

imagining motives

Adam Morton

late draft of paper for del Mar & Amaya *Virtue, Imagination, and Emotion in Legal Reasoning*

pagination, some references, and the occasional wording in the main text are likely to be different in the printed version

abstract: We want legal systems to treat people in terms of the attributes they actually have, as opposed to those they claim or those that others attribute to them. But what we know about human attributions of mental states encourages a scepticism about this process. Salvation from scientific psychology seems both a long way off and rather undermining of our institutions if it ever arrives.

introduction I shall explore a tension in legal procedures between two ideals. One is *impartiality*; people should be treated 'the same', in accordance with general principles expressed in terms of morally and legally relevant objective impartial concepts. The other is *accuracy*; people should be treated 'differently', in accordance with real and relevant differences between their situations, rather than differences or commonalities that are ascribed to them in error or from prejudice. Both are aspects of the ideal of fairness. I am not going to state these principles in a careful or sophisticated way. My concern is with the application of the concepts involved, and in particular with the ascription of states of mind, including motives and emotions. Can we grasp the minds of others, for example by a kind of imagination, well enough to see at once how they fall under the same principles and how the details of each person are different? I shall be defending a sceptical attitude to this task. I shall argue that in a number of legal contexts our capacities to describe general attributes of people and our capacities to grasp what is individual about them do not mesh. In some of these contexts the form that this takes is that we are just not capable of applying the relevant concepts, the ones that would link a particular individual with a general principle, accurately enough. We are not equipped to live up to both of these ideals at once.

A regime of strict liability, where individual motives and constraints are ignored because a bare description of the legally relevant aspects of your act is all that counts, is wonderfully impersonal, but in many cases unfair for reasons of inaccuracy. A regime that takes the slightest differences between people as bases for differing ways of excusing apparent culpability *may* be wonderfully fair, but imposes a need for accurate assessment onto its participants. It is easy for people to claim that their situation is special, and tempting for those who identify with them to suppose that it really is. One might think

that the tension was easily resolved: accurate attributions of legally relevant differences should allow equal treatment of people who are alike in legally relevant ways, ignoring irrelevant differences. But this easy resolution is complicated by my second theme. Our usual ways of attributing states of mind to people and anticipating their actions are very imperfect. They form a loosely connected bundle of disparate resources which we tie together with wide labels of 'empathy' and 'imagination'. But each strand in the bundle gives accurate results only under very particular circumstances. It is far from clear that these include many of the circumstances where they could be legally useful. So the difficulty of getting any guarantee of the accuracy of ways of attributing states of mind to people makes it harder to find a compromise between impartiality and fairness. The conclusion is that the competing demands on our resources produce a grave and intractable difficulty in legal procedure, which we should not gloss over. Schematically, a scepticism about the accuracy of imagination leads to a pessimism about the fairness of our practices.

imagination and attribution

Since human beings are profoundly social creatures, we incorporate routines for anticipating one another's actions and their tendencies to play various roles in various shared projects. Given a normal human upbringing, normal humans tend to know pretty accurately when others are sad or angry, and whether they are likely to cooperate or obstruct many-person activities. It is not obvious in detail how we do this, and there was a lively interdisciplinary discussion between philosophers, developmental psychologists, and anthropologists about this in the last quarter of the twentieth century. But all parties agree that the innate human capacities that develop in children as they master the social routines of their cultures allow them to attribute states of mind to others in the same cultures.¹ Beyond that point there is less agreement.

There is an optimistic and a pessimistic way of interpreting these developments. The optimistic way focuses on our innate but tunable capacities to make sense of one another. On this picture we have in our heads software that meets the needs of human social life.

The pessimistic way also stresses our pre-equipped nature. But it expresses doubts about how fitted for the task, especially for the contemporary form of the task, our innate equipment – biologically recent but formed by ways of life that no longer exist – might be. As background to this, note that until the last quarter of the twentieth century and in some circles much later, it was generally accepted that humans understand other humans by putting them on a grid of rational thought and behaviour. Humans are most

1 I have written a handbook article surveying the field (Morton, 2009). I do not there discuss my less consensual suggestions, which are found in Morton (2001) and Morton (2007). See also Gopnik and Meltzoff (1997).

intelligible when they are being rational. This relic of the older view, that there is a discrete faculty of rationality whose presence marks humanity, had been sublimated in philosophy to a view that by treating people *as if* they were fully rational – whatever that means – we get our best predictive grasp over them.² This assumption was common to a lot of economics and philosophy of mind and social theory. The central reason for its failure is the awareness of ubiquitous sub-rational behaviour in human beings, which might have been evident from the psychoanalytic tradition but became mainline with the work of Kahneman and Tversky.³ Doubts about the exact content of the rationality claim also took their toll. Is it rational, for example, to be risk-averse, or does rational choice for every person consist in maximising the expectation of each action according to her probability and utility functions? That would result in predictions very different from what we actually do, and might be socially disastrous. The claim is also unhelpful about the effect of our limited power. Nearly all people at nearly all times are in situations where a vastly more intelligent creature would do something very different. Is nearly all of our behaviour thus inexplicable?⁴

With this in mind, consider the competing approaches to inbuilt behaviour-predicting capacities. Nichols and Stich (2004) contrast information-rich and information-poor accounts of folk psychology. Information-rich accounts postulate a source of articulated principles to guide attribution, understanding, and prediction. The principles may not be formulated in spoken language or available to conscious reflection, but they guide our more explicit judgements. They may originate in innate information-processing routines, reflection on the actions of others, or innate constraints on explanations. According to information-poor accounts, on the other hand, the fundamental features of our thinking about mind are determined by routines and capacities which do not embody any assumptions about how one state of mind leads to another.

Earlier versions of these two kinds of account referred to information-rich accounts as the ‘theory theory’. The idea was that children readily develop an implicit psychological theory from which they can derive predictions as they need them. Although six-year-old children cannot do astronomy – they can hardly articulate a description of the passage of the sun through the sky – on this account they possess something from which they derive predictions and explanations of human actions in a way that is analogous to the easier task of getting a prediction of an eclipse from an astronomical theory. Some of the arguments in defence of this account emphasise how it can include specifications of innate cognitive and motivational biases. The alternative is to make the theory in

2 Davidson (1982).

3 Kahneman and Tversky (1979), Kahneman (2013).

4 See ch 2 of Morton (2002).

question one of perfect rationality, which is a nonstarter for predicting the individual behaviour of one's neighbours, or for that matter of oneself.⁵

Information-poor accounts, on the other hand, were referred to as simulation theories. Their characteristic assumption was that faced with another person's actions, one experiments to discover what settings in one's own mind would result in similar actions, and then attributes them to the other. It is tempting to label these procedures as imagination: one imagines getting oneself to do the action and takes the other to be in the state that one has imagined for oneself. At this point, the contrast between two versions of this approach becomes relevant. The first is that of Heal,⁶ which applies best to predictions of other people's intellectual output. The idea here is that when you want to know how someone else would solve a problem, you solve it yourself and predict that they would do what you would do. I have no doubt that this is a strategy that we sometimes use, though its limitations are fairly clear. The other is my own approach, on which people are hardwired to enter into small-scale cooperative activity, and will act on the assumption that others will do the mutually advantageous thing.⁷ (Of course there are uncooperative and downright malign actions that this says nothing about.) And its direct application is limited to prediction of behaviour rather than prediction of thought or feeling.

While these lines were being developed and refined, experimental work on children's acquisition of adult concepts proceeded fruitfully. The central strand of this work – which stems from papers by Perner in the dying days of explanation as making rational – concerns small children's attribution of beliefs, and to some extent desires, to others.⁸ The famous false belief task plays a big role here. As a result we do have a fairly good timetable of how the capacity to attribute central cognitive concepts develops.

So different theories of our grasp of one another emphasise different capacities – different *kinds* of capacities – which are directed at different mental states: thoughts, emotions, actions. That suggests a deep question. How are these different parts of 'folk psychology' tied together? Are they fragments of a comprehensive predisposition in people for interpreting one another's actions, or are they independent of each other? We do not know the answer. I have defended the second position, which I suspect is the minority view, in a paper arguing that people learn to improvise connections between a

5 'Individual' because sometimes the differences between different people cancel or compensate out so that anticipating the average behaviour of large numbers of people is easier than predicting particular acts of particular people. Perhaps this is a factor that makes classical economics possible.

6 Heal (1998).

7 Morton (2002), especially chapters 3 and 5.

8 Perner (1991).

number of innate capacities.⁹ It is not at all clear that this view is correct but it needs to be explored. (Though not to the present point, such a view would allow more room for individual and cultural variation.)

Two themes from this history will be particularly relevant in what follows. The first is the tribulations of the concept of rationality. In philosophy, psychology, and economics, it is much less foundational than it once was. The other concerns the imaginative capacities involved, the ability to produce in one's own mind what is going through someone else's. Since the days when theory of mind/folk psychology/mind-attribution was a new and hot topic, there has been a flowering of interest in imagination, empathy, and sympathy.¹⁰ I take it that the earlier interest has led to the present one. But the ambiguities and uncertainties remain. There is no consensus about the nature of the imaginative capacities, or about their importance in our understanding of one another, and a widespread doubt that rationality can play the role that was once given, without any confidence about what if anything can replace it.

Legal systems rely on the ability of judges, barristers, and jurors to attribute motives, assess contributions to shared projects, distribute blame and negotiate matters of common concern. Without the ability to do these things, the application of law and judicial judgement, from earlier times to its current versions, would be crippled. We may be at point of transition, though. In the past we had a simpleminded trust in everyday capacities for attributing states of mind, at any rate as used by reflective reasonable people. Now, however, we are becoming aware of how problematic they can be.¹¹ An analogy is the naïve reliance that until recently was placed on eyewitness testimony. We thought that the only reason such testimony can be false is deliberate dishonesty or self-preservation. But the work of Loftus and others has convinced many that context and deliberate manipulation can influence beliefs that present themselves as simple direct recollection.¹² I strongly suspect that something similar is true of the imaginative attribution of motive and emotion. I do not have solid psychological evidence for this. So what I shall do is to describe some ways that the law, as seen by this particular outsider, has a problematic dependence on these skills. To the extent that they are unreliable, the institutions in question have a problem they should face.

Throughout their history legal institutions have relied on everyday capacities to do these things, in deliberate reflective form aided by specific devices, for example to encourage people to tell the truth by threatening them. So the connection between law and 'folk psychology' ought to be an important and central topic. I shall assume that all

9 Morton (2007).

10 Two comprehensive anthologies are Maibom (2017) and Coplan and Goldie (2011).

11 See Levesque (2006) which argues for the relevance of experimental psychology to legal practice. I owe this reference to Maksymilian Del Mar and Amalia Amaya.

12 Loftus (1979).

people interacting through legal systems apply the mind-attributing skills they have acquired through normal social life to those they interact with. I shall assume that the skills can be described not too misleadingly as exercises of the imagination. And I shall assume that legal procedures have the use of explicit evidence-based psychological theories to the extent that these are seen as relevant (which I am told is largely confined to cases where there is a mental health element to a defence, or the like). The crucial facts are, first, that we have no substitute for these skills of psychological imagination – for the capacity to represent states of other people’s minds in one’s own intuitively and more or less vividly, when faced with a large range of puzzling behaviour. It is a flawed capacity in obvious ways but it is often the best we have, and scientific substitutes for it may be generations away (or may never be usable by human beings in an informal practical way). Second, what we get from everyday psychological imagination – this informal imaginative capacity – generally does not combine well with what we get from psychology, neurology, or other sciences. Generally speaking, when we begin with an intuitive grasp of someone’s action we continue it in the same intuitive terms.

We do not do a lot of conscious reasoning about these things, even in the law. We get very quickly to the judgement that someone acted out of jealousy (an emotion that is typically based on a thought – that someone has gained what is rightfully one’s own – which is itself based on a standing character trait or an occurrent emotion), or malevolent lust (an emotion, inasmuch as it has an object, otherwise a very controlling mood), or desire for profit (a traditional motivating state), and then we move on to judging the moral quality of the action. Once we have done this, we can backtrack a little and consider excuses or exculpating facts, which typically retain the moral judgement of the act and qualify its implications as a judgement of the person. When we do this we normally do not consider *what* the person has done or even *why* they have done it, but *how* they could have done it – how they overcame various obstacles to heeding the motive in question. For example, it might be obvious that someone poisoned their grandmother in order to inherit a fortune, but far from obvious how they managed to overcome their affection and family feeling. This gives an explicit focus for imagination. There are many motives available to most people in most situations, and we have very little beyond platitudes to tell us which motive a person will be seized by, and how they will choose the means to the ends that it dangles before them – why the murderous grandchild acted in accordance with the chance of gain rather than respect for someone familiar from childhood.¹³ So we approach this task with an improvised variety of techniques, sometimes by trying to imagine ourselves into a similar position.

Excuses are then an important context for imagination, including neglected excuses – ones that it never occurs to us to apply in particular cases. For our purposes questions of

¹³ These are issues discussed in the literature on evil, especially with authors concerned to avoid the image of mysteriously irredeemable villains nearly all of whose actions are terrible. See Card (2002) or Morton (2004).

accuracy are important: whether the excuse that we are inclined to apply when things are presented in a particular way does in fact connect with the psychology of the person concerned. This takes us into largely unexplored territory. Most work on our capacities to attribute states of mind focuses on their effectiveness. This is not surprising given that the background in philosophy traces to 'other minds' problems; sceptical doubts that we have adequate evidence for taking people to have the beliefs, desires, emotions, and the like, that are normally attributed to them. And in the culture in general during these decades scientists were reconsidering the naïve empiricism of earlier generations in the light of the psychological possibility that much of the way we think is shaped by our genetic inheritance. But talking about excuses, and the attribution of motive in a legal context more generally, possibilities of misattribution – getting things wrong – are at least as relevant.

But attention to this side of attribution is not much studied.¹⁴ The importance of issues about accuracy seems to have been missed, probably to a large extent because accuracy is so hard to assess.¹⁵ As I note in the concluding section of this chapter, there is a common sense way of gauging accuracy in everyday life. It consists in interacting with the person and seeing whether one's expectations are fulfilled. But such interaction is not only not available but explicitly forbidden in many legal procedures, for the sake of fairness.

The conclusion of this section is that what we have learned about human capacities to attribute states of mind is rather fragmented. We have competing accounts, each of which focuses on some particular class of states of mind, and very little guidance about how to glue these together. We also have a very rough concept of imagination without much assurance of its accuracy or the conditions under which what we imaginatively ascribe to someone does capture the way they actually are. This gappy patchwork invites a sceptical response.

the legal connection

The problem of accuracy arises when we need something analogous to an everyday grasp of another person, but also need judgements that can stand up to the scepticism that

14 There is a literature on scientific substitutes for intuitive judgment on these topics. See Goldstein, Morse, and Shapiro (2003) and Levesque (2006). However, there is not much on the (in)accuracy of pre-scientific judgment on them. Undermining this point slightly, there is a precedent for studies of attributional error in influential work on infants, associated with the false belief task, where toddlers have surprising difficulty grasping the fact that people operate on the basis of wrong information (Gopnik and Meltzoff, (17)).

15 Intuitive imaginative judgements about what people generally think or feel are easier to test. For example, people generally expect that victims of violent crimes or their families will feel retributive emotions, but surveys suggest that this expectation is based on faulty imagination.

fairness and impartiality require. Here is a list. I will not discuss any of them in detail, let alone engage with the vast specialised literature on each.¹⁶

provocation

Acts are more culpable if premeditated, on standard doctrine.¹⁷ There are situations that we can easily imagine to produce rage or vindictiveness, and the actions that result are not taken to indicate a standing intention to do harm. Husbands discovering their wives with a lover are no longer excused from murderous rage. But some degree of annoyance or upset would hardly be surprising. Suppose H comes upon W in the arms of L, and as a result slaps L or W, or scratches L's car, or writes a damning letter of unrecommendation in connection with a job application. As a result, H finds himself in court, charged with assault or criminal damage or professional misconduct. He is likely to be found guilty, but the sentence is likely to be minimal. In this case the decision is in the hands (mind) of a judge. (There is no law against writing vicious letters about a job candidate. But it is easy enough to imagine a situation where there is a legal aspect to a professional misconduct or harassment case. Some letters of 'recommendation' might result in libel suits.)

It may seem that there is no problem here. We can all think ourselves into these shoes and can imagine losing our tempers. But. What we imagine ourselves into may not resemble H's state at all closely. Perhaps instead of rage he felt calm, or sadness, or relief, and perhaps he acted out of professional rivalry, a calculated strategy for some other purpose, or just plain nastiness.¹⁸ These things can be discovered, and given enough other evidence of a familiar psychological sort, they would be likely to undermine the plea. But this further evidence is unlikely to emerge in court, and would require anecdotes and descriptions of behaviour patterns, which would have an air both of irrelevance and of inadmissibility.

irresistible impulse

The idea is that some people have a lesser capacity to deny the urges that are common in the rest of us to strike out at enemies, take retribution for old injuries, or react to

¹⁶ I am not pretending to have more legal knowledge than the average informed citizen. So I shall not provide faked up erences to hastily digested legal works. I shall illustrate this below by the out- of-dat reference to Holmes (1882).

¹⁷ Is this intuitively correct in all cases? Contrast two killers. One person has for decades mused about ways of killing her abuser, an, when the opportunity presents itself, although she is a gentle soul she steels herself and takes it. The other person knows nothing of the victim, but sees a similar sudden opportunity for undetected destruction and takes it because the thought of a crushed skull appeals so evocatively. I think I would be more appalled by the second.

¹⁸ This is a theme of Sartre's play *Les mains sales*. Sartre's suggestion is that there is no answer to the question of why his protagonist killed his rival. It is all determined retrospectively, Sartre implicitly argues. There is a subtle germ of truth here, but most of us are determined to resist the line.

provocations that are not really that extreme. The reason is usually that the impulsive person is in some way damaged: some specific insanity, a deeply flawed personality, or diminished intelligence. The desired result is not full exculpation but a lesser guilt: manslaughter instead of murder in the most dramatic cases.¹⁹ So if you are foolish enough to tease a mentally handicapped giant about her infantile clothes, you should not be surprised at the result but will find your hospital rest disturbed by news of her negligible penalty.

This is one of the less problematic cases under this heading. Expert testimony from a social worker or psychologist should be able to clarify whether the giant was able to restrain her impulse, and the court is likely to defer to a better informed view where common sense gives little guidance. The expert is in effect translating between scientific knowledge, or at any rate a position backed up by some thought and evidence, and what ordinary folk are likely to say given an ordinary description. Other cases are more worrying.

Consider a 'normal' person under the influence of drugs or alcohol, which we can take to be administered surreptitiously so that the person bears no responsibility for being in a diminished state. The charge is assault or rape, and the defence is that this person could not be expected to resist the provocation involved. Some people would have resisted it well, so the defence has to be that the offender is a somewhat impulsive person of strong instantaneous feelings. The hope is not acquittal but a verdict of guilt followed by a reduced sentence. For this defence to succeed the person's character and history has to be presented in just the right way. So we can expect character witnesses and we can expect the defence to struggle to introduce anecdotal material from the mouths of suitable witnesses. All going well, though perhaps unjustly, the result will be an impression of the person as basically non-malevolent but prone not to weigh all the options and their consequences.

This was a more problematic case, and it really is. It is easy to present a personality under a variety of aspects, because we are complicated inconsistent creatures who respond more to circumstances than we like.

testamentary capacity to think, and as a result a little careful selection from the wide-ranging precedents in anyone's history will produce whatever impression is convenient. A rich old woman falls under the influence of a charming but unscrupulous young man, who in her eyes makes her see life completely differently, and in the eyes of her family deludes her about her attachments and values. She dies, leaving him a fortune. The will

¹⁹ Murder is the topic for most of the examples. Perhaps this is because the crime is particularly abhorrent and the penalties particularly extreme. Insanity defences against parking tickets are rare.

is contested, on the grounds that she was subject to influences that made her incapable of understanding her actions, especially given her advanced age. This is a rough civil law analogue of the irresistible impulse defence to a criminal charge. The young man's lawyer is likely to produce witnesses to show that the older woman had entered into serious argumentative conversations, so she was still thinking for herself, and that she had made responsible decisions during the relevant time period. The family's lawyer is likely to produce witnesses that she was unreasonably indulgent of her lover and had made gifts that no responsible person would have accepted. And that he had boasted that he was onto a good thing.

It could work out either way. The important point is that the judge, as it most likely will be, will have to get a grasp of what it was like for this aging woman to feel overwhelming affection later in her life than she would have thought possible. And a grasp of what the young man's motives might have been, and whether there was a deliberate plan to diminish her capacity for judgement or alienate her from her family. I do not envy the judge's task. Both sides are likely to have intuitive appeal, and there is no expert opinion available.²⁰

the reasonable person

According to an indispensable legal device, there is a normal standard of rationality which adequately functioning citizens can live up to without superhuman effort and which it is in the public interest for everyone to hold others to. (I suspect it is also widely supposed that this, whatever it is, is a normal mode of operation for most people, and deviations from it are a matter of momentary impulse, bad character, or incompetence.) It involves seeing consequences of actions and developments and understanding the implications of information received. Attributions of negligence as well as many other determinations depend on this. But of course no one sees or understands *everything* that has a necessary connection with what they know. A superhuman genius would just take it as obvious that if you use your credit card four times in this one store this particular morning, then the Yuan will have a catastrophic slide three days later. And in a society of such creatures, not to take this into account would be considered highly irresponsible. (If, with such an elevated grasp of things, they used concepts at all like this.) But not among twenty-first century humans with our history and our limits.

It is a puzzling though indispensable concept. It becomes even more puzzling when we consider limitations of reasoning and imagination in even intelligent people, let alone the whole human race. (There is a philosophically problematic 'should' attached to reasoning, perhaps built into our culture from days when we thought that our species

20 If the young man was- canny he would have ensured that the will was made in his absence but in the presence of a lawyer, ad, ideally, that this was videotaped. I do not have the impression that the failure to do this would fatally undermine the case for leaving the will in force.

was just a couple of steps down from divinity rather than a couple of steps up from other primates, to use a very tendentious sense of up and down. But it is worth noting that a concept can be objective although it is relative.²¹) I gather that juries are not supposed to judge what a reasonable person would think or do by reflecting on their own reactions. (Judges are perhaps taken to be nearer the postulated target in this respect!) Instead, they are supposed to apply the standards that people normally expect other people to meet. As Holmes puts it:

If, for instance, a man is born hasty and awkward, is always having accidents and hurting himself or his neighbors, no doubt his congenital defects will be allowed for in the courts of Heaven, but his slips are no less troublesome to his neighbors than if they sprang from guilty neglect. His neighbors accordingly require him, at his proper peril, to come up to their standard, and the courts which they establish decline to take his personal equation into account.^[22] The rule that the law does, in general, determine liability by blameworthiness, is subject to the limitation that minute differences of character are not allowed for. The law considers, in other words, what would be blameworthy in the average man, the man of ordinary intelligence and prudence, and determines liability by that.²³

But there is no codification of what these standards are. No one has ever made precise sense of what ideal rationality would consist in, or how near to ideal rationality an acceptably imperfect person would expect other acceptably imperfect people to be, let alone which deviations from such a standard would be irresponsible dangers to the public interest. But we do in some improvised, shifting way hold people to standards of thinking about the consequences of their actions. And this is reflected in the legal reliance on the concept of the reasonable person. Judges and juries are often required to consider what they would expect a reasonable person to have done in the position of some participant in a trial. Though it is not usually put this way, my take on the issue is that juries are supposed to *imagine* what people would on reflection expect or find startling in other people's actions and considerations, on the assumption that the other person knows the law and is trying conscientiously to abide by it. (This raises the troubling possibility that someone might suffer for being too farsighted and too subtle to be easily intelligible to the mass of humanity, especially if she happened to have arrived at a false conclusion in this particular case, as can happen to anyone whatever their skills.²⁴) If this is right, the

21 What is right and left given a direction is objective on the surface of the earth because there is a privileged sense of 'us'; in space it needs to be relative to both a direction and to an up/down axis.

22 'Personal equation'. Not relevant for present purposes but curious: the term comes from the astronomy of his time, where the timing and accuracy of a particular person's use of a telescope was noted and included in calculations based on their observations.

23 Holmes (1882), Lecture 3.

24 The only real case at all like this that occurs to me is that of the great logician Kurt Goedel, whose claim to have discovered an inconsistency in the U.S. constitution almost scuppered his

judge or juror is asked to do something very delicate: to imagine not what they would expect of themselves but what they would expect of someone else, or even what they would expect a well informed and reflective other person to expect another well-informed affected further person to do. They have to imagine what other people would imagine under rather special circumstances that are not found in their regular lives.

It is (somewhat) supportive of this reading that the law also uses conceptions of what a reasonable child of a given age, or a reasonable person of a particular degree of mental handicap, or for that matter a reasonable expert, would consider and do. For these can be paraphrased as what people would as a matter of fact normally expect of others in these categories. If someone in the category concerned understands what the law expects and is trying their best to live up to it, how will they usually think and act? Understanding the concept this way makes it somewhat clear what its extension is – what the characteristics of someone it applies to are. But it does not tell us how anyone, from experienced judges to jurors thrown into the situation, can reliably tell when it applies in particular cases.²⁵

It is a delicate and tricky concept, but what are the deeper roots of its slipperiness? There is a connection, at any rate, with the contrast between the imaginative capacities nurtured by everyday life and those based on largely unintuitive scientific evidence. When a jury is asked to consider whether someone should have been expected to have paid attention to certain considerations, all they can do is to imagine themselves into the situation as a dispassionate observer and note their imagined reactions when the person ignored the considerations. If their imagined reaction would be that something was amiss and the person would not have been acting with care, then they are distinguishing that person from the hypothetical standard. The question is not whether the juror would personally ignore the considerations, since she may know that she is personally hasty and sloppy – a negligence charge waiting to happen. It is not whether on imagining herself observing the situation of the person in question she imagines approval or acceptance. For she may know that she would, in fact, be too permissive or too condemning. And it is not whether given just the conclusion the accused person arrived at she would consider it reasonable, careful, dutiful or whatever, since that amounts just to the full question of whether it is a matter of negligence (or whatever). More probing than this into possible thought processes is required. The jury is proposing to itself a hypothetical person who operated in the reported way and reacting through the eyes of a hypothetical thoughtful human observer. I do not see how they can do this except by

American citizenship hearing. Luckily he had brought along his best friend, Albert Einstein, who was able to defuse things.

25 See King (2017) for a better informed view about why it is not realistic to expect jurors to simulate the reactions that a reasonable person would have in given circumstances.

imagining they are faced with such a person and asking what they would conclude of her. And the only plausible way I can come up with for them to do *this* is by imagining a constrained imagination of the person. The jurors ask themselves what conclusion someone of a vaguely specified kind would form about someone else committed to satisfying certain norms and performing the act in question. The question is whether they can succeed.

It is a hard task – hard both to follow the instructions and to do so in such a way that gets an accurate answer, one that does say whether a legally conscientious person could have operated in this way. It is hard not to be cynical about how well typical juries can do it. It will depend a lot on how the jury is instructed, and thus on the skill of the judge in setting the jury to do this psychologically delicate thing rather than any of the other tasks they might suppose. I would guess that judges often approach their task by considering how they would tackle the jury's assignment were they on it, and then turning their reflection on this into instructions.²⁶ We might then ask how often judges and juries negotiate all these obstacles successfully. But how would we even begin to answer this? Success is rarely given by a straightforward fact such as whether or not a defendant performed some act. More often it is a question of whether someone *should* have been more careful, or the like, presupposing something close to the very concept concerned.

Everyday social/psychological capacity combined with guided introspection is thus a very fragile tool here. (The connections between capacities for social life and capacities for attributing states of mind are controversial. On the consensus view social capacities draw on independent capacities to attribute states of mind. A more radical view, to which I subscribe, puts the needs of social life at the heart of our mind-to-mind thinking. This would further encourage scepticism about our accuracy in imagined cases, especially if not mediated by an imagined social or practical connection.) But scientific psychology is no more promising. There are no experts and there is no experimental data on what people of normal intelligence trying hard to abide by the law and their responsibilities will do. How could there be? The selection of experimental subjects would depend on the result of the inquiry, what a reasonable person would do.

In the previous section I invited scepticism about the accuracy of our mind-ascribing capacities. In this section I have suggested that legal procedures give a lot of ground for such scepticism.

²⁶ As far as I know expert witnesses are never brought in at this stage. And in fact no existing expertise would fit it. Perhaps the reactions of a judge with a knowledge of the precedents and experience in managing cases where the distinction is relevant are the best kind of expertise. But there is a potential speciality here, for someone with a background combining philosophy of mind, psychology, and law. If only there were people with the right combinations of skills and if only courts had the remotest inclination to make use of them.

problems with imagination

We tend to trust our imagination of people when we can identify with them, or otherwise think of them as people rather like us, and we trust it much less with people we find less familiar. Part of this is a matter of moving the goalposts: when we have been in a situation we often think that reactions that came easily to us do not need much explanation; when the situation is outside our experience we find it harder to grasp. This is largely a matter of being able to summon corroborating details in familiar situations, and knowing what details to look for to establish a plausible conclusion.

In recent years in English-speaking countries, and probably elsewhere, there has been a spate of episodes in which middle-aged male judges have showed remarkable naivety in discussing the motives of men accused of rape (and other similar offences) and their accusers.²⁷ In a typical case the judge criticises the accuser for foolishly allowing the man opportunity or expectation, and expresses some sympathy with the man for succumbing to what would naturally have seemed like an invitation (or so the judge thinks). So the judge ends up by saying either 'under these conditions, it wasn't a crime' or 'well, technically you broke the law, but your situation was so understandable that I'm giving you a minimal sentence.'

The judge in these cases is citing what seems to him the easy imaginability of a relatively innocent or ordinary motive as a mitigating factor. In effect he is saying 'if you are to be punished severely then so should almost everyone'. There is obviously a big topic here on imagining ordinary motives as mitigating factors in cases where there is judicial latitude in sentencing (which I realise varies a lot from one jurisdiction to another). I am not going to discuss it because it requires knowledge I do not have. It also requires a sophistication about the philosophy and psychology of motive attribution, for which I am only a little better equipped. Quite likely there is no single person who is perfectly equipped. This is an important topic waiting for an interdisciplinary approach.

But the imagination is asymmetrical. The judge imagines the man's motives and not the woman's. It is not plausible that many of the self-aware, sophisticated, and responsible people who become judges are simply acting from crude prejudice. I suspect that something less obvious is going on, and that it connects with the failures of imagination I mentioned just above. Sexual attraction is a mysterious thing in cases where you are not yourself subject to it. It is hard for exclusively straight people to imagine same-sex

²⁷ The examples I have particularly in mind are well-publicized and controversial lenient sentences in rape cases in Calgary Alberta in 2011, in Stanford California in 2016, and in Winnipeg Manitoba in 2016. But there is no shortage of cases where middle-aged male judges have suggested victims could and should have prevented the events by dressing differently and physically resisting more.

attraction, and I take it that exclusively gay people think of other-sex attraction as just one of those peculiar things that happens with much of the human race. (This is too formal and pretend precise. We can retreat to 'we all find it hard to get a good grasp on who other people fancy'. And add that six-year-olds faced with all adult sexuality just say: yuck.) The reason is that these motives are not continuous with or well integrated with our other motives, at least not with the ones that we can easily simulate in others. So the enterprise of putting yourself in another person's shoes is difficult when the other person is in this respect not like you. In this way it is like neurology or academic psychology rather than the everyday grasp we have of everyday motives.

Can one learn to imagine across such differences? There certainly is an abstract virtue here, a description of a valuable capacity. What people can actually learn, though, is usually more specific. One overcomes an inability to grasp the motives of one kind of person in one kind of situation. If one is interacting with people of this kind on a regular basis then one gains evidence about the accuracy of the procedure. Without this interaction there is very little available from everyday social resources to give assurance about accuracy. The danger, then, is just that the vividness of imagination will make them think that the process has captured the other person's motivation when in fact it has not. The list of *attainable* virtues and how they are acquired is not a matter of obvious common sense, and not given by philosophical reflection, but a matter of hard serious psychology.²⁸ And unfortunately much of the psychology does not exist.

a general pattern

In all of the cases I have discussed there is a tension, to put it mildly, between what we need in order to get a sense of individual motives and what we need in order to get conclusions that will apply equally and accurately to a large range of people. Fairness-as-accuracy versus impartiality. There is a fundamental reason for the tension. Our everyday grasp of other people depends on innate routines, refined but not essentially altered by growing up in particular human cultures, which have evolved to serve us in the human ecological niche of thought-out cooperative activity. We are fairly accurate about whether someone is inclined to help, or likely to be indifferent or aggressive, what information is available to them, and which aims will appeal to them. We are much less good at predicting what specific means people will take to their ends. And we are even less good at matching combinations and compromises between all the potential motives that are lurking in any person with the actions they have in fact performed. (How they could have done it rather than why they did.) Still, the methods we learn as children work fairly well among people set on getting things done together, and one reason is that when we

28 Virtue ethics, from Hursthouse to Swanton via Nussbaum, while capturing important points about the accessibility to humans of their values always runs the risk of running together our terms of praise with the qualities that actually describe our capacities. See Harmon (2009). Imaginative capacities surely can be educated.

operate socially we are able constantly to check and reconsider the meaning of what others are doing, and – a generally neglected point – to fine-tune our side of the cooperation, or even explicitly to intervene, in order to test hypotheses (suspicions, fantasies) of what the other person may be up to.

This applies to much of our knowledge of individuals, and a lot of it takes the form of shallowly conceptualised expectations of what a person will do and how they might react to what one oneself does. (The point is reinforced when we consider the role of informal experimentation – fine-tuning, intervention – in assessing people we are interacting with.) It does not yield many nontrivial generalisations about human action and motivation in general.

Moreover, there are whole areas of human life where any thinking directed at cooperative activity, whether or not based on capacities that we can think of as imaginative, fails. It misleads us about depression and other mental vulnerabilities, social panic and other interactive styles, the long-term effects of damaging experiences, and the ways that a person's history affects their performance. It is tempting to say that these are areas for neurology, or psychology taken in a narrowly scientific sense, rather than every day psychological acuteness. But there is a circularity in saying this: we need further resources here because this is where our inbuilt capacities fail.

So what about the alternative, evidence based psychology and the like? That does not tell you much about why individual people did particular things at particular moments. It focuses on general human motivation and cognition. Perhaps in the future we will know enough that we can feed individual differences into general theories to get predictions and diagnoses about particular people. Perhaps. (I referred above to how work such as that of Loftus (1976⁹) of how false memories can be implanted have had an impact on legal procedures.²⁹) We certainly cannot get such useful information very often at the moment. And if we could very hard issues about tensions between it and long-established routines for assessing motives would arise. The difficulty of reconciling impartiality with accuracy remains.

references

- Card, C (2002) *The Atrocity Paradigm: A Theory of Evil* (New York, Oxford University Press).
- Davidson, D (1982) 'Rational Animals' 36(4) *Dialectica* 317.
- Coplan, A, and Goldie, P (eds) (2011) *Empathy: Philosophical and Psychological Perspectives* (Oxford, Oxford University Press).

²⁹ Psychiatric testimony might be adduced here. One reason I am not considering it is that it does not compete with every day routines of motive attribution. The domains are too different.

- Gopnik, A, and Meltzoff, G (1997) *Words Thoughts and Theories* (Cambridge MA Mass, MIT Press).
- Harman, G (2009) 'Scepticism About Character Traits' 13 *Journal of Ethics* 235.
- Heal, J (1998) 'Co-Cognition and Off-Line Simulation' 13 *Mind and Language* 348.
- Holmes, Jr, OW (1882) *The Common Law* (London, Macmillan).
- Kahneman, D (2013) *Thinking Fast and Slow* (New York, Anchor).
- Kahneman, D, and Tversky, A (1979) 'Prospect Theory: An Analysis of Decision Under Uncertainty' 48 *Econometrica* 263.
- King, M (2017) 'Against Personifying the Reasonable Person' 4(11) *Criminal Law and Philosophy* 725.
- Levesque, R (2006) *The Psychology and Law of Criminal Justice Processes* (New York, Nova).
- Loftus, EF (1979) *Eyewitness Testimony* (Cambridge MA, Harvard University Press).
- Maibom, H (ed) (2017) *The Routledge Handbook of Philosophy of Empathy* London, Routledge
- Morton, A (2002) *The Importance of Being Understood: Folk Psychology as Ethics* (London, Routledge).
- Morton, A (2004) *On Evil* (London, Routledge).
- Morton, A (2007) 'Folk Psychology Does Not Exist' in M Ratcliffe and D Hutto (eds), *Folk Psychology Re-assessed* (Dordrecht, Kluwer/Springer) 211.
- Morton, A (2009) 'Folk Psychology' in B McLaughlin and A Beckermann (eds) *The Oxford Handbook of Philosophy of Mind* 713.
- Nichols, S, and Stich, S (2003) *Mindreading* (Oxford, Oxford University Press).
- Perner, J (1991) *Understanding the Representational Mind* (Cambridge MA, MIT Press).