattern and/or may not be underlain by a causal relationship involving motivational states. However, the linear pattern just presented seems a plausible explanation for a vast array of generalisations relied upon in ‘OB inferences’: in particular, generalisations concerning membership in a racial, ethnic, national or social group.⁴⁸ The linear pattern can be pictured as follows.

Picture 2



Again, the question to ask is whether Pundik’s claims would have any purchase with respect to a behavioural generalisation of this kind. In other words, would compatibilists and/or libertarians argue that this generalisation and the respective inference deny the individual’s responsibility? Again, compatibilists like Wallace, Fischer and Ravizza would answer the question in the negative, thus contending that Pundik’s second claim does not apply to ‘OB inferences’. To see why, it is sufficient to restate a conclusion previously

⁴⁷ The fact that the motivational states are not expressed in the generalisation may give rise to a version of the ‘argument from autonomy’ that I have formulated, discussed, and criticised in Picinali (2015) at 5-7.

⁴⁸ It is plausible to claim that a triangular pattern – as the one described earlier – is involved in ‘OB inferences’ where the individual can choose to become a member of the relevant group: the motivational states may help to explain both the choice to become a member and the criminal behaviour. Possibly, in these cases there is a combination of a triangular and a linear pattern. I will briefly deal with these cases later.

reached: reflective self-control and guidance control are not denied by the linear causal chain, because they are not denied by the fact that the motivational states are themselves entailed by causal antecedents.⁴⁹

As far as libertarians are concerned, a negative answer seems again plausible. To begin with, it would be open to the libertarian to claim that if the membership in the group is the result of the individual's choice and if this choice is not itself entailed by causal antecedents, the linear causal chain linking membership to behaviour does not deny the individual's responsibility. For instance, according to Robert Kane an individual is 'ultimately responsible' for an action – e.g. the criminal behaviour – even if the action is causally determined – e.g. by the membership and the motivational states – as long as any determining cause is, or results, from an action of the agent that is not itself determined – e.g., the choice to become a member.⁵⁰ Moreover, in cases where the membership is the result of the individual's choice, it would be plausible to argue that the causal relationship is best captured by a triangular pattern,⁵¹ where the motivational states account for the decision to join the group and for the subsequent criminal behaviour. If so, the libertarian may adopt the very arguments available to her for the case of 'inferences from behaviour', that I explained above.

These solutions, however, would not work for cases where the individual has not chosen to become a member, possibly because she has been a member of the group since her birth – consider the case of racial, ethnic, national and social groups. As regards some of these groups – e.g. a national group – the individual may have the opportunity to opt out of the group later in her life. If so, the libertarian may argue that the agent is responsible for the

⁴⁹ So called 'historical compatibilists' may argue that it matters for moral responsibility how the agent came to satisfy the 'ahistorical' conditions for responsibility. In particular, it matters whether an individual has reflectively endorsed the motivational states or whether these states are not reflectively endorsed but inculcated as a result of abuse and indoctrination. See Levi and McKenna (2009) 108. This does not deny that the presence of a causal link between membership in a group and the motivational states is compatible with moral responsibility.

⁵⁰ See Kane (1996), in particular Chapter 3. Notably, Kane argues that '*[n]ot all* of our morally responsible choices or actions (those for which we are truly praiseworthy or blameworthy) have to be such that we could have done otherwise with respect to them directly' (at 40, italics in the original).

⁵¹ Or by a combination between the triangular and the linear pattern.

criminal behaviour if: the agent endorsed the membership by deciding not to opt out of the group or by not considering this option at all; and if the decision, or the failure to consider the option, are not themselves entailed by antecedent events. If these conditions obtained, Kane's requirements for responsibility would seem to be met again, because the causal chain that produces the crime would involve an indeterministic behaviour of the agent. As for groups that an individual does not have the opportunity to leave – e.g. a racial group – the libertarian may follow an argument presented in section 3. What if the motivational states caused by the membership were not a sufficient cause of the criminal behaviour? What if another causal factor were necessary for the defendant to engage in criminal behaviour? It seems perfectly reasonable to argue that, even though the membership in a group may cause an individual to have certain motivational states, these states alone cannot account for her criminal behaviour. After all, it is reasonable to believe that if the member of a particular group – especially a group that the individual belongs to since her birth and has no opportunity to leave – engages in criminal behaviour, this behaviour cannot be completely accounted for – if at all – by her membership. It seems also necessary that the individual choose to engage in criminal behaviour. This choice may be based on the motivational states caused by the membership, but need not be entailed by these states. In other words, the choice need not be determined by these motivational states; therefore, it may not feature in the linear causal chain. If so, it would be open to the libertarian to argue that the individual is responsible for the criminal behaviour, as long as her choice is indeterministic, or the deliberation preceding the choice involves an indeterministic mental event.⁵² Positing either of these facts is not a stretch for the libertarian, as she assumes that indeterminism is true.

What range of application is left for Pundik's claims?

As a result of the arguments advanced in this section, it would seem that Pundik's claims would not apply to character inferences. However, these inferences exhaust the potential

⁵² See n 32 above.

range of application of the claims: after all, they encompass all the inferences that involve a behavioural generalisation concerning criminal conduct. Therefore, one may conclude that no range of application is left for Pundik's claims. This would be too fast. As I pointed out earlier, the linear causal pattern pictured above may not capture the generalisations involved in 'OB inferences' such as the inference concerning skin marks that was at issue in John's case. If so, the arguments presented above may not apply to these inferences. The generalisations on which these inferences rely express a propensity to engage in criminal behaviour, but don't seem to be underlain by a causal relationship that involves the motivational states or the choice of the individual: the criminal behaviour is *directly* triggered by a medical condition or an external factor. When confronted with generalisations of this kind, it is likely that both the libertarian and the compatibilist would argue that the generalisation and the respective inference deny the moral responsibility of the individuals to which they apply.⁵³ In fact – referring again to Wallace's, Fischer's, and Ravizza's theories – if behaviour is directly triggered by a medical condition or an external factor, there is no room for either reflective self-control or guidance control. Moreover – for what we know of causal relationships of this kind – it would not be reasonable to argue that the medical condition or the external factor are not sufficient causes and that an indeterministic choice of the agent may be present as a concomitant causal factor. Therefore, both the compatibilists and the libertarians would probably conclude that Pundik's claims successfully apply to these cases. However, the important point to realise here is that in these cases we don't really need to delve into the intricacies of the free will debate in order to conclude that the agent is not responsible if the generalisation applies to her case. There is a solid and rather convincing body of law telling us precisely that the defendant is not responsible, or only partially so: this is the law on insanity, diminished responsibility, and automatism. The directives formulated by Pundik don't seem to add anything new to the directives that we already receive from this body of law,⁵⁴ nor does Pundik's theory seem to

⁵³ Provided that they express a sufficient probability.

⁵⁴ Apart from the fact that, according to Pundik, the problem raised by behavioural generalisations should be dealt with already at the admissibility stage. I have criticised this aspect of his theory in section 3 above.

improve our understanding of these cases. Focusing on free will and determinism unnecessarily complicates questions that we are capable of answering otherwise.

5. Conclusion

A couple of years ago, while preparing a seminar for the evidence course at LSE, Mike Redmayne and I came across a fascinating passage from Conan Doyle. The passage reads: ‘...while the individual man is an insoluble puzzle, in the aggregate he becomes a mathematical certainty. You can, for example, never foretell what any one man will do, but you can say with precision what an average number will be up to. Individuals vary, but percentages remain constant.’⁵⁵ Although the passage may sound like an overstatement, we have used it since in order to convey to the students that using generalisations in order to draw inferences about individual behaviour is a very complex matter. This task involves epistemic and possibly also moral problems, which have been examined with admirable sophistication in Mike’s *Character in the Criminal Trial*. I doubt that I will have a last word on this, but I don’t think that the task of drawing inferences about individual behaviour involves the problems that Pundik has raised.

Bibliography

Arpaly N (2002) *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford: OUP.

Clarke R and Capes J (2013) Incompatibilist (Nondeterministic) Theories of Free Will. *Stanford Encyclopedia of Philosophy*. Available at <http://plato.stanford.edu/archives/spr2014/entries/incompatibilism-theories/> (last accessed 18 September 2015).

Conan Doyle A (2003) *Sherlock Holmes. The Complete Novels and Stories*, New York: Bantam Classics.

⁵⁵ Conan Doyle, (2003) at 202.

Duff RA, Farmer L, Marshall S, Tadros V (2007) *The Trial on Trial: Volume 3. Towards a Normative Theory of the Criminal Trial*. Oxford: Hart Publishing.

Fischer JM and Ravizza M (1998) *Responsibility and Control: An Essay on Moral Responsibility*. Cambridge: Cambridge University Press.

Frankfurt HG (1971) Freedom of the Will and the Concept of a Person. *The Journal of Philosophy* 68: 5.

Ho HL (2008) *A Philosophy of Evidence Law: Justice in the Pursuit of Truth*. Oxford: OUP.

Kane R (1996) *The Significance of Free Will*. New York: OUP.

Levy N and McKenna M (2009) Recent Work on Free Will and Moral Responsibility. *Philosophy Compass* 4: 96.

McKenna M and Coates DJ (2015) Compatibilism. *Stanford Encyclopedia of Philosophy*. Available at <http://plato.stanford.edu/archives/sum2015/entries/compatibilism/> (last accessed 29 October 2015).

O'Connor (2010) Free Will. *Stanford Encyclopedia of Philosophy*. Available at <http://plato.stanford.edu/archives/fall2014/entries/freewill/> (last accessed 29 October 2015).

Picinali F (2012) Structuring Inferential Reasoning in Criminal Fact-Finding: An Analogical Theory. *Law, Probability & Risk* 11: 197.

Picinali F (2015) Base-Rates of Negative Traits: Instructions for Use in Criminal Trials. *Journal of Applied Philosophy* doi: 10.1111/japp.12109.

Pundik A (2015) Freedom and Generalisations. (unpublished manuscript, draft of 16 July 2015). The abstract is available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2556690 (last accessed on 28 October 2015). The paper is forthcoming in the *Oxford Journal of Legal Studies*.

Redmayne M (2015) *Character in the Criminal Trial*. Oxford: OUP.

Strawson PF (1962) Freedom and Resentment. *Proceedings of the British Academy* 48: 187.

van Inwagen P (1975) The Incompatibility of Free Will and Determinism. *Philosophical Studies* 25: 185.

Wallace RJ (1994) *Responsibility and the Moral Sentiments*. Cambridge (MA): Harvard University Press.

Wasserman DT (1991) The Morality of Statistical Proof and the Risk of Mistaken Liability. *Cardozo Law Review* 13: 935.